Ref. Ares(2019)4339302 - 08/07/2019

INBOTS
Inclusive Robotics for a better Society

INBOTS WHITE PAPER

ON

INTERACTIVE ROBOTICS LEGAL, ETHICS & SOCIOECONOMIC ASPECTS

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

www.inbots.eu

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Table of Contents

Versioning and Contribution History

Version	Date	Modified by	Modification reason
v.01			
v.02			
v.03			

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

1. Executive Summary

This deliverable aims at providing an overview of the legal, ethical and socio-economic aspects of 'interactive robotics' - an emerging field where robots are conceived to perform tasks in close proximity with humans, cooperating with them both physically and cognitively. Within the INBOTS consortium, several companies work in different fields of the interactive robotics such as exoskeletons, humanoids, prosthetics, collaborative robots, etc. This document is intended as a general review of WP2 (i.e. the preliminary report). The next version (i.e. the final white paper of INBOTS) will go further in detail and will finalise the analysis.

The INBOTS consortium, together with the Project Coordinator, agreed to focus on the following issues for this preliminary report: from the philosophical and legal perspective, should we conceive of robots - including 'smart' robots possessing artificial intelligence - as subjects or agents in their own right? Is this approach sustainable? What alternative approaches exist? What would the consequences of these approaches be for important legal matters such as apportionment of liability and protection of intellectual property? What are the consequences of the growth of interactive robotics for labour laws and the future of work? What elements of tax law and corporate social responsibility should be examined in this context?

Regarding the important issue of agency/subject-hood, we note in this report that if we are to define what a robot can and cannot do by referring to the notions of agency, responsibility and liability, it is first necessary to understand what we mean by these concepts, which have complex and possibly indeterminate meanings. In this report we clarify that for an entity to be deemed an agent, it ought be able to instantiate intentional mental states capable of directly causing performance, and that for it to qualify as a moral agent, it ought to display what is usually referred to as 'strong autonomy', i.e. the ability to decide freely and coordinate one's action towards a chosen end, as well as the moral awareness needed for understanding the moral significance of one's actions. In this regard, we explain in this report that at present robots, conceived to complete a specific task identified by their user, ought not qualify either as agents (absent the consciousness required for them to have intentional mental states), nor as moral agents (given that they have no capacity to engage in moral judgments, and lack strong autonomy). At present robots can determine how to reach the goals they are programmed to achieve, but said goals are still defined by an external agent – most likely, the designer, producer or programmer. Therefore the only moral agents involved in the functioning of the machine remain the humans behind it, who are responsible for both the goals chosen, the model of functioning designed for the robot, as well as the very choice to grant to it a certain degree of autonomy in determining how to perform intended tasks.

Following on from this conclusion, we consider the issue of how liability can be apportioned under the law. We note that positioning robots as 'products' brings questions of product liability to the fore; meanwhile, viewing robots as 'legal subjects' (if not moral subjects) could provide a means, via a type of legal personhood, for considering legal questions of liability of the human agents (designers, programmers, owners) and the apportionment of e.g. damages.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

We further explore the question of how intellectual property (IP) laws – particularly the laws of copyright and patents – have traditionally been premised on the existence of a human author. The existence of co-creation and collaboration with interactive robots thus brings a legal challenge: how to recognise the works and inventions co-created with robots when these are not entirely 'human creations / inventions'. We examine whether legal reform is required to clarify questions of IP authorship and ownership.

On robotics and the labour market, we consider four key trends that impact the future of work: technological progress and automation; international trade and urbanization; a rising diversity of work forms; and population aging. We reflect upon the importance of automation for the long-term economic prospects of the EU and the role that interactive robotics will play in this regard, in an era when 'new work forms' such as the 'gig economy' are becoming highly prevalent.

Furthermore, with regard to the labour relationship, we examine labour Law and the possibility that an ('intelligent') robot could be considered a 'worker'; and we further consider the principle of equality and non-discrimination, i.e. the real and effective equality of people, including the groups of people who are most vulnerable to losing their jobs to automation or robotization. On the basis of this first reflection, we consider the adoption of labour legislation as it applies to workers' rights and obligations both from the perspective of guaranteeing people's employability and of guaranteeing the rights of workers in their interaction with robots in the workplace.

Regarding financial and tax Law, we note that States will have to allocate rights and responsibilities among human beings for the actions of non-human beings, and fight inter-personal and inter-national inequality through strengthened cooperation, e.g. investing in re-skilling and/or via a basic income, now at an experimental phase.

We present a philosophical orientation and analysis of the ethical issues, the positions and the arguments in the field, with an outlook on policy. Robotic artefacts are mediations arising from a certain socio-cultural context and which open up new possibilities in human functioning. The introduction of robotic devices in domestic environments or institutional ones (schools, hospitals, etc.) must be done in such a way that the overall dimensions of the care offered to create, consolidate and support personal autonomy contribute to good development of human beings in all their different life stages, avoiding a 'robotic divide'. We focus on clarification of key concepts and values: 'inclusive robotics' in relation to the Convention on the Rights of Persons with Disabilities -particularly caring for people with intellectual and developmental disabilities, and in relation to society as a whole. Societal inclusion (for instance of people in elderly care or rehabilitation) relates to the general human existential situation of vulnerability and finitude, seen as universal human feature which also enables creativity, arts, and joy of life. We pay attention to theoretical frameworks such as the capability approach, facilitating autonomy in the achievement of human functioning in inclusive environments. Additionally, we stress the importance of the Convention on the Rights of the Child in order to safeguard the greater interests of the minor.

Furthermore, we note that if environmental values are built into the economies, markets, institutions and practices relating to interactive robotics, the environmental harms can be minimised, and interactive robotics can serve as a model for other novel technologies.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Finally, we note that corporate directors and company managers will play an important role -they will need to work on advancing and reinforcing corporate social responsibility in the governance and administration of their organisations. The support for gender diversity and the observance of women's rights in the workplace are essential aspects that inform the concept of this corporate social responsibility that must be present in the approach to the new challenges posed by interactive robotics and artificial intelligence today and in the future.

Notwithstanding the need to regulate the future impacts of robotics and artificial intelligence, the asymmetries in knowledge and the awareness of its impacts among society, experts, government authorities and industry; the objective of not reducing the innovation and competitiveness of a sector that is still incipient in many States; the lack of flexibility and agility of regulatory mechanisms and processes; and the different capacity of action of the affected interest groups, have led to a wide body of statements by experts and industry sectors, which represent a first step towards the establishment of agreed ethical principles and minimum guidelines for the future regulation of the sector.

Some voluntary initiatives have arisen so far, and we can find some basic principles of a future due diligence system that could be articulated by the private sector as complementary to the regulation.

The measures of transparency established in these statements should contribute to giving visibility not only to the risks but also to the sequencing of responses to them, also facilitating the establishment of monitoring and follow-up systems, making changes in the way of proposing some corporate activities; increasing the awareness of workers and users about particular risks and increasing the supervision of the media and other sources of information on human rights, among others.

On the other hand, the application of the principles of prevention, mitigation, and protection is relevant at a time when the principles and process contained in the Guiding Principles of Business and Human Rights are beginning to be widely understood by companies, also helped for the development of national business and human rights frameworks. But together with the establishment of due diligence processes by public and private actors, the involvement of fundamental rights invokes the role of the State in the establishment of economic and institutional incentives (for example, in public procurement mechanisms that encourage the use of responsible technologies) for systems that respect human rights in the deployment of artificial intelligence and robotics. Along with this, the establishment of certification programs or labels on responsible artificial intelligence and robotics, taking as background the programs provided in the European Commission Regulation on the protection of personal data. Companies or organizations that deserve this recognition based on objective criteria could benefit from a very significant and reassuring competitive advantage for people (consumers, customers, employees, service users, etc.).

The WP2 team has tested most of these ideas at the European Robotics Forum 2019, through the organization of two Workshops: one on Ethics and Corporate Social Responsibility for Inclusive Robotics; and the other on Sustainable Public Policies for Innovation and the Future of Work. The results of the debate held with the robotics community are here offered as an annex.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

2. Clarifying liability, risk management and insurance issues

In order to define new approaches to interactive robots' liability, insurance and risk-management, a first overview of the legal, economic and ethical profiles of such technologies is provided in this sub-section.

1. Agency, moral responsibility and liability: an analytical inquiry

a. Introduction

It has been claimed that some robotics and artificial intelligence applications (henceforth, R&AI) are so technologically advanced, that they require "a systemic change to laws or legal institutions in order to preserve or rebalance established values"[1]; in other words, it is argued that R&AI actions are essentially beyond human control, and as such, we should deem the R&AI as responsible for the wrong caused, instead of blaming the producer, the owner or the user[2].

The above thesis displays a series of flaws - and over the course of this document it will be proved both theoretically incorrect and practically inadequate.

Rather, in this document we argue that legal reforms can be grounded in two different approaches.

The first is the "ontological" or "essentialist" perspective - entities have a clear-cut legal qualification based on their inherent features, which in turns determines the applicable legal rules. In this view**, we may need to adopt new rules, or change existing ones, when the object of regulation** (in this case, R&AI applications) **is so different** from what we have been regulating so far (other, less advanced forms of technology), that a distinct legal qualification is due.

The second represents a functionalist point of view: **legal frameworks shall be developed according to their adequacy in performing the functions** attributed to them, as well as the broader consequences deriving therefrom[3].

The aforementioned claim – that R&AI should be held morally and legally responsible for any wrong caused – is often grounded on a combination of functional and ontological perspectives, but it lacks sufficient analytical clarity,

---

[1] Calo, Ryan. "Robotics and the Lessons of Cyberlaw." California Law Review (2015): 513; Christophe Leroux et al., "Suggestion for a Green Paper on Legal Issues in Robotics. Contribution to Deliverable D.3.2.1 on Els Issues in Robotics," (2012); ibid.; Luciano Floridi and J.W. Sanders, "On the Morality of Artificial Agents," *Minds and Machine* 14 (2004).
[2] Santosuosso at al., "Suggestion for a green paper on legal issues in robotics."
[3] Bertolini, Andrea. "Robots as products: the case for a realistic analysis of robotic applications and liability rules." Law, Innovation and Technology 5.2 (2013): 214; Bertolini, Andrea. Robots and liability: justifying a change in perspective. Rethinking responsibility in science and technology. (Pisa University Press, 2014. 143; Schulze, Reiner, and Dirk Staudenmayer, eds. *Digital revolution: Challenges for contract law in practice*. Nomos, 2016)

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

and thus needs to be specified. Indeed, arguing an identical conclusion in terms of policy recommendation, but based on one or the other of the perspectives, alternatively bears radically different theoretical and practical consequences. If the robot is to be deemed a *subject* – not an object – thence not only his or her liability will follow but also a complex bundle of rights and obligations intended to protect his or her own interest. Instead, if the robot is treated as a juridical person, with the sole aim of segregating selected assets, shielding single human beings from the legal and economic consequences of its operations, and eventually providing a diversified taxation scheme, then the overall legal – and ethical the like – implications radically differ, and the two theoretical stances (ontological and functionalist) would not be confused.

Yet, the lack of clarity with respect to the choice each single author makes among these competing perspectives causes the arguments to become obscure, thence hard to affirm or contrast with the adequate degree of scientific precision that a legal or ethical debate would demand in any other matter.

Indeed, the idea that we shall avoid the so called "responsibility gap"[4], where humans are called to respond for damages upon which they have no or very limited control, and that machines shall behave as responsibly as possible, according to the principles elaborated through "machine ethics"[5] is often expressly grounded on the belief that the peculiar features displayed by advanced R&AI – their asserted autonomy and ability to modify themselves –, make them agents, and, more specifically, moral and legal agents.

However, neither the functional nor the ontological assumptions are sufficient to ground this exceptionalist claim - and the proposal for a reform of liability rules associated with it. On the one hand, **it is disputable whether holding R&AI directly accountable for the damages caused is preferable, everything considered, to holding the humans behind them liable[6]. On the other hand, the ontological claim according to which new robots' essential qualities make them subjects, rather than mere objects, is far from being proved.**

As already mentioned, in the current debate R&AI are said to be autonomous *agents*, and since autonomous agents are considered moral and legal subjects, it is claimed that they shall be held *morally responsible* and *legally liable* for their actions, for example by making them artificial legal persons.

Nonetheless, if we are to define what a robot can and cannot do by referring to the notions of *agency*, *responsibility* and *liability*, it is first necessary to understand what we mean by these concepts, which have complex and possibly indeterminate meanings.

Indeed, when discussing the challenges and opportunities brought about by R&AI, both economic, legal, ethical, philosophical and engineering considerations come into play,

---

[4] Andreas Matthias, "The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata," *Ethics and Information Technology* 6 (2004).
[5] Fossa, Fabio. "Artificial moral agents: moral mentors or sensible tools?." Ethics and Information Technology 20.2 (2018): 115
[6] For social and ethical considerations on this point: Joanna J Bryson and Philip P. Kime, "Just an Artifact : Why Machines Are Perceived as Moral Agents," in *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence: Barcelona, Catalonia, Spain, 16–22 July 2011*, ed. Toby Walsh (Menlo Park, CA, USA AAAI Press, 2011).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

leading the debate to merge the methodological and analytical background of heterogeneous disciplines. Yet, economists, engineers, philosophers and lawyers may use terms which have both a common, a-technical understanding, and one which is peculiar of their own subject. Therefore, engineers or lawyers may speak of autonomy to denote different qualities than the ones that philosophers understand as associated with said notion. This constitutes a case of semantic ambiguity:[7] both the meaning of a concept and the conditions of its use depend on the peculiar contexts in which the latter is used.

As highlighted by the studies on legal reasoning and linguistic indeterminacy[8], unclear and under-specified terminology may undermine the acceptability of the warranties used to back a specific argument, which in turn affects the correctness of the overall claim. **For any debate on the social and legal implications of robotics' uptake to be framed correctly, a preliminary account of the notions of agency, responsibility and liability shall thus be offered.**

In the following pages, each concept will be discussed, first by offering a general analysis – through insights from philosophy and legal theory – and then by considering how they shall be declined when R&AI applications are involved (§1.2, §1.3). Although agency, responsibility and liability are traditionally said to be strongly interconnected, focusing on one concept at a time enables us to identify the essence, as well as the relationships among them. Also, it allows us to understand whether they shall be framed in terms of a bi-functional determination, or whether one notion is constitutively independent from the others.

Finally, we will use the insights and conclusions developed in the first part of the inquiry to offer a comprehensive analysis of the legal status of robotics (§1.4); here, a functional, instead of an ontological perspective will also be taken into consideration, to further develop the results achieved so far.

b. Definition of "Agency"

i. Philosophical notions of "Agency"[9]

From a philosophical perspective, agents are traditionally defined as subjects who have the capacity to act, i.e. to perform action, while agency denotes the manifestation of such capacity[10].

However, "actions are doings, but not every doing is an action"[11]: according to (the main variations of) the standard conception, an event may be deemed as an action only if

---

[7] Waldron, Jeremy. "Vagueness in law and language: Some philosophical issues." *Cal L. Rev.* 82 (1994): 509; On the different forms of linguistic indeterminacy and their implications on the legal discourse, see also Thomas Endicott, "Law and Language," in *The Stanford Encyclopedia of Philosophy*, ed. Edward Zalta (2016)., https:plato.stanford.eduarchivessum2016entrieslaw-language; *Vagueness in Law* (Oxford: Oxford University Press, 2000).,

[8] Stephen Toulmin, *The Uses of Argument* (Cambridge: Cambridge University Press, 1958). Endicott, *Vagueness in Law*; Stephen E. Toulmin, *The Uses of Argument*, vol. 34 (Cambridge University Press, 1958).

[9] This issue is more deeply analysed by philosophers in section 5.

[10] For a comprehensive account of the philosophical conceptions and theories on this topic, see Markus Schlosser, "Agency," in *The Stanford Encyclopedia of Philosophy (Fall 2015 Edition)* (Edward N. Zalta). https://plato.stanford.edu/archives/fall2015/entries/agency.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

brought about intentionally[12], thus not being mere the result of causal determinations among naturalistic events.

Whatever the case, we need to understand under which conditions an entity can be said to act intentionally.

Intentionality is often defined as the "the determination of a specified end that implies the necessity of actions of a specified kind"[13], resulting from a combination of volition and rational interference.

In particular, according to some authors, the kind of rationality required consists in being capable of rationally justify one's actions in reference to determined and determinable purposes, which, in turn, requires deliberative and argumentative skills that only human beings possess, in part because of their linguistic abilities. Under this view, only humans can perform actions, being able to reason and decide intentionally[14].

Other theories set a lower threshold, describing intentionality as a *mental state* – such as belief, desire, will – that does not necessarily entail the rationality humans display, and also covers the spontaneous initiation of actions which do not follow rationally justifiable desires[15].

Some authors have instead theorized a "minimal agency" which qualifies as "agent" any unified entity that is distinguishable from its environment and that is doing something by itself according to certain goals: according to this view, very simple organisms can be said to have the intrinsic goal of continuing their existence, even if they lack the ability to rationally elaborate and justify their aims and actions[16]. However, by putting into brackets any reference to mental states, this account merges the very distinction between naturalistically caused behaviours and actions, and – at least for the purpose of this paper – shall thus be refused.

Indeed, it seems more sensible to accept the medium approach, which implies that "X is an agent if and only if X can instantiate intentional mental states capable of directly

---

[11] Himma, Kenneth Einar. "Artificial agency, consciousness, and the criteria for moral agency: What properties must an artificial agent have to be a moral agent?." Ethics and Information Technology 11.1 (2009): 19

[12] G.E.M. Anscombe, *Intention* (Oxford: Basil Blackwell, 1957)., D. Davidson, "Actions, Reasons, and Causes," (1963).

[13] Mathias G. et al., "Action and Autonomy: A Hidden Dilemma in Artificial Autonomous Systems," in *Robo- and Informationethics. Some Fundamentals*, ed. Michael Decker and Mathias Gutman (Lit Verlag, 2012). The rational and volitional elements of intentionality as a way of performing one's action is also the base of the seminal Anscombe, *Intention*.; Davidson, "Actions, Reasons, and Causes."

[14] C. Taylor, "What Is Human Agency?," in *He Self: Psychological and Philosophical Issues*, ed. T Michel (Oxford: Blackwell, 1977).; H. Frankfurt, 1971, "Freedom of the Will and the Concept of a Person," *Journal of Philosophy* 68, no. 1 (1971). The same position is held by Gutman, Rathgeber, and Syed, "Action and Autonomy."

[15] C. Ginet, *On Action* (Cambridge: Cambridge University Press, 1990).

[16] X.E. Barandiaran, et al., "Defining Agency: Individuality, Normativity, Asymmetry, and Spatio-Temporality in Action," *Adaptive Behavior* 17, 5 (2009).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

causing performance"[17]. Even if the very essence of mental states is difficult to grasp, still, it seems reasonable to consider them as requiring some sort of capacity of introspection[18]: in this sense, rather than fully fledged, human-like rationality, agency would presuppose *consciousness*, as only conscious beings can have intentional mental states.

What we have said so far is important for two reasons. Firstly, it helps us sketch a basic definition of agency, and understand that the latter constitutes a more basic notion than other compound concepts, such as those of rational and autonomous agency[19]. Secondly, it leads us to affirm that, until R&AI can be said to be conscious (assuming that this would ever be possible), they could not be considered as agents, but rather as mere objects.

It is worth noting, however, that some authors suggest a radically alternative approach: given the difficulty in understanding whether any entity, other than ourselves, is conscious or not, they claim that we should imagine R&AI as having such capacity, unless proved otherwise.[20] This argument takes into account a difficult epistemological question, and it would go beyond the scope of the present paper to engage in said analysis. In contrast, we may say that the more demanding conditions for intentionality – such as those connected with the autonomy of action and its rationality – despite setting a narrower account of agency, are probably the most relevant threshold against which to test R&AI essential features, as they could work as a proxy for the assessment of the conscious nature of the action.

Indeed, the idea of intentionality certainly goes towards (without necessarily overlapping) that of autonomy. One entity is said to act autonomously, when its actions are (i) free, lacking determination, and (ii) are means to achieve ends which are set *by the subject himself*[21]. Condition (i) sets the standards that we have already discussed, namely, that an action is to be contrasted to a mere behaviour, a deterministically caused event that was not brought about intentionally[22]. What differentiates the notions of intentionality and autonomy is that the latter puts major importance on the origin of the goals for which the actions are performed. Defining an entity as an autonomous agent – instead of a mere agent – implies that the former has acted in order to obtain *its own* goals.

If robots are not autonomous, then, we might assume that they are not conscious, and cannot act intentionally. In the case of software that is interacting with a human being, even if deprived of a physical body, could be deemed a robot, because in a social science perspective the presence of a physical body is but one of the element of analysis that may or not be of relevance[23]. **How a robot which is not autonomous (capable of identifying its preferences and set its goals and pursue them) could be deemed conscious?** If it falls in the realm of simulation (thence the ability of the machine to

[17] Himma, Kenneth Einar. "Artificial agency, consciousness, and the criteria for moral agency: What properties must an artificial agent have to be a moral agent?." Ethics and Information Technology 11.1 (2009): 19
[18] Ibid.
[19] Ibid.
[20] Ibid.
[21] Gutman, Rathgeber, and Syed, "Action and Autonomy."
[22] Ibid.
[23] Bertolini, Andrea. "Robots as products: the case for a realistic analysis of robotic applications and liability rules." Law, Innovation and Technology 5.2 (2013): 214

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

mimic or give the appearance of consciousness) then it is a mere issue of user's deception[24].

As we will see in the following sections (§ 1.4), this specification is of crucial importance also because, despite the variety of discourses which are made on the topic, the statement that R&AI applications should qualify as agents – and thus be held morally and legally responsible – is based precisely on the (not always explicit) assumption that they are not mere agents, but rather *autonomous agents*, possessing free will.

ii. Legal notions of "Agency"

In the legal world, the term agency has its own peculiar meanings.

In a broad sense, being an agent equates to having "legal capacity", whereas a narrower version of this notion merely covers the "legal capacity to act".

Some entity has general legal capacity if it may be entitled with rights and duties (as well as other types of legal situations). According to the modern western legal tradition, each person has said status, at least from the moment of birth, being banned from forms of *capitis deminutio*, such as those related to slavery in ancient Rome or to political and racial prosecution of Jews in the Nazi regime[25]. However, legal capacity is not an exclusive feature of human beings: non-human entities – such as corporations and associations – may be granted general legal capacity, thus being capable to bear those rights and duties which do not require the holder to be a human being (thus excluding, e.g., those arising from marriage).

Despite having legal capacity, legal subjects may still lack the legal capacity to act, i.e. the ability to autonomously modify one's rights and duties by performing legal acts.[26]

In order to be correctly understood, such a notion shall be complemented with a taxonomy of legally relevant facts and acts, which characterizes the European continental legal tradition:

- By *facts* we denote naturalistically caused events or human behaviours producing specific legal effects, where – if having human origin – it is immaterial whether they were brought about intentionally or not;
- By *legal acts* we define intentional actions which the law considers as the basis for the production of given legal effects. Among the latter, we could further distinguish among:
    - *mere acts*, where the action itself is intentional, but the legal effects are produced regardless of whether the author intended to bring about such legal consequences or not;

[24] Bertolini, Andrea. "Human-Robot Interaction and Deception." Osservatorio del diritto civile e commerciale 7.2 (2018): 645
[25] In Italy, for example, natural persons acquire legal capacity with birth (art. 1 of the Italian Civil Code), and no one can be deprived of it for political reasons (art. 22 Italian Constitution).
[26] According to our previous example (the Italian legal system), one subject acquires the capacity to act when he or she become of age – turns 18 years old – (art. 2 Italian Civil Code) and can be limited or revoked by the courts, for example through interdiction, i.e. by depriving the person of the right to handle his or her own affairs because of mental incapacity (artt. 414 ff. Italian Civil Code).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

- *juridical acts*, which produce their peculiar legal effects only if the action was performed intentionally as a means to achieve specific consequences; said otherwise, the production of legal effect is not a mere by-product of the action, by rather the reason why the latter was undertaken.

What has been said so far does not mean that the actions of those who lack the legal capacity to act have no legal effect, or that they do not have the power to perform legal actions at all. On the contrary, any entity – even non-human entities – may cause events, for which the law sets specific legal consequences, despite no legal capacity being required therefor. For a person to perform mere acts, it is necessary to have what is called as "natural capacity", i.e. having the ability to understand the meaning and consequences of one's own actions, and to act accordingly. For example, if a 17-year old boy or girl, having full intellectual capacity, caused a physical damage to another person with fault or malice, s/he would still be liable for the wrong caused (even though, under certain conditions, his or her parents would be called to respond for the damage as well). On the contrary, full legal capacity is required for entering into a valid contract or performing other juridical acts. If we assume that the same under-age boy or girl may be a real-estate owner and wanted to sell a property, despite having legal capacity (as far as the ability to be entitled with property rights is concerned), s/he would lack the power to enter into a legally valid contract,[27] and need someone else acting on their behalf, namely an agent. Which leads us to another point worth of discussion.

In a more specific sense, the term "agency" also refers to that institution, or rather set of norms, allowing and regulating the fiduciary relationship whereby a subject – the agent – is expressly or implicitly authorized to act on behalf of another subject – the principal – to create legal relations between the latter and third parties. Thus, an agent who acts within the scope of authority conferred by his or her principal – or so long as a third party in good faith may legitimately believe him or her to do so – binds the principal to the obligations s/he creates *vis-a-vis* third parties. However, for such effects to be produced, it is not necessary for the agent to have legal capacity, but only for the principal.[28]

This last feature is particularly important for the purpose of our analysis. From a legal perspective, **the role performed by the machine may resemble that of the agent, who acts towards the end set by the principal, and thus produces effects within the legal sphere of the latter**, being able to choose how to perform the intended task.

This does not entail that machines should be deemed "agents" of human beings: rather, it highlights how the law allows the production of effects on another subject, **who is held responsible for having identified the desired results, regardless of the level of**

[27] Legal capacity is required for a person to conclude a contract; should it be lacking at the moment in which the contract was concluded, then the latter may be annulled: see art. 1425 Italian Civil Code.
[28] See artt. 1389 and 1390 Italian Civil Code. "Since the will of the agent is the important factor in relation with third parties, it is the agent's will which has to be examined in the event of defects impairing the validity of the transaction. Art. 1390 of the civil code states that the contract is voidable if the will of the agent is vitiated. Defect in the principal's consent are immaterial, unless he predetermined any terms (price, for example) of the transaction subsequently concluded by the agent. As to the capacity to exercise rights, it is sufficient that the principal possesses it. Therefore, a principal can avail himself of an agent who simply has natural capacity (art. 1389 civil code)": Guido Alpa and Vincenzo Zeno-Zencovich, *Italian Private Law* (Routeledge-Cavendish, 2007)., 186.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

autonomous agency displayed by the entity who performed the action[29]. Just like a person may be bound to the legal effects produced in his or her legal sphere by the contract signed by a representative, an adult with full legal capacity, who has the maximum autonomy in determining the content of the agreement, there should be no reasons why the same person may not be legally bound by the effects produced by the action of a machine showing a lower degree of autonomy than the one displayed by the human agent.

c. Definition of "Responsibility"

For the purpose of this analysis, we shall not refer to responsibility to denote the moral responsibility of a subject, as defined by the traditional philosophical debate[30]. On the contrary, we will use the concept of liability to denote that specific form of legal responsibility which is connected to the infringement of one's rights or duties. Both these forms of responsibility will be distinguished from that pertaining to the deterministic domain of cause-effect relationships, for which the term causal connection will be used.

i. Moral notions of "Responsibility"

According to the traditional philosophical discourse, moral responsibility is the state which characterizes the subject whose actions are judged as worthy of praise or blame[31].

According to the perspective adopted, moral responsibility may be either merit-based – so that praise or blame would be an appropriate reaction toward the candidate only if s/he deserves such reactions – or consequence-based – so that moral judgement would be appropriate only when they are likely to have a desired effect in the agent's actions and dispositions[32]. In this paper, we will take into consideration the merit-based approach, as the major reactions to morally reprehensible actions take the form of legal sanctions (broadly intended, i.e. considering different form of liabilities)[33], and the consequence-based approach to moral responsibility shall thus be reframed as a peculiar form of functional approach to the ascription of liability, which will be considered in the following section.

In this sense, one's action may be candidate for moral evaluation, only if s/he (i) could exercise control over his or her actions and dispositions, and (ii) was aware of what s/he was bringing about. These are generally referred to as the control and the epistemic conditions[34].

---

[29] Bertolini, Andrea. "Robots as products: the case for a realistic analysis of robotic applications and liability rules." Law, Innovation and Technology 5.2 (2013): 214
[30] For an overview, see Andrew Eshleman, "Moral Responsibility," in *The Stanford Encyclopedia of Philosophy (Winter 2016 Edition),* (T Edward N. Zalta, 2016).
[31] The traditional account of this thesis dates back to Aristotle, *The Nicomachean Ethics*, trans. Terence Irwin (Indianapolis: Hackett Publishing, 1985).
[32] Eshleman, "Moral Responsibility."
[33] Norberto Bobbio, "Sanzione," in *Novissimo Digesto* (Torino: UTET, 1969). For an account of the sanctioning and behaviour-shaping functions of liability rules, see §**¡Error! No se encuentra el origen de la referencia.**.
[34] Eshleman, "Moral Responsibility."

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

For the sake of this argument, we will leave aside the deterministic problems connected to one's ability to control one's actions and dispositions[35], and merely assume that (i) agents have a certain degree of freedom of determination, and that (ii) the practice of holding someone responsible needs no external justification in the face of determinism, since moral responsibility is based on social intrinsic reactive attitudes[36].

That being said, it is necessary to ask ourselves whether a machine could meet the control condition. Again, this question has to be addressed in the light of the peculiar form of "weak autonomy" that current robotics display. **Even in a scenario where the machine learns from the environment, possibly adapting its own functioning as a result of this interaction and learning, the machine cannot be said to be in control of its actions:** even if it is free to determine the way in which to act, its choice is still determined by the need to interactively adjust its functioning to the environment and, on the basis of the available data, plan the most efficient way of performing its tasks. Given that the machine does not have control on **the goals which it is programmed to achieve,** since the latter **are heteronomously set** by the human behind it (most likely, the programmer), it cannot be deemed in control of the end itself[37].

Likewise, artificial moral responsibility could not be recognized because it would still lack the epistemic conditions. In the philosophical debate, the issue of awareness is separated by that of the possible deviancy of the causal chain initiated with one's own actions, which, if anything, shall be traced to the definition of agency, not of moral responsibility[38]. Awareness is rather to be understood as "the interpretive process wherein the individual recognizes that a moral problem exists in a situation or that a moral standard or principle is relevant to some set of circumstances"[39]. One entity's complete and unavoidable lack of moral awareness equals to the impossibility of its moral consideration.

Machines, by definition, lack cognitive skills[40]: therefore, **even if we were to encode moral principles in software** – a version of machine ethics which is highly problematic, first and foremost because of the difficulties encountered in the translating normative statements into strings of commands –[41] **robots would not be able to distinguish**

---

[35] For a general account, see Timothy O'Connor and Christopher Franklin, "Free Will,"ibid. (2018)., Tomis Kapitan, "Free Will Problem," in *The Cambridge Dictionary of Philosophy*, ed. Robert Audi (Cambridge: Cambridge University Press, 1999).
[36] Peter F. Strawson, "Freedom and Resentment," in *Proceedings of the British Academy, Volume 48: 1962*, ed. Gary Watson (Oup Oxford, 1962).
[37] Gutman, Rathgeber, and Syed, "Action and Autonomy.", Bertolini, "Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules."
[38] Schlosser, "Agency."
[39] Reynolds, Scott J. "Moral awareness and ethical predispositions: investigating the role of individual differences in the recognition of moral issues." Journal of Applied Psychology 91.1 (2006): 233
[40] Gutman et al., "Action and autonomy: A hidden dilemma in artificial autonomous systems." Robo-and Informationethics. Some Fundamentals (2012): 231-257.
[41] Indeed, a series of problems arise: the first one, lays in the very definition of the ethical principles to be encoded, upon which disagreement is likely to be found; the second one, is related to ambiguities connected to the use of natural language, which may lead to gaps and incongruences between what the robot is told to do, and what the designer actually intended it to do; the third one, is rather connected to the peculiar functioning of ethical norms, as well as many legal norms, which do not apply once and for all, but may be subject to conflicts, exceptions and balancing, which require processes of prioritization and proportionality assessments, which are far from easy to be pre-defined in a way as to be hard-coded in the machine. For an account of these difficulties, as well as some

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

them from other forms of command, because they would lack the capacity to understand the moral significance of their actions.

On the other hand, machines can certainly perform actions which are, in abstract terms, worthy of reactive moral attitudes; however, since they cannot engage in moral considerations, they will not qualify as moral subjects, and thus may not be attributed moral responsibility[42].

In this sense, it is worth highlighting how the theories which accommodate artificial moral agents are often based on formal definitions and behaviouristic tests that aim at proving that there is no qualitative difference between artificial and human agents. A famous example for this is the thesis offered by Floridi and Sanders, who claim that moral responsibility shall be equated to the ability to cause moral effects, which arises when an entity satisfies the formal criteria of interactivity, autonomy, and adaptability[43].

However, it has been recently demonstrated how such claims shall be read within the perspective of the machine ethics projects, and do not hold absolutely. The theoretical possibility of constructing a theory that is functional to the attribution of moral agency to robots, assimilating robots and humans, does not mean that, in absolute terms, there is no significant difference between the two, nor that there is a pragmatic reason why artificial moral agency shall be constructed[44].

R&AI applications do not share human's autonomy and moral awareness necessary according to an absolute – i.e., non-instrumental or sector-specific – definition of moral agency, as the latter "cannot abstract from the very determination of ultimate ends and values, that is, of what strikes our conscience as worthy of respect and concretization"[45].

ii. Legal notions of "Responsibility": the concept of "Liability"

In legal terms, being liable means to be responsible or answerable for something at law. It rests on the idea that there are specific sources of obligations, which bound one subject to do something, denoted as the object of the obligation.

In criminal matters, liability arises because of a court decision, when the prosecutor demonstrates beyond reasonable doubt that the defendant's conduct meets both the mental and the physical element required for offence to be punished under criminal law, and consists in fines and imprisonment, as well as other non-custodial punishments.

---

methods used to overcome them (whenever possible), see Gaetano Aurelio Lanzarone and Federico Gobbo, "Is Computer Ethics Computable?," in *Conference Proceedings of Ethicomp 2008: Living, Working and Learning Beyond Technology*, ed. Terrell Ward et Al. Bynum (Mantova: Tipografia Commerciale, 2008).

[42] A similar position is held by Himma, "Artificial Agency, Consciousness, and the Criteria for Moral Agency: What Properties Must an Artificial Agent Have to Be a Moral Agent?.", when he correctly notes that all the three capacity of moral agency – rationality, ability to know the difference between right and wrong, and the ability to apply correctly these rules to certain paradigm situation that constitute the meaning of the rule –, and indeed the very concept of agency, requires the agent's consciousness.

[43] Floridi, Luciano, and Jeff W. Sanders. "On the morality of artificial agents." Minds and machines 14.3 (2004): 349

[44] Fossa, Fabio. "Artificial moral agents: moral mentors or sensible tools?." *Ethics and Information Technology* 20.2 (2018): 115

[45] Ibid.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Under western legal tradition, criminal liability has a sanctioning, as well as a re-educative aim[46].

Administrative liability is a type of financial responsibility posed by agents of the public administration for damage, in order to sanction the infringer and compensate for the wrong caused.

Civil liability rules determine who is supposed to bear the negative economic consequences arising from an accident, and under which conditions[47]. Typically, the party that is deemed to have caused the accident is held liable, and thence bound to compensate, and is therefore responsible for it. Liability is established after a trial, where the claimant, who sued the wrongdoer, has to prove the existence of the specific constitutive elements that ground the liability affirmed. Under the English civil law of torts, for example, to hold a person liable for negligence, the claimant needs to prove that the defendant had a duty, that s/he breached it, and that such breach caused an injury, resulting in recoverable damages (e.g. because the harm is not too remote a consequence of the breach)[48].

Civil liability rules pursue three distinct functions, namely: (i) *ex ante* deterrence, since they aim at making the agent refrained from the harmful behaviour, given that s/he will have to internalize the negative consequences caused; (ii) *ex post* compensation of the victim, as they force the person responsible for the damage to make good for the loss suffered; (iii) and *ex post* punishment, since the compensatory award also constitutes a sanction, making sure that the infringer does not get away with the illicit behaviour.

Many different theories have been elaborated to justify civil liability, as well as to shape liability rules within a legal system according to specific ideologies; most of them are related to different notion of justice. According to a retributive account of justice, the blameworthy deserve to suffer, because of the socially reprehensible character of their conduct, and liability rules shall be framed to serve as sanctions[49]. Theories of corrective justice, instead, understand tort law as a system of second order duties, setting obligations to make good the wrong caused by the breach of first-order duties[50]; under this view, liability rules shall rather be elaborated and interpreted so as to assure that the victim is put, as much as possible, in the position s/he would be, had the damage not occurred. Thus, for a loss to be wrongful and worthy of being compensated, it needs to

---

46 See art. 27 Italian Constitution: "Le pene (…) devono tendere alla rieducazione del condannato".
47 Similarly, liability means «the law determining when the victim of an accident is entitled to recover losses from the injurer». See Steven Shavell, "Liability for Accidents," in *Handbook of Law and Economics*, ed. A. Mitchell Polinsky and Steven Shavell (Amsterdam: Elsevier, 2007).
48 Walter Van Gerven, Jeremy Lever, and Pierre Larouche, *Tort Law* (Oxford: Hart Publishing, 2000). As leading cases on the tort of negligence and on compensatory damages arising therefrom, see *Donoghue v Stevenson* [1932] AC 532, 580; *Nettleship v Weston* [1971] 2 QB 691; *Smith v Leech Brain & Co* [1962] 2 QB 405; *The Wagon Mound No.2* [1967] 1 AC 617 Privy Council.
49 Alec Walen, "Retributive Justice," in *The Stanford Encyclopedia of Philosophy*, ed. Edward Zalta (URL = <https://plato.stanford.edu/archives/win2016/entries/justice-retributive/>.Winter 2016 Edition).
50 Jules Coleman, Scott Hershovitz, and Gabriel Mendlow, "Theories of the Common Law of Torts,"ibid. (\url{https://plato.stanford.edu/archives/win2015/entries/tort-theories/Winter 2015).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

derive not from a morally reprehensible conduct, but rather from a damaging violation of the victim's right[51].

In Law & Economics (L&E) theories, liability rules constitute economic incentives, leading agents to adopt economically efficient behaviours, which increase the overall social benefit. In this sense, paying damages is almost equal to buying the right to obtain the benefit associated with the wrong[52].

Nowadays, legal systems do not commit to only one theory of tort and justice, but rather to a combination of the three: the same normative framework will feature different models of liability rules, displaying a variety of imputation criteria (causation/remoteness, subjective element), which in turn reflect the peculiar rationales underlying the attribution of liability.

Many tort law systems – such as the Italian one[53] – have a general rule prescribing liability for damages caused by reprehensible behaviours on the basis of fault. This solution is moved by all the different goals defined above: not only *ex post* compensation and sanction, but also *ex ante* deterrence, since fault-based liability incentivizes agents to adopt the standard of care necessary to avoid harmful behaviours, as to avoid the negative economic consequences deriving from the duty to compensate.

Sometimes, however, the defendant is held liable in tort even though s/he did nothing blameworthy, merely because of the particular position that the s/he held towards the cause of the damage: i.e. person who holds a duty to watch over some other entity – such as the keeper, owner or user of a dangerous thing, the keeper or user of an animal – or the person who benefits from having or using a things, or running a specific activity[54]. The basic idea underlying the ascription of liability is that **whoever has the**

---

[51] Ibid. Under some version of this theory – developed to object other forms of liability, as developed by the school of law and economics – the principle of corrective justice that justifies the link which tort law creates between the victim and injurer, since it takes the injurer to have the duty to repair the wrongful losses that he causes, and neatly considers compensation as the primary function of liability, against that of inducing efficient behaviour.
[52] Calabresi, G. et al., "Property rules, liability rules, and inalienability: one view of the cathedral." Harvard law review (1972): 1089
[53] Art. 2043 Italian Civil Code: «*Risarcimento per fatto illecito.* Qualunque fatto doloso o colposo, che cagiona ad altri un danno ingiusto, obbliga colui che ha commesso il fatto a risarcire il danno».
[54] Examples from Italian civil code: artt. Article 2047. «Injury caused by person lacking capacity: If an injury is caused by a person incapable of understanding or intending, compensation is due from those who were charged with the custody of such person, unless they prove that the act could not have been prevented. If the person injured is unable to secure compensation from the person charged with the custody of the person lacking capacity, the court, considering the financial conditions of the parties, can order the person who caused the injury to pay equitable compensation» art. 2048 « Liability of parents, guardians, teachers, and masters of apprentices: The father and mother, or the guardian, are liable for the damage occasioned by the unlawful act of their minor emancipated children, or of persons subject to their guardianship who reside with them. The same applies to a parent by affiliation. Teachers and others who teach an art, trade, or profession are liable for the damage occasioned by the unlawful act of their pupils or apprentices while they are under their supervision. The persons mentioned in the preceding paragraphs are only relieved of liability if they prove that they were unable to prevent the act.»; art. 2049 «Liability or masters or employers: Masters and employers are liable for the damage caused by an unlawful act of their servants and employees in the exercise of the functions to which they are assigned.»; Article 2050. «Liability arising from exercise of dangerous activities: Whoever causes injury to another in the performance of an

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

economic or similar benefit associated with possessing or running a dangerous thing or activity, should also make sure that no damages are caused, and pay whenever this happens. This model is often associated with a strict or semi-strict liability basis, depending on whether or not the defendant may exclude his or her duty to compensate – i.e. by demonstrating that he took all the necessary measures to prevent the harm from occurring, or by demonstrating that the latter was caused by an act of God. The stricter the liability, the more compensation-oriented, instead of deterrence- and punishment-oriented, the rationale.

Along similar lines, **sometimes liability is ascribed to the person who is best positioned to manage and internalize the risk**, preventing its occurrence and minimizing its consequences, as well as to compensate the victim once an accident occurs. Such model is particularly common in L&E literature.[55]

A peculiar version of this model is the so called Risk Management Approach (henceforth RMA), which is grounded on the idea that liability should not be attributed on the basis of considerations of fault – defined as the deviation from a desired conduct – typical of most tort law systems, but rather on the party that is best positioned to (i) minimize risks and (ii) acquire insurance. It moves from the basic consideration that – although liability rules may well work as incentives or disincentives towards specific behaviours – they might not ensure sufficient and efficient incentives towards a desirable *ex ante* conduct, be it a safety investment – such as in the case of producers' liability – or a diligent conduct – such as the driver's in the case of road circulation – and that end is best attained through the adoption of detailed *ex ante* applicable regulation, such as safety regulation. According to this view, liability rules should thus be freed from the burden of incentivizing the agents towards desired conducts, and rather be shaped as to ensure the maximum and most efficient compensation to the victim. In extreme cases, this could also be designed to avoid the difficulties and burdens connected to traditional judicial adjudication, and rather be based on no-fault compensatory funds[56].

2. Robot

a. As Agents, and more specifically, as Moral Agents (Exclusion)

In the previous analysis (§§ 1.2.1, 1.3.1), we have clarified that for an entity to be deemed an agent, it shall be able to instantiate intentional mental states capable of directly causing performance; and that for it to qualify as a moral agent, it shall display what is usually referred to as "strong autonomy", i.e. the ability to decide freely and coordinate one's action towards a chosen end, as well as the moral awareness needed for understanding the moral significance of one's actions.

---

activity dangerous by its nature or by reason of the instrumentalities employed, is liable for damages, unless he proves that he has taken all suitable measures to avoid the injury»; art. 2051 «Damage caused by things in custody: Everyone is liable for injuries caused by things in his custody, unless he proves that the injuries were the result of a fortuitous event».

[55] Mitchell A Polinsky and Steven Shavell, *Handbook of Law and Economics*, vol. I (North-Holland, 2007).

[56] Palmerini and Bertolini, "Liability."; Andrea Bertolini, "Insurance and Risk Management for Robotic Devices: Identifying the Problems," *Global Jurist*, no. 2 (2016).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

In doing so, we have also explained why current robots, conceived to complete a specific task identified by their user, shall neither qualify as agents, absent the consciousness required for them to have intentional mental states; nor as moral agents, given that, at this stage, they have no capacity to engage in moral judgments, and lack a "strong autonomy", because while they can determine how to reach the goals they are programmed to achieve, these goals are are defined by an external agent – most likely, the designer, producer or programmer. The only moral agents involved in the functioning of the machine remain the humans behind it, who are responsible for both the goals chosen, the model of functioning designed for the robot, as well as the very choice to grant to it a certain degree of autonomy in determining how to perform intended tasks.

b. As Things (Products)

Having excluded any ontological reason why robots shall be deemed autonomous agents, thus moral and legal subjects, they shall be qualified as products: "artefacts crafted by human design and labor, for the purpose of serving identifiable human needs"[57]. Therefore, should a robot cause any damage, ordinary product liability rules would apply. Since the latter rest on the idea the producer shall be responsible because, and as long as, he is in full control of the features and actions of the products, **some authors have claimed that product liability rules are inadequate for the purpose of regulating the consequences deriving from a damage caused by R&AI applications, because the degree of autonomy they display creates a 'responsibility gap'**[58].

Regardless of the complexity of its functioning, **as far as the machine performs the tasks it was designed for, it is still under the control of the producer or the programmer**: even in the case of machine-learning technologies – such as neural based systems and genetic algorithms – the unpredictability of the learning behaviour does not create and actual lack of control, but rather requires the training and associated evolution of the robot to be included in the development phase, so that the product reaches the market only when it is supposed to have learnt or perfected the skill to function safely[59]. Should such threshold be impossible to reach, so that the machine seems not to be able to develop in a predictable way, the moral and legal responsibility for the damage caused still lays on the producer/programmer, who has a duty not to put unsafe products into the market.

---

[57] "Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules."
[58] Ibid.
[59] Bertolini, "Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules", *Law Innovation and Technology,* (2013): 214. Bertolini emphasizes the importance of testing and makes a policy consideration: it is not sound policy to incentivize people to put on the market things that are inherently dangerous, that those who designed do not understand or control. Each and every legal system imposes responsibility precisely to avoid such kind of scenarios. However, O'Brolcháin believes that the epistemic condition is unlikely to be met if genetic algorithms are used. The implication then would be that most products utilising them would be unsafe. This would certainly impede the field as there would be significant liability burdens on producers/programmers. Morally, if producers and developers cannot know what their product is going to do, it is difficult to say they are responsible. It might be the case that poor upkeep contributed to an event caused by a robot, so even liability becomes hard to prove with semi-autonomous robots. There is an ongoing debate on this point.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

What has been said so far against the alleged responsibility gap served to prove that **there are no compulsory ontological reasons why ordinary product liability rules shall not apply to advanced R&AI**. However, it could still be the case that changes to the existent paradigm shall be made, so as to address the regulation of new technologies, in a way which both fosters technological innovation, while being respectful of and driven by the respect of European values and principles[60]. Social and policy considerations, as well as constitutional law may suggest the **adoption of different liability models, favouring the development of applications which are particularly valuable for society**, such as prosthesis or devices intended to help the otherwise disabled in their everyday tasks.

Likewise, current liability rules may be reformed, in order to better pursue the goal, they are meant to achieve[61]. Indeed, the Product Liability Directive – which constitutes the European framework on the issue – has recently been evaluated as to assess whether it is still adequate for regulating contemporary advanced technological products. Some critical elements have been identified, primarily uncertainty as per the qualification of software as product, the implications and effectiveness of the development risk defense, and the cost and difficulty of exactly ascertaining the existence of a defect – in particular in design – as well as of a causal nexus between the fact and the damage. The latter, in particular, burdens the claimant substantially, discouraging litigation. Also, when advanced robotics is considered, **tight human-machine interaction causes different bodies of law to overlap**. Indeed, if a single task is handled together by the human agent and by a machine, when an accident occurs it might be due to the fault of the former or a defect (or malfunctioning) of the latter. Apportioning liability among the two – human agent or manufacturer – might therefore require complex factual ascertainment and articulate legal analysis. For this purpose, different approaches to liability – such as the abovementioned RMA – have been elaborated, with the purpose of modifying current product liability rules as to better address the new challenges brought about by technological innovation.

c. As Legal Agents: Electronic Personhood (Discussion)

Even if robots cannot qualify as autonomous beings and, thus, there is no ontological reason why they should be considered as "subjects" at law, this does not mean that they may not qualify as such, because of discretional choice of the legislator. Indeed, the constitutive independence between the notion of agency, moral agency and that of legal subjectivity is such that functional reasons could very well justify dissociation between the different states. For example, *ad hoc* legal personhood could be awarded to robots, exactly as it is granted to corporations. However, to justify this choice specific end needs to be identified, and a comparative judgment on the pros and cons of this alternative, as well as other tools, shall be taken into account. For example, it may be useful to attribute it to a robotic application, such as software agents, or drones[62], which would then be registered, as to identify the limits of its allowed tasks and functions, and eventually the (physical or legal) person it is representing.

[60] "Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. Artificial Intelligence for Europe," (Brussels: European Commission, 2018).
[61] Bertolini, Andrea. "Robots as products: the case for a realistic analysis of robotic applications and liability rules." Law, Innovation and Technology 5.2 (2013): 214-247.
[62] Ibid.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

With respect to liability issues, the recognition of legal personhood would mainly serve as a liability capping method; yet it would neither necessarily change the person bearing the costs of its functioning nor the cases when compensation is awarded. In fact, **unless the robot was capable of earning a revenue from its operation, its capital would have to be provided by a human, or a corporation, standing behind it, thus not necessarily shifting the burden from the party that would bear it pursuant to existing product liability rules[63]. Such result could also be achieved through insurance mechanisms or with a simple damages cap**. Should the robot be allowed to earn a fee for its performance, this would only constitute a cost on the user, producing an overall risk-spreading effect which could be effectively achieved otherwise, for instance through the adoption of a no-fault scheme funded by the product's users in various fashions[64]. Which of the different alternatives is preferable is still a matter of correctly specifying particular circumstances, among which are the size of the market for the given application and the existence of evident failures which could be designed around through *ad hoc* regulation; much less would depend on the machine being weakly autonomous or even able to learn.

3. Clarifying the role of IP law

1. Introduction

In light of the above discussion on agency and legal subjecthood, this sub-section discusses intellectual property (IP) protection in the field of robotics and focuses on the legal and philosophical challenges of interactive robotics, namely addressing the challenge of how the law should address the impact of human-robotic co-creation on the ownership of IP rights.

First and foremost, it is worth emphasizing that IP protection – via patents, copyrights, designs, trade secrets, trademarks, etc. - is key to the field of interactive robotics[65]. The lengthy and expensive **process of designing, developing, producing and delivering interactive robotic products relies on IP protection to recoup up-front investments and to fend off competitors** seeking to capitalize on the R&D investments of their rivals. IP is also important for investment and raising finance: a company subject to due diligence, because of - for instance - a strategic investment plan, acquisition or initial public offering (IPO), will likely have its IP portfolio reviewed as part of this process, with potential investors likely to view robotics firms without a strong IP portfolio as less attractive[66]. **Investors tend to not only want proof of a company's potential** for developing promising robotic applications **but also a policy on IP**

[63] Ibid.
[64] Ibid.
[65] See for example C. Andrew Keisner, Consultant, Julio Raffo and Sacha Wunsch-Vincent, Breakthrough Technologies – Robotics and IP, December 2016, Economics and Statistics Division, WIPO 2016, available at http://www.wipo.int/wipo_magazine/en/2016/06/article_0002.html.
[66] See the article "Making Your Robotics Company a More Attractive Investment", in Robotic Business Review of 21 October 2012, at https://www.roboticsbusinessreview.com/unmanned/making_your_robotics_company_a_more_attractive_investment.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

protection - i.e. positive indicator of the company's scientific inventiveness and its strategic economic planning[67].

Furthermore, in line with the above discussion on agency and legal subjecthood, the rise of interactive robotics brings a considerable challenge for the law: how should the law deal with robot-assisted invention and human-robot co-creativity? In other words, as robots become more and more capable of intelligent interaction with humans, to the extent that they begin to 'perform' semi-independent creative and inventive acts that produce new works and inventions (capable of being protected by copyright, patents, designs and trade secrets), how should the law protect such works/inventions? Who should own them?

2. How Intellectual Property Currently Applies to Robot-Assisted Creativity and Innovation in the EU

a. Patents

Patents on new inventions represent one of the main legal instruments used by robotic firms to protect their technologies. In Europe patents are granted by the non-EU European Patent Office (EPO) and by national patent offices. As a result, patents constitute an area of law that is not entirely within the jurisdiction of the EU, though there are several EU directives that apply directly to patents e.g. the Biotechnology Directive, the Enforcement Directive, etc. A system to enable unitary patents to be granted for participating EU member states and a Unified Patent Court to enforce such patents is currently being set up, but its start date is in doubt due to the impact of Brexit and various court challenges around Europe, including in Germany.

Robotic companies active within the European market often apply for a European Patent (EP). An EP is not a single patent, but rather a bundle of national patents granted by the European Patent Office that can be subsequently validated in multiple designated jurisdictions (including all 28 EU member states)[68]. Companies make decisions on their filing plans depending on whether a national market is of particular interest and based on the likely of infringements by direct or indirect competitors[69].

Patents give patentees – the owners of patents - the right to prevent others from exploiting the patented technology for a limited period of 20 years. Therefore, they exist as legal monopolies which give innovators a tool to maximize profits from a new technology for the first two decades of product/service development, distribution and sale. All technology companies – whether large, medium or small enterprises – tend to rely on patents to attract investors as well as to protect their investments in technology. For example, smaller, and more specialized, firms often use patents to protect their IP

[67] Ibid.
[68] The European Patent Office is based in Munich (Germany), its activity, and the patents it grants, being regulated by the European Patent Convention (Convention on the Grant of European Patents (European Patent Convention: "EPC") of 5 October 1973 as revised by the Act revising Article 63 EPC of 17 December 1991 and the Act revising the EPC of 29 November 2000).
[69] Linda J. Thayer; Rachel L. Emsley, Be Competitive: Patent Planning for Robotics Companies (2011) *Robot Magazine*, available at https://www.finnegan.com/en/insights/be-competitive-patent-planning-for-robotics-companies.html.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

assets defensively against larger players[70]. From the perspective of the public interest, a patent (and the monopoly coming with it) is granted in return for the disclosure of technical information so that the public at large, including patentees' competitors, will be able to exploit the invention after the 20-years term of protection expires.

The patent route can be particularly valuable for companies whose robots, or their elements, can be easily reverse-engineered (as is also known, reverse-engineering is the process whereby a product can be deconstructed to disclose its elements and the way it is manufactured). Indeed, in situations where reverse-engineering is simple, **filing for a patent may be favoured over the alternative tactic - trying to protect the process of manufacturing and/or the relevant product by keeping them secret -** with that patent being enforceable against any third party that exploits the invention without the patentee's consent. Thus, the decision to apply for a patent may be influenced by the complexity of the company's products and whether the company's competitors are likely to get their hands on such products and subsequently reverse engineer them. For example, are the robots likely to reach millions of private homes or will they merely be deployed behind closed factory doors? These are factors that need to be considered when it comes to protecting robotics innovation through IP.

There are examples of prominent robotics companies that have enforced their patents in Europe – notably, iRobot. In 2013 iRobot filed a lawsuit before the Court of Dusseldorf (Germany) against several companies including Solac GmbH, asserting that the Solac Ecogenic AA3400 vacuum robot had infringed five of its European patents. The case was subsequently settled. In 2013 iRobot brought another legal action, this time against the Chinese company Shenzhen Silver Star Intelligent Technology Co., Ltd., again before the Court of Dusseldorf. It obtained four preliminary injunctions, based on the German portions of four European patents[71], preventing the sale by the defendant of vacuum cleaner robots of the types XR210, M-H688 and M-788 in Germany[72]. **Nonetheless, patenting robotic technologies does not always produce benefits.** During the 1980s several companies in this field obtained numerous **patents that ended up expiring before the owners could commercialise the protected products**[73]. Thus, the successful use of patents in the field of interactive robotics requires further efforts within the innovation life cycle, including further R&D and link up with relevant technologies, such as the 'internet of things'.

b. Trade Secrets

As mentioned, robotics firms may rely on trade secrets and the legal protection given to such information, to protect their investments in technology. Trade secrets are protected in most countries of the world, although the type and degree of protection varies. In the EU, Directive 2016/943 was approved in June 2016 with the aim of harmonizing the laws

---

[70] See Keisner, Raffo, Wunsch-Vincent, above note 2, at p.27.
[71] The patents in question are EP 1 331 537 B1, EP 2 251 757 B1, EP 1 969 438 B1, and EP 1 395 888 B1.
[72] See iRobot's press release at http://investor.irobot.com/news-releases/news-release-details/irobots-preliminary-injunctions-against-shenzhen-silver-star.
[73] Keisner, Raffo, Wunsch-Vincent, above note 2, at p. 32.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

that protect undisclosed know-how and business information against unlawful acquisition, use and disclosure[74].

There are a number of reasons why a robotics company may choose to protect certain technologies **via trade secrets** rather than by seeking patent protection[75]. First, trade secrets confer protection without the need to adhere to certain prescribed formalities, such as filing an application with an office. **Robotics companies can therefore avoid certain costs and complexities associated with patent filing and prosecution**. This may be particularly useful in Europe where, due to fragmentation, obtaining and enforcing patents is more expensive than in other jurisdictions such as US and Japan. For example, not only the European Patent Office requires the payment of several fees to get European patents, including filing fees, search fees, fees per Designated States, fees per claim over ten claims, examination fees and finally fees for granting/printing[76]. The costs of enforcing European Patents may also be substantial as **patentees that seek to take action over infringements in various countries must file multiple legal actions before the national courts of those jurisdictions**, which is not only expensive, but also brings legal uncertainty as courts in different countries sometimes reach divergent decisions regarding the alleged infringement and the validity of the national portion of the European patent[77].

Second, trade secrets (rather obviously) do not require disclosure, as the patent system does. **For robotics inventions that are more difficult to reverse-engineer, the trade secrets option may prove a superior alternative as the protection could potentially last indefinitely**[78]. As has also been stressed, "trade secrets can be critical to [robotic] inventions that may not gain market acceptance and momentum for a long time"[79].

Third, **trade secrets can protect subject matter that patents cannot**[80], for example innovation related to software and computer code. This option could be particularly beneficial in light of the fact that protecting software inventions via patents has proven to be a contentious (and complicated) at national and EPO levels.[81]

---

[74] The aim of this piece of legislation is to harmonise national regimes on the protection of confidential information.
[75] Keisner, Raffo, Wunsch-Vincent, above note 2, at p. 27.
[76] Enrico Bonadio, The EU Embraces Enhanced Cooperation in Patent Matters: Towards a Unitary Patent Protection System (2011) European Journal of Risk Regulation, Vol. 3, p. 416.
[77] Luke McDonagh, European Patent Litigation in the Shadow of the Unified Patent Court (Edward Elgar, 2016).
[78] See discussion at the website of the Finnegan law firm https://www.finnegan.com/images/content/8/6/v3/866/IntellectualPropertyConsiderationsfortheRoboticsIndustry-revised.pdf, p.3
[79] See discussion at the website of the Finnegan law firm at https://www.finnegan.com/images/content/8/6/v3/866/IntellectualPropertyConsiderationsfortheRoboticsIndustry-revised.pdf at p.3
[80] Ibid.
[81] See also C. Leroux et al, EU Robotics Coordination Action: A green paper on legal issues in robotics (2012), p. 29, available at https://www.unipv-lawtech.eu/files/euRobotics-legal-issues-in-robotics-DRAFT_6j6ryjyp.pdf.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

c. Copyright

Certain elements of interactive robotic devices, especially software code embedded within, can be protected by copyright if they satisfy the relevant requirements, including originality. Copyright is a common legal tool to protect software in Europe - and the EU has harmonised such protection since 1991[82]. Software code enables interactive robots to engage in acts such as pathfinding, control, locating and sharing data; moreover, some programming code also aims to imbue robots with the ability to create artistic, literary and musical works. Relying on copyright to protect such software is therefore key for the robotic industry[83].

Firms in this field may also rely on 'technological protection measures' to restrict access to, and prevent copying of, a robot's copyright-protected code[84]. This makes it difficult for third parties, both competitors and users, to get their hands-on relevant software code, by inserting electronic barriers to prevent access. Copyright laws allow this construction of barriers. Moreover**, circumventing electronic barriers to gain access to copyrightable computer code is considered a violation of copyright**. Take for example the EU Directive 2001/29 on Copyright in the Information Society, which provides that adequate legal protection must be given "against the circumvention of any effective technological measures, which the person concerned carries out in the knowledge, or with reasonable grounds to know, that he or she is pursuing that objective"[85]. It is a type of protection that may be useful against users or competitors that want to access commercially valuable software code.

d. Trademarks

Trademarks - such as distinctive logos and brand names - are administered by the EU Intellectual Property Office (EUIPO) and at national offices. Registering trademarks is crucial to protect a robotic product's goodwill and reputation, especially in business-to-consumer industries. Notably, interactive robotics is increasingly becoming an industry where products are sold directly to millions of end-users (consumers). The commercial success of products such as nanny-robots, pet-robots, caretaker-robots and medical-

---

[82] See **Council Directive 91/250/EEC of 14 May 1991 on the legal protection of computer programs, repealed by** Directive 2009/24/EC.
[83] Alleged copyright infringement has also been the focus of the above-mentioned dispute *iRobot v Urus Industrial Corporation* (which was eventually settled). That case was not just about alleged infringement of the patents covering certain functional aspects of the Roomba. iRobot also claimed that Urus' vacuum cleaner robot infringed copyright in iRobot's product literature and system interface, including its musical audio feedback features.
[84] Keisner, Raffo, Wunsch-Vincent, above note 2, at p.34.
[85] Article 6 of Directive 2001/29. A similar provision is set forth by the US Digital Millennium Copyright Act (DMCA): a provision which was invoked in 2001 by Sony when a programmer created and distributed via a website free software to enhance the capabilities of the robot dog named Aibo produced by said Japanese company (such user basically decrypted the code defining the robotic dog's abilities). Sony complained that the website in question provided the means to circumvent the copy protection protocol of Sony's AIBO memory stick to allow access to the relevant software, and therefore constituted a violation of the DMCA anti-circumvention provision. On this case see Matthew Rimmer, Respect the Code or the Dog Gets It (December 2001) InCite, p. 31, available at https://eprints.qut.edu.au/98504/1/aibo.pdf

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

robots also depends on a reliable brand which consumers know, trust, appreciate and remember86. For this reason, **robotics companies with a strong brand name and solid reputation are indeed investing in and registering trademarks, especially with the EUIPO**, which grants registrations valid and effective in all EU Member State. Brands such as 'iRobot'87, 'ABB'88 and 'Kawasaki'89 as well as 'Roomba' (the robotic vacuum cleaner from iRobot)90 have all been registered with the EUIPO. Given the growing propensity of companies in this sector to register trademarks and build overarching brand identities, and the increasing availability of robots amongst final consumers, disputes about robotics trademark infringements may soon reach courts, in Europe and elsewhere.

e. Designs

Interactive robots are consumer-facing - and as such **a robot's physical appearance and its 'look and feel' plays a central role in influencing consumers' choic**e[91]. Robot designs that meet certain requirements, including novelty and individual character, can be registered with the EUIPO, such registrations protecting the ornamental features of the machines. Under EU law, for example, it is possible to obtain an EU design registration which is valid in all Member States (up to 25 years), with a shorter protection of 3 years also offered to unregistered designs[92]. The exclusive rights given by the registrations can then be enforced against third parties that use designs that are perceived by an informed user as giving the same overall impression.

Some robotics companies in Europe have indeed taken advantage of this chance and obtained EU design registrations protecting the ornamental features of products such as vacuum cleaners[93], robotic lawnmowers[94] and transportation robots[95]. Also, design rights may soon be regularly sought by companies active in the field of wearable robots, i.e. devices that are used to enhance people's motion and physical abilities. Despite having functional elements, these products may be devised in a way which makes them more appealing to final consumers – and **design rights could exactly be the appropriate legal tool in the hands of such firms to protect the eye-catching elements of their products**. In other words, these rights may help these companies to keep pace with the likely "fashionisation" of the robotic industry.

---

[86] Luke McDonagh. 'From Brand Performance to Consumer Performativity: Exploring TradeMark Law in the Aftermath of Anthropological Marketing' Journal of Law & Society (2015).
[87] See the webpage https://euipo.europa.eu/eSearch/#details/trademarks/W01353068.
[88] See the webpage https://euipo.europa.eu/eSearch/#details/trademarks/002628964.
[89] See the webpage https://euipo.europa.eu/eSearch/#details/trademarks/000814681.
[90] See the webpage https://euipo.europa.eu/eSearch/#details/trademarks/002995108.
[91] See Meenakshy Chakravorty, Elizabeth D. Ferrill, Linda J. Thayer, and Robert D. Wells, Design-Patent Protection for Modern Robotics Companies: What to Do When the Face of Your Robot Becomes the "Face" of Your Company, published in Robotics Business Review, July 1, 2014, at p.6.
[92] EU Regulation 6/2002.
[93] See https://euipo.europa.eu/eSearch/#details/designs/004680866-0025; https://euipo.europa.eu/eSearch/#details/designs/004680866-0026.
[94] See the webpage https://euipo.europa.eu/eSearch/#details/designs/002524462-0002.
[95] See the webpage https://euipo.europa.eu/eSearch/#details/designs/005418506-0001.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

3. Dealing with The Challenges of Interactive Robots

We now turn our attention to the fascinating issue of whether interactive robots can generate unpredictable output that can be protected by IP. Indeed, as the software-hardware integration becomes increasingly central for the robotic industry, robots with an express ability to create and invent in their own right could be soon become widespread and of common use. After all, there are already interactive 'smart' robots embedded with AI which show such abilities[96]. A striking example is Paul, a robot that uses its camera eye and arm to draw portraits of human subjects[97].

Should these outputs be protected by copyright and patent laws? Should they be left to the public domain? If they are protectable, who should be deemed the owner of the resulting copyright or patent? The programmer? Or the user?[98] The key IP issues are essentially the same whether we talk about pure and intangible AI-empowered software (e.g. algorithms) or if we consider AI-equipped interactive robots that rely on physical embodiment to enhance creativity. For this reason, scholarship and discussion of AI is relevant to the questions of whether works and inventions created by interactive robots should be protected by IP rights, and regarding who should own such rights.

a. Machines, creative works and copyright

Some European copyright statutes such as the UK Copyright Designs Patent Act (CDPA) explicitly include computer generated works amongst the subject matter of copyright[99]. Indeed, works of music, literature and art are already being produced by computers and machines today. Examples abound: from Jukedeck (a startup that uses AI to produce music)[100] to the Cybernetic Poet (a software which allows a computer to write poetry)[101] and The Next Rembrandt (a 3D printed painting made solely from data of Rembrandt's body of work)[102], amongst many others.[103]

---

[96] For the sake of clarity, we will use in the remainder of the article the terms "machine", "robot" and "computer" interchangeably.
[97] See the BBC webpage http://www.bbc.co.uk/programmes/articles/1f4Z6k7Clz6qY6Q2K56nkzZ/robotticelli-the-mechanical-marvel-creating-extraordinary-works-of-art.
[98] See, amongst others, Ramalho, Ana. "Will Robots Rule the (Artistic) World? A Proposed Model for the Legal Status of Creations by Artificial Intelligence Systems." (2017); Abbott, Ryan. "Artificial intelligence, big data and intellectual property: protecting computer-generated works in the United Kingdom." Research Handbook on Intellectual Property and Digital Technologies (Tanya Aplin, ed), Edward Elgar Publishing Ltd, Forthcoming (2017); Denicola, Robert C. "Ex Machina: Copyright Protection for Computer Generated Works." Rutgers UL Rev. 69 (2016): 251. McCutcheon, Jani. "The vanishing author in computer-generated works: a critical analysis of recent Australian Case Law." Melb. UL Rev. 36 (2012): 915; Bridy, Annemarie. "Coding creativity: copyright and the artificially intelligent author." Stan. Tech. L. Rev. (2012): 5; Abbott, Ryan. "I Think, Therefore I Invent: Creative Computers and the Future of Patent Law." BCL Rev. 57 (2016): 1079.
[99] Section 178 of the UK Copyright, Designs and Patents Act 1988.
[100] See the webpage https://www.jukedeck.com.
[101] See the webpage http://www.kurzweilcyberart.com/poetry/rkcp_poetry_samples.php.
[102] See the webpage https://www.nextrembrandt.com.
[103] Artworks produced by AI have also started being auctioned by auction houses: see for example a sale by Christie of an AI-created portrait in a gilt frame in October 2018 (see Christie webpage at https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

To help understand the copyright debate around these categories of works, **a distinction should be made between computer-*aided* and computer-*generated* works**, with the former category representing works that are produced by humans with the mere help of machines and the latter referring to output autonomously created by AI. Ryan Abbott explores this distinction by reference to a spectrum: "On the one end, computers may function as simple tools that assist human authors …, much the way that a pen … can help someone to write"[104]. He continues by noting that "[a]t the other end of the spectrum, computers generate works under circumstances in which no human author … can be identified"[105]. In other words, **the level of machine autonomy in producing the work is inversely proportional to the presence of human input in the creative process**: more machine autonomy means less human input.

This debate is not entirely new, and judges in the past have looked at situations where the creation of works occurred with the help of machines. The 1980 English case *Express Newspapers plc v Liverpool Daily Post* noted that a programmer that instructed a computer to produce a series of sequences and grids was entitled to copyright[106]. The key point made by the judge was that the technology was a tool of the author. Suffice to say, technology has advanced significantly since then. Recent developments have led to machines that can independently learn and create, with the human input in the creative process becoming more and more redundant[107].

Take the 'Painting Fool', a striking example of a creative machine, and an "aspiring painter"[108]. It is a computer program which can simulate the physical painting process and detect emotions of people as well as use its abilities to paint portraits and invent visual scenes by means of generative techniques[109]. Another example is the above-mentioned Paul, described as an interactive robotic installation which uses its camera eye to create portraits of people[110]. As opposed to the Painting Fool, which is a computer program, Paul consists of a physical robotic arm. As explained by its creators Tresset and Leymarie, "[t]he drawings we are aiming to produce with an embodied system such as Paul are distinct from those made by a human hand, and yet it is our experience that they have comparable emotional and aesthetic artistic effects on the observer"[111]. Reactions by critics have been positive, with collectors and artists accepting Paul's productions as artworks of good quality (a drawing by Paul is part of the Victoria and Albert museum collection).[112]

---

[104] Abbott, above note …, p.2.
[105] Abbott, above note …, p.3.
[106] Express Newspapers Plc. v Liverpool Daily Post & Echo Plc. and Others [1985] 1 W.L.R. 1089.
[107] de Cock Buning, Madeleine. "Autonomous Intelligent Systems as Creative Agents under the EU framework for Intellectual Property." European Journal of Risk Regulation 7.2 (2016): 310
[108] See the webpage at http://www.thepaintingfool.com. Simon Colton, a Professor of Computational Creativity in the Department of Computing of Goldsmiths College, University of London, is the academic behind the Painting Fool.
[109] See the webpage at www.thepaintingfool.com/about/index.html.
[110] Patrick Tresset and Frederic Fol Leymarie, Portrait drawing by Paul the robot, Computers & Graphics 37 (2013) 348–363 at p.348.
[111] Idem.
[112] See also Bridy, above note …, p. 4 (discussing on the generative art movement, which aims at exploring 'computational creativity' via "a set of creative practices whereby the artist cedes control to a system that is self-contained enough to operate autonomously").

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

We now assess whether machine-created outputs - e.g. works created by interactive robots - are capable of meeting the requirements for protection[113].

i. Originality and authorship

Copyright protects works that satisfy the originality requirement. The EU originality criteria as affirmed in the 2009 CJEU case of *Infopaq* requires the work to be the author's own intellectual creation: "... works such as computer programs, databases or photographs are protected by copyright only if they are original in the sense that they are their author's own intellectual creation"[114]. Following *Infopaq*, this is considered a generalised standard of originality applying not only to computer programs, photographs or databases[115], but to all literary, dramatic, musical and artistic works. In *Eva-Maria Painer* (focusing on copyright protection of portray photographs), the Court of Justice of the European Union (CJEU) clarified that an intellectual creation is an author's own if it reflects her personality[116]. This would be the case, the court added, if the author were able to express her abilities in the production of the work by making free and creative choices[117]. These choices would therefore enable the author to stamp the work with her 'personal touch'[118].

The emphasis on the author's 'own intellectual creation' and 'personal touch' suggests that the originality requirement involves some degree of human authorship[119]. This is reinforced by what Advocate General Trstenjak noted in *Eva-Maria Painer;* "… only *human creations* are … protected, which can also include those for which the person employs a technical aid, such as a camera" (emphasis added)[120]. This comment calls to mind the distinction between computer-*aided* works and computer-*generated* works. It also seems to suggest that only computer-aided works can be protected by copyright, where computer-generated works cannot qualify as 'human creations' due to lack of human input. **One may reasonably doubt about whether a machine can stamp its output with its personal touch by making free and creative choices**, and in general whether it can have a personality at all, let alone legal personality, and all the rights that status would bring.

Yet, a look at how some machines actually work may lead to another conclusion. Think again about Painting Fool as an example: during a festival on computational creativity in 2013, the machine was used to create 'mood-driven' portraits of the guests. The software's mood was determined by its analysis of newspaper articles from *The Guardian*. The average sentiment was then used to simulate the Painting Fool, resulting in positive,

---

[113] The fixation requirement, provided by several copyright acts (for example, in the US and UK), will not be analysed here. There is indeed little doubt that most of machine-produced works meet this requirement.
[114] Case C-5/08 Infopaq International A/S v Danske Dagblades Forening at [35]
[115] See respectively Article 1(3) of the Software Directive (Directive 2009/24), Article 6 of the Copyright Term Directive (Directive 93/98) and Article 3(1) of the Database Directive (Directive 96/9).
[116] Case C-145/10 Eva-Maria Painer v Standard Verlags GmbH and Others at [88].
[117] Case C-145/10 Eva-Maria Painer v Standard VerlagsGmbH and Others at [89]. See also Case C-604/10 Football Dataco at [39] (citing also, by analogy, Case C-393/09 *Bezpecnostni* Softwarova *Asociace* at [48] and [49], and Joined Cases C-403/08 and C-429/08 *Football Association Premier League and Others* at [98]).
[118] Case C-145/10 Eva-Maria Painer v Standard VerlagsGmbH and Others at [92].
[119] Madeleine de Cock Buning, above note …, p.314.
[120] Case C-145/10 Eva-Maria Painer v Standard VerlagsGmbH, Opinion of Advocate General Trstenjak, 12 April 2011, at [121].

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

very positive, experimental, reflective, negative or very negative moods[121]. As explained by Colton and Ventura: "[i]f in a positive/very positive mood, the software [the Painting Fool] chooses one/two of nine upbeat adjectives (e.g. *bright, colorful, happy*) and directs the sitter to smile while it extracts their image from a video recording. ... If in a negative mood, the software chooses one of six downbeat adjectives (e.g. *bleary*, *bloody*, *chilling*) and directs the sitter to express a sad face. If in an experimental mood, it chooses one of 11 neutral adjectives (e.g. *glazed*, *abstract*, *calm*) and asks the sitter to pull an unusual face ..."[122]. The chosen adjective is used to select a filter to achieve an appropriate visualisation. The program also selected one of seven rendering styles involving the simulation of paints, pencils and pastels to produce the portrait.

That said, could one argue that the behaviour of the Painting Fool displays an ability to make creative and independent choices which renders its works original? After all, in *Eva-Maria Painer* the CJEU suggested that an author could stamp a portrait photograph with her personal touch by making choices such as the applicable background, lightning, angel and atmosphere of the portrait[123]. Similarly, in light of the software's freedom to make creative choices by reference to its mood, it could be said that the Painting Fool could be considered as imbuing the portraits with its personal touch.

Yet, one may counter-argue that machines such as the Painting Fool or robots like Paul lack a fundamental ingredient of any copyright work, namely a human being as author/personality. This point brings into the picture another requirement for copyright protection, i.e. authorship, which entails that **no creation that does not entail at least some degree of human intervention is eligible for copyright protection[124]. The concept of authorship is intertwined with the originality requirement**[125]. Indeed, the latter's focus on the 'author's own intellectual creation' seems to imply a *human* author behind the work. Several copyright laws limit authorship to natural persons. Spanish law for example provides that the author is the natural person creating the work[126]; French law states that only a natural person can be the author[127]. Likewise, the US Copyright Office emphasises the importance of the human element in the creative process, with no claim that does not satisfy the Human Authorship Requirement being

[121] See webpage at http://www.thepaintingfool.com/galleries/you_cant_know_my_mind/ICCC_YCKMM.pdf.
[122] Ibid.
[123] Case C-145/10 Eva-Maria Painer v Standard VerlagsGmbH and Others at [91]-[92].
[124] Most copyright acts, including in the US and UK, however, still do not provide a definition of "authors". Just judicial decisions address what authorship means, and who an author is (see Jane C. Ginsburg, The Concept of Authorship in Comparative Copyright Law, 52 DePaul L. Rev. 1063, 1066 (2002)). In the US for example the Supreme Court has defined an author as 'he to whom anything owes its origin; originator; maker; one who completes a work of science or literature' (*Burrow-Giles Lithographic Co. v. Sarony*, 111 U.S. 53, 58 (1884)). US lower courts have also noted that an author must be more than one who contributes creativity or originality to a work (*Aalmuhammed v. Lee*, 202 F. 3d 1227, 1233 (9th Cir. 2000)); basically, it must be one 'who superintended the whole work, the "master mind"' (Ibid. (citing *Burrow-Giles Lithographic Co.*, 111 U.S. at 53, 61)).
[125] See also Ramalho, above note … p. ….
[126] See Preambulo, *Ley 22/11 sobre la Propiedad Intelectual de 1987*: "los derechos que corresponden al autor, que es quien realiza la tarea puramente humana y personal de creacion de la obra y que, por lo mismo, constituyen el nucleo esencial del objeto de la presente Ley".
[127] The French Code of intellectual property defines protectable subject matter as 'oeuvres de l'esprit': see Article L112-1 of the French Code de la Propriete' Intellectuelle.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

registered. In other words, the US Office only registers an original work of authorship "…provided that the work was created by a human being"[128].

The authorship requirement will not raise significant issues when the machine merely *aids* a natural person in the creative process (and indeed in *Express Newspapers plc v Liverpool Daily Post* Mr Justice Whitford found no authorship issue). This is also the point made by Advocate General Trstenjak in *Eva-Maria Painer*, i.e. that copyright protection is available for human creations, including those for which the person employs a technical aid, such as a camera[129]. Another example is that of a person using the aid of editing software to produce a picture[130]. Here, the end product will reflect the creative impulses of the artist and thereby her personality[131].

Authorship issues may instead arise where the machine input materially outweighs that of the human which uses it, i.e. where the level of human intervention in the creative process is minimal and when that of the robot is predominant. This is when the line has been crossed from a situation where the machine has merely aided the human to create towards a scenario where the work has been generated by the robot itself. Verifying when such line is crossed may not be easy, and **a thorough case-by-case analysis may often be required to determine the level of human input in the whole creative process**. In other words, deciding whether a work qualifies for copyright protection under EU copyright law (as well as under US law, which as mentioned does require authorship)[132] requires an assessment of the level of human as opposed to machine input[133]. The aim of this analysis is to determine whether the work can be said to reflect predominantly the (human) author's own intellectual creation, and thus attract copyright. If such human element lacks, copyright will not subsist.

ii. The UK solution

An approach that aims at finding a pragmatic solution has been adopted by the UK. Section 9(3) of the CDPA provides that "[i]n the case of a literary, dramatic, musical or

---

[128] U.S. Copyright Office, Compendium of U.S. Copyright Office Practices § 101 (3d ed. 2017), 306. Examples of works that fail the authorship requirement include, for instance, a photograph taken by a monkey and works "... produced by a machine ... that operates randomly or automatically without any creative input or intervention from a human author" (see Compendium 313.2).

[129] Case C-145/10 Eva-Maria Painer v Standard VerlagsGmbH, Opinion of Advocate General Trsteniak, 12 April 2011, at ……

[130] Guadamuz, above note …, p.178.

[131] Guadamuz, above note …, p.178.

[132] That authorship is a requirement under US copyright law is confirmed by a string of cases that have dealt with who should be considered author in situations where "celestial voices" may have played a role (psychography cases): *Oliver v Saint Germain Foundation*, 41 F. Supp. 296 (S.D. Cal. 1941); *Urantia Foundation v Burton*, No K 75-255 CA 4, 1980; *Urantia Foundation v Maaherra*, 895 F. Supp. 1337 (D. Ariz. 1995); *Penguin Books, Inc. v. New Christian Church of Full Endeavour*, No 96 CIV. 4126. Indeed, when faced with claims of supernational authorship, US courts in those disputes concluded that only humans can own the copyright. Similarities can also be drawn with claims of "animal authorship", where US courts have confirmed this line of argument: see *Naruto v. Slater*, 2018 U.S. App. LEXIS 10129 (9th Cir. Cal., Apr. 23, 2018) (rejecting the argument brought by the plaintiff, an animal rights organisation, that US copyright law does not prohibit an animal – in the case at issue, a monkey that have taken a selfie - from owning a copyright. The court held that while the animal had constitutional standing it "lacked statutory standing to claim copyright infringement of photographs)."

[133] Guadamuz, above note …, p.179.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

artistic work which is computer-generated, the author shall be taken to be the person by whom the arrangements necessary for the creation of the work are undertaken". A few other common law jurisdictions have followed this approach[134]. The CPDA also defines a computer-generated work as one generated by a computer in circumstances where there is no human author[135]. What these provisions do is basically to broaden the concept of author in a way which is sufficient enough to subsume human beings that simply instigate and trigger the creation of the work. In other words, the author (and owner of the relevant economic rights) will be considered a person (either natural or legal) who may have played no role at all in the actual production of the work. **The law here basically introduces a legal fiction as it considers author a person who has not directly produced the word but has merely made the necessary arrangements for such production**.[136]

This solution clearly departs from the anthropocentric and human focus of copyright laws in jurisdictions such as the EU and US. Guadamuz notes that these provisions that recognise copyright in computer generated works with no human input constitute an exception to the originality requirement, as the works in question do not directly originate from an author[137]. Indeed, such solution does not look at whether a human being has actually produced the work. It just considers the objective creation of the output, and then "finds" who the author (and thus the copyright owner) should be, i.e. the person who has come up with the necessary arrangements, which could also be a company. Yet, it is not quite a revolutionary approach, as copyright laws in several jurisdictions, including the UK, sometimes attribute ownership to persons who have not directly created the work. This is the case with regard to employees' works and commissioned works, with the employer and commissioner, respectively, being the copyright owners.

The UK regime is certainly helpful and may be used to identify authorship, and accordingly ownership, in many scenarios where original works are produced by computers or robots with no or little human input[138]. Yet, it might not be always easy to identify who the person who has made the necessary "arrangements" is, namely who puts the machine in the condition to create the work. Is this person the machine's programmer? Or the user? Section 9(3) of the CDPA may not help much here. It could be argued that **determining who that person is entails a careful analysis of the facts, circumstances and specific robotic application that generated the work**. A case-by-case analysis again is therefore necessary. For example, in *Nova Productions v Mazooma Games*[139]*,* the only UK case where s. 9(3) CDPA has been applied so far, the court had to determine whether copyrights had been infringed in the graphics and frames

---

134 Namely, Ireland, New Zealand, India, South Africa and Hong Kong.
135 See again Sec. 178 of the Copyright, Designs and Patents Act 1988.
136 McCutcheon, above note …, at pp. 44-45.
137 Guadamuz, above note …, p.176.
138 For an opposite opinion see Lionel Bently, mentioned by Begoña González Otero and Joao Pedro Quintais Before the Singularity: Copyright and the Challenges of Artificial Intelligence (2018) Kluwer Copyright Blog, available at http://copyrightblog.kluweriplaw.com/2018/09/25/singularity-copyright-challenges-artificial-intelligence/ (stressing that the CDPA provisions on computer-generated works do not offer a useful model for protecting AI outputs, because of their incompatibility with the EU copyright *acquis* and failure to address the issue of originality).
139 *Nova Productions v Mazooma Games* [2007] EWCA Civ. 219.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

generated and displayed by the users on a screen when playing a videogame. It was held that such frames were computer generated works, and that the programmer was the person making the arrangements and therefore the author and owner of the copyright. As the Court of Appeal put it, "the player is not ... an author of any of the artistic works created in the successive frame images. His input is not artistic in nature ... and he has contributed no skill or labour of an artistic kind ... All he has done is to play the game"[140].

With that said, what about other more robot-focused examples? Let's take again 'Paul', the robotic arm that produces observational face drawings of people, by using an eye camera to focus on and take snapshots of the person to be drawn and then executing the drawing with a pen hold by the arm. Whoever uses Paul does not really have the power to change the settings, for example to direct or change the process that leads to the artistic output. It could thus be argued that Paul's creator (and not the person who actually uses it) should be considered the person that makes the necessary arrangements. The same is true of the aforementioned 'Painting Fool' - a computer program that simulates the physical painting process without giving users much control on how to drive such process.

On the other hand, whenever users of the machines or robots have actually the chance to manage or influence the whole or part of the process (in other words, where they do not have to just press a button), they may be considered the authors and owners of the resulting work. **Two scenarios may here be distinguished**. First, we could have a situation **where the user merely manages the creative process by running the program and generating the output**. An example might be Deep Dream, a popular computer vision program that employs a convolutional neural network to find and enhance patterns in images through algorithms. It has been argued that, as the creator of the program (Google) has released the code as open source and any user can run the program and actually generate art (for example, by choosing predetermined styles), it would be the users the ones that make the necessary arrangements and thus should be considered authors under Section 9(3) CDPA[141]. Second, we could also have a scenario **where users' input is more creative than managerial**, for instance when it's the user that makes the most important creative choices with the machine or robot merely executing such decisions[142]. In this case we would probably move towards the other side of the spectrum identified by Ryan Abbott, where computers function as simple tools that assist human authors. Such works would therefore be more comparable to machine-aided works, with no need to trigger the "authorial" legal fiction engineered by Section 9(3) CDPA[143].

---

[140] Ibid., at [106].
[141] Guadamuz, above note ..., p.177.
[142] This scenario may soon become a common reality, for example in the field of computer games, taking into consideration the increasingly relevant roles of users in such interactive games: see Ramalho, above note …, p. ….
[143] See also Denicola, pp.283-285 (noting that, should users be considered the authors of computer-generated works, there would not be need anymore to distinguish between machine-generated works and machine-aided works. Indeed, if the creative output is considered to be authored by someone other than the user of the machine (or is not deemed copyrigthable at all), it would be necessary to differentiate situations where the computer is simply a tool from scenarios where the machine itself is the creator. Denicola thus believes that the "users' option" should be preferred. The market – he stresses - already supplies programmers with the incentive to create software and machines which are

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

iii. Copyright or Public Domain?

We have considered whether copyright *can* subsist in machine-generated works, and who would be considered author and owner of the economic rights that result under the law. But what also ought to be discussed is whether copyright *should* apply to these outputs at all, as a matter of policy. Some scholars have proposed leaving such works – works generated autonomously by machines - in the public domain, with the result that everyone would be able to use them, even for commercial purposes[144]. As has been noted, outputs produced by robots would be comparable to things found in nature, like music that the wind generates when it moves through wind chimes or the sounds of a waterfall, or birds singing at dawn[145].

But is this outcome desirable? Would the refusal to offer copyright to machine-generated works discourage investments in, and dissemination of, these "creative" technologies?[146] It could indeed be argued that without the lure of monopolistic rights offered by copyright not many persons and/or their employers would bother to develop machines or robots capable of creating musical, literary or artistic outputs. As this is already a field where investments are necessary (let's think of companies that offer music production software), the need to secure copyright may become apparent with the industry, especially to avoid the inevitable market failure that would be caused by an absence of exclusive rights over the created content[147].

A couple of proposals have been mentioned and put forward by Ana Ramalho that take into account the above concerns[148]. The first is the introduction of a *sui generis* **right of a limited period of time protecting output created by AI embedded machines and robots,** similar to the EU database right which aims at protecting investments[149]. This right would thus incentivise and reward the investments made in the development of those machines and robots. The second proposal, more concerned with enhancing the accessibility of the content produced by AI, is to create **a sort of "disseminator's right"** comparable to the publisher's right in the publication of previously unpublished works provided by the EU Term of Protection Directive[150]. While this right would be

---

used to create works, such incentive mainly coming from the prospect of maximising sales. Also, should programmers be deemed as authors and owners of the copyright over the work produced by the machine – his argument goes - users may not be encouraged to purchase and use the software to create works as they would not be the copyright holders: which would eventually jeopardise the interests of programmers that may see sales of their machines decrease. Finally, attributing programmers copyright may often turn out to be useless as programmers themselves would frequently be anaware of the creation of works by users.

[144] Amir H. Khouri, Intellectual Property Rights for Hubots: on the Legal Implications of Human-like Robots as Innovators and Creators (2017) Cardozo Arts & Entertainment Law Journal, Volume 35, Issue 3.

[145] Khouri, above note …, p. ….

[146] See also McCutcheon, above note ..., p.952 (noting that "while the incentive of copyright may be irrelevant to a novelist compelled to write their opus, it may well explain why an expensive computer-generated production is made. Without that reward, the work may not be made ... or disseminated").

[147] See Tatiana Synodinou, mentioned by González Otero and Quintais, above note ….

[148] Ramalho, above note …, p. ….

[149] Database Directive (Directive 96/9).

[150] Article 4 of Copyright Term Directive (Directive 93/98) provides that "Any person who, after the expiry of copyright protection, for the first time lawfully publishes or lawfully communicates to the public a previously unpublished work, shall benefit from a protection equivalent to the economic rights

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

inherently economic, which would allow the right holder to extract value out of the creations, it would also intend to stimulate the dissemination of output that is increasingly perceived as being meritorious.

b. Inventions developed by machines, and patents

This sub-section explores the issue of patentability of robot-generated inventions. Again, the focus here is not on machines which merely *aid* humans during the inventive process, but rather on robots that *generate* inventions without human input. The burning question is whether such inventions can be protected under patent law. No patent statute, and related case law authority, addresses computational subject matter and no patent office has adopted specific policies on these issues. This is not a purely academic exercise - and **the debate on whether machine-generated inventions should be considered patentable is much needed also in light of the fact that patent protection in this field may further encourage the development of creative computers and systems, which may be deemed a socially desirable target**.

It has been noted that computers have already come up with inventions. For example, Hattenbach and Glucoft reported that a company named Cloem use brute-force computing to mechanically compose text for thousands of claims covering potential new inventions[151]. More precisely, automated software here employs automated drafting techniques which can create tens of thousands of alternative patent claims[152]. Also, Ryan Abbott mentions the so-called Creativity Machine, a patented[153] computational machine created by US computer scientist Stephen Thaler that is capable of generating novel (patentable) ideas via software concept referred to as artificial neural networks[154]. Importantly, one of these machine-generated ideas – called the "Neural Network Based Prototyping System and Method" - was actually patented[155], but it is notably that the designated inventor in the application was not the Creativity Machine itself, but its creator Stephen Thaler, who listed himself as the actual inventor. Abbott uses this example to make the point that patent offices (in this case, the US patent office) have already granted patents for inventions developed by non-human inventors, probably without being aware of such non-human participants; and to stress that this has probably occurred because the applicants did not want to risk losing the opportunity to get the patent, as there is a lack of legal clarity about whether machine-generated inventions can qualify for patent protection[156]. Thus, **omitting to disclose the role of the machine in the inventive process can be seen as an appealing option to avoid the patent being challenged on the grounds of lack of human inventorship. There appears to be a problematic gap in the law**.

---

of the author. The term of protection of such rights shall be 25 years from the time when the work was first lawfully published or lawfully communicated to the public".

[151] Hattenbach – Glucoft, above note …, p. 35 (reporting that said brute-force computing has also been used to create defensive publications aimed at preventing others from patenting in the same field).

[152] Hattenbach – Glucoft, above note …, pp. 36 and 51 et seqq. (noting however that many of these claims appear non-sensical, and that therefore their validity needs to be assessed on a case-by-case basis).

[153] See US patent No 5,659,666.

[154] Abbott, above note …, pp.1083-1086.

[155] See US patent No 5,852,815.

[156] Abbott, above note …, p.1081.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Indeed, in European law there is little clarity with regards to whether inventors need to be human beings. The European Patent Convention does state that applications need to mention the designation of the inventor[157], which must include among other information the inventor's family name, given names and full address[158]. This would suggest that the inventor is to be a natural person. Likewise, the UK Patents Act 1977 refers to natural persons many times[159]. Section 7(1), for instance, confirms that any *person* may make an application for a patent, with section 13(2) requiring an applicant to identify the *person* who is believed to be the inventor. It should also be reminded that in March 2013 the EU IPR Helpdesk published a Fact Sheet relating to inventorship, authorship and ownership, which interestingly noted: "…the inventor is always a *natural person* and the first owner" (emphasis added)[160]. While this statement does not expressly point out that machine and robot-generated inventions cannot obtain patent protection, it epitomises the importance of considering inventors as human beings.

On the other hand, one may also argue that no human inventorship requirement exists, at least in Europe, by pointing to the Guidelines for the examination of European Patents. Part A, Chapter III, Section 5(2) of such Guidelines refer to the possibility for inventors to waive the right to be mentioned as inventor: "The inventor designated by the applicant may address to the EPO a written waiver of his right to be mentioned as inventor in the published European patent application and the European patent specification, in which case his name is not mentioned"[161]. This guideline echoes what the Guide for European patents' applicants that choose the so-called PCT route[162] suggest in relation to the designation of the inventor: "It is recommended that the inventor always be identified … *unless there are special reasons for not doing so*" (emphasis added)[163]. These guidelines confirm that patent applications can be prosecuted, even until the final issuance of the patent, without designating any *human* inventor: which in turn may seem to indirectly suggest that patentable inventions could also be developed by non-human agents (the above emphasised sentence strengthens this argument).

That said, **it remains to be seen whether AI-triggered inventions satisfy the patentability requirements, especially the inventive step** (or non-obviousness) requirement. Indeed, no invention is patented if it is obvious to a person skilled in the art (as provided for example by Article 56 EPC and most patent statutes around the world). This requirement aims at raising the bar to getting patents, excluding from protection activities that are within the reach of the average expert of the field. Can then a machine

---

[157] EPC Rule 19(1); See also Article 41(2)(j) of the EPC confirming that the request for a grant of a European patent shall contain the designation of the inventor, where the applicant is the inventor.
[158] EPC Rule 19(1)
[159] Abbott, Big Data, above note …, p.7
[160] European IPR Helpdesk, 'Fact Sheet Inventorship, Authorship and Ownership' March 2013 see https://www.iprhelpdesk.eu/sites/default/files/newsdocuments/Fact-Sheet-Inventorship-Authorship-Ownership.pdf, at p.3.
[161] See the EPO webpage at https://www.epo.org/law-practice/legal-texts/html/guidelines/e/a_iii_5_2.htm.
[162] PCT stands for Patent Cooperation Treaty, an international treaty administered by the World Intellectual Property Organization (WIPO) which assists applicants in obtaining patent protection internationally for their inventions and helps offices, including the EPO, with their decisions to grant or refuse the patent.
[163] See the EPO webpage at https://www.epo.org/applying/international/guide-for-applicants/html/e/ga_b_18.html.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

or robot-generated invention satisfy this criteria? One may argue that - taking into account a machine's or robot's potentially high level of intelligence – the inventions reached by such machine or robot would often meet the threshold in question[164]. Take for example the question-answering computer system Watson, developed by IBM, which is capable of answering questions posed in natural language[165]. Ryan Abbott noted that some Watson's results have been so surprising to its creators that they may be considered non-obvious and therefore as meeting this requirement[166].

The attraction of machine and robot-developed inventions into the realm of patentable subject matter may also change the way the inventive step criteria is assessed by patent examiners and judges. Currently, patent officers and courts take into account, as benchmark, the "person skilled in the art", i.e. somebody who is considered as having good knowledge of the relevant prior art, and an understanding of whether the invention to be examined departs significantly from that existing body of knowledge. Yet, as interestingly noted by Ryan Abbott, **the acceptance of computational innovation by patent laws may trigger a "substitution" of the concept of "skilled (human) person" with the notion of "skilled computer", with the inevitable result that patenting inventions may become more difficult**[167]. Indeed, because of the way they have been programmed, machines and computers have an incredible extensive knowledge of the prior art, much broader than the knowledge a human being may have, even in fields not strictly related to those of the inventions to be analysed: this would likely raise the bar to obtaining patents. While stricter patent procedures are certainly to be welcomed (patent offices have often been accused of being too generous for granting patents for trivial inventions), a system where the benchmark becomes the "skilled machine" would also pose challenges. For example, patent examiners and judges would have "to put themselves in the shoes" of the skilled computer and take into consideration the prior art which may be within the reach of powerful machines, but not human beings[168]: which might turn out to be practically (and humanly) undoable.

Finally, issues of ownership of patent rights over the inventions developed by machines or robots should also be mentioned. We have seen above that applicants tend not to disclose the role of creative computers to patent offices, with (human) inventors/applicants ending up owning the economic rights to the patent. But if and when robots and machines will be considered as capable of autonomously coming up with patentable inventions (for example, when applications designating computers as inventors will be accepted), the question of who should be the owner of the patent will be crucial. As computers cannot (at least, yet) own legal rights, possible "candidates" for holding ownership rights would be: (i) the programmer of the AI software which produced the invention; (ii) the user of such program who gives the AI tasks; or (iii) even the owner of

---

[164] Yet, it has also been noted that certain mechanically generated claims may be considered obvious. For example, the linguistic manipulation software devised by the company Cloem (see above note …) often merely adds or deletes sentences. As noted by Hattenbach and Glucoft, above note …, p. 45, many of these claims are "the result of relatively slight rearrangements, and these minor modifications that work in predictable ways would by definition be considered obvious".

[165] See Watson related webpages at https://www.ibm.com/watson.

[166] Abbott, noted above …, pp.1091

[167] Abbott, noted above …, pp.1124-1125.

[168] Abbott, noted above …, pp.1125.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

the machine or robot themselves[169]. The issue is undoubtedly difficult to resolve - and **once more a case-by-case analysis aimed at finding out who has substantially contributed the most to the invention will likely be needed to identify the owner of the resulting patent**.

4. Conclusion

Patents, trade secrets, copyright, trademarks and designs are tools that robotics companies can use to attract, and recoup, the big investments that are needed in this field. **While there are phases within the life cycle of robotic firms where a cooperative and non-IP focused policy is better suited to support the growth of the venture, IP strategies based on the exploitation of proprietary rights are certainly key in shaping and strengthening this industry.** Just filing a patent, design or trademark application may not be enough – taking legal action against competitors that try to free-ride on the investments made by someone else is also increasingly necessary. We have mentioned some of these lawsuits, several of which have been settled with a consent judgement that has left the complainant that enforced its IP in a stronger position than the alleged infringer. Yet, some legal actions aimed at protecting IP in robotics have also been criticised, particularly where patents have been granted by patent offices with broadly drafted claims and aggressively enforced against competitors[170].

Robots, especially those embedded with AI, may also be capable of coming up with works or inventions that are usually created by human beings, being them painters, illustrators, poets or engineers. Whether these artistic, literary, musical or technical outputs meet the requirements for copyright or patent protection, and who the author or the inventor is, are thorny issues that are currently giving IP scholars and lawyers food for thought: issues that will become even more burning as technology keeps advancing and allows robots and machines to perform behaviours or tasks with a higher degree of autonomy. What we may be witnessing in the not so distant future is not only the continued progress of robotic technology which will increase AI capabilities, but also an evolution of copyright and patent regimes that will put at the centre of the debate the creation of "works" or the development of "inventions", possibly diminishing the roles of "authors" and "inventors". **Are we moving from the "law of authors and inventors" to the "law of copyrightable works and patentable inventions"?** The robotics and AI industry may be the innovative sectors that will provide an answer to these questions[171]. Time will tell.

In the meantime, it is recommended that a number of **new legal reforms, to be enacted at the European level, should be brought forward to bring a measure of clarity to this rapidly developing area**. In the areas of copyright and design law, where the EU has substantial competence, new legislative guidance on (i) whether interactive robot-created works of e.g. music, literature, art can be protected by copyright

---

[169] Abbott, above note …, p.1114.
[170] See article "The IP Battle Continues for Robotics Companies", in Robotic Business Review of 7 January 2018, at https://www.roboticsbusinessreview.com/legal/the_ip_battle_continues_for_robotics_companies.
[171] See also Massimo Maggiore, Artificial Intelligence, Computer-Generated Works and Copyright, in Enrico Bonadio – Nicola Lucchi, Non-Conventional Copyright – Do New and Atypical Works Deserve Protection? (2018 Elgar), p. ….

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

and (ii) if so, who should own these works. Similar reforms could be brought forward by the European Patent Office to clarify the circumstances regarding (i) whether an interactive robot-created invention is patentable or whether a human inventor is required; and (ii) who should be the owner of such a patent (e.g. the consumer/user of the robot, the developer of the robot).

4. Determining the impact of interactive robotics in the labour market

1. The Future of Work and the Digital Transformation: An Overview

The nature of work in modern economies is being transformed as a result of ongoing trends such as automation, urbanization and globalization, and workforce ageing, sparking public debate and concern over the future of work. This part of the Preliminary Report sketches an overview of these trends with a focus on digitalization, their implications for work and workers, and the societal challenges they bring. Lastly, we outline some avenues for policy discussion.

Recent decades have witnessed important changes in our labour markets, influencing the nature, quality, and productivity of work. Entrepreneurs, policymakers, and other thought leaders face the challenge of making use of these developments to foster economic growth, while at the same time ensuring decent working conditions, social protection, and equal opportunities for all. To add to an understanding of and discussion about these ongoing changes, this report outlines the current debate. The structure of this report is shown in Figure IRLESA 1.

In Section 3.2, we draw on an active scientific literature to highlight key trends that are shaping the future of work. These include automation, a force which is at the center of the debate on the future of work, but also the shifting geography of work through globalization and urbanization, and workforce aging. Section 3.3. highlights the chances and challenges these forces raise for our labour markets today and in the future, such as rising inequality, job reallocation and polarization, and skill gaps. To illustrate this, Section 3.4. estimates the number jobs that are expected to be lost and created due to digitalization by 2030 in advanced economies. Section 3.5. then connects these to three broad societal themes along which thinking about the future of work can be organized: inclusion, decent work, and skills. Lastly, Section 3.6 provides starting points for a policy discussion on how to address these challenges, aimed at safeguarding the three themes of labour market inclusion, decent work, and skill acquisition.

2. Key trends impacting the future of work

The future of work is shaped by many factors, from technological to institutional and cultural. Here, we outline three key driving forces that are impacting our labour markets today and into the future: technological progress and automation (section 1.1); globalization and urbanization (section 1.2); the rise of new work forms (section 1.3); and population ageing (section 1.4).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

a. Technological progress and automation

Chief among the trends impacting current and future labour markets is the progression of innovations that imbue machines with ever-increasing capabilities. While automation is by no means a new phenomenon, the digital revolution has brought its own set of labour market impacts through the introduction of desktop computing, the Internet, and other Information and Communication Technologies (ICT). These impacts are expected to continue into the future from further developments in robotics, as well as in machine learning and other Artificial Intelligence (AI) technologies.

It can be useful to classify these digital automation technologies in two broad types: firstly, rules-based technologies, and secondly, prediction-based technologies. The former are technologies which automate tasks by codifying them into a series of if-then statements, which are then embodied in software. As such, these technologies can be used to automate tasks that follow a set protocol, sometimes referred to as "routine" tasks (Autor, Levy and Murnane 2003). Examples of such routine (or codifiable) tasks are performing calculations and assembling products in an assembly line. Prediction-based technologies, on the other hand, use big data and machine learning techniques to predict likely outcomes (Agrawal, Gans, and Goldfarb 2018). These technologies belong to the class of so-called Artificial Intelligence (AI). An example of AI is machine translation: algorithms trained on large databases of existing translated texts can learn to predict how a particular sentence would be translated in another language, without having encoded this translation in a set of explicit if-then rules. Similarly, machine vision can learn to recognize images of cats from on a database of pictures previously tagged as containing cats or not: as such, AI can tell us whether a new picture contains a cat, without having been given explicit rules on what makes a cat cat-like.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

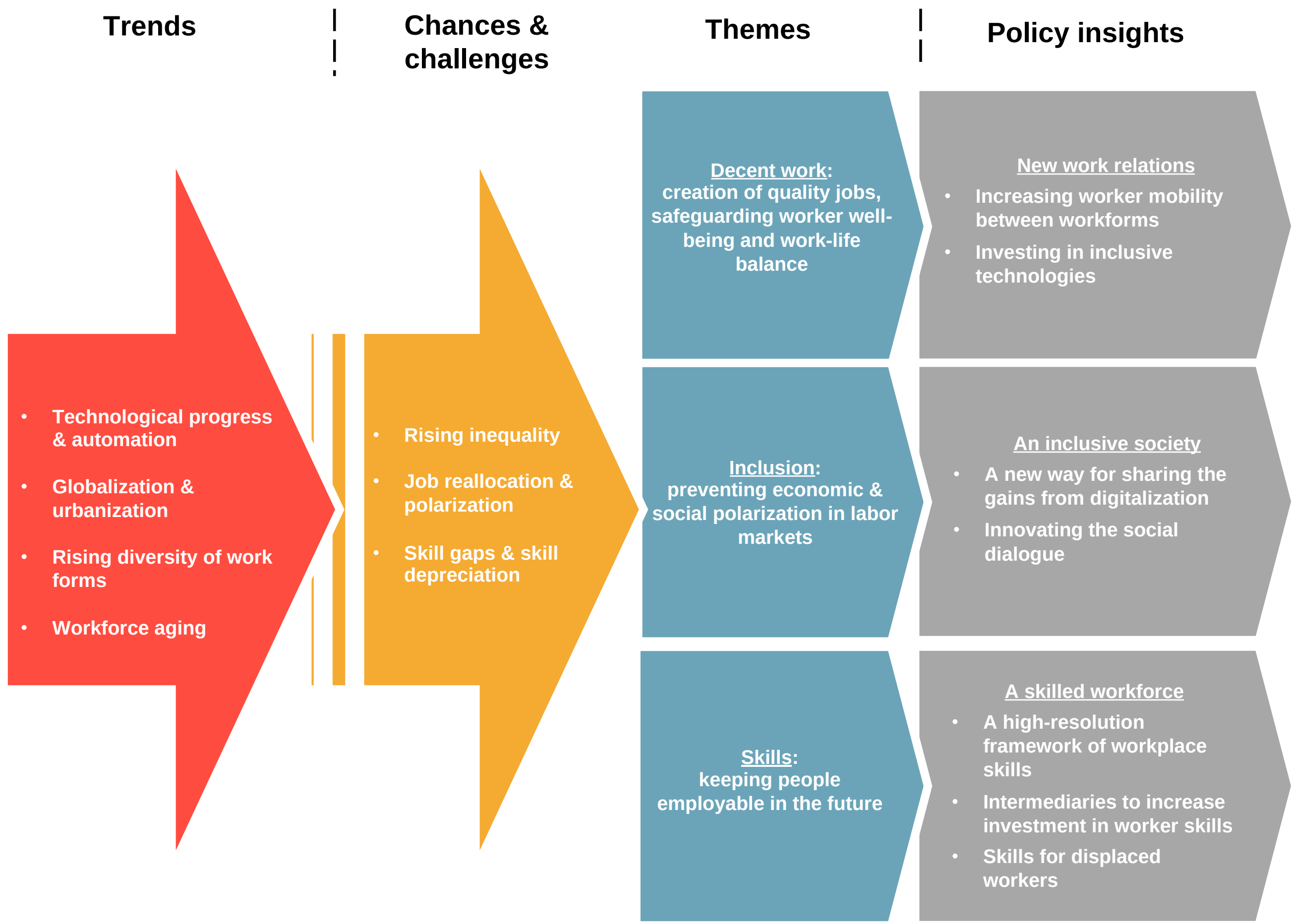


FIGURE IRLESA 2 THE FUTURE OF WORK: A ROADMAP

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Both types of technologies aim to automate tasks, as well as improve their speed, quality, and/or diversity relative to what human workers can produce unaided by technology. The main difference is that unlike rules-based technologies, prediction-based technologies do not require a precise description of the separate steps that need to be performed to complete the task. Rather, prediction-based technologies learn about how to automate a task by first observing how it has been performed previously (using what is called "training data"), usually by humans, and then replicating the outcome using a statistical model. This is why prediction-based technologies can go beyond the automation abilities of rules-based technologies: given enough data and a clear determination on the outcome to predict, they can flexibly perform processes that are highly variable. Further, these technologies can learn over time (hence "machine learning"), increasingly calibrating their decisions to the desired outcome.

For example, AI is being used to classify credit card charges as fraudulent. A machine learning algorithm uses a training database of existing credit card charges (both fraudulent and legitimate) and their characteristics (such as the expenditure amount, geolocation, time of the expense, as well as the previous expenditure history of the card), to predict in real-time whether charges should be flagged or not. Driverless cars are an example of an emerging AI technology: based on previous data from human drivers, online maps, and aided by machine vision, driverless cars analyse road situations and predict the best driving decision in real time.

In both cases, this is very different from an algorithm that would pre-program all possible credit card expenditures or road situations and what to do in each specific case. For one, there may simply be too many situations to program; and new ones may emerge which have not been pre-programmed and would therefore cause the automation to fail. Further, humans often make decisions without being able to summarize their reasoning in a set of explicit rules. This phenomenon is known as Polanyi's Paradox, also described as "we know more than we can tell". For example, a bank employee reviewing credit card charges may not be able to exactly explain their intuition on what makes a charge suspect – but such a rulebook is not needed for AI. Similarly, there is no way to explicitly encapsulate the task of driving into a finite set of pre-determined rules.

However, it would go too far to consider prediction technologies as equivalent to human intelligence. While they may be cheaper and better at predicting outcomes, the limits of these technologies lie in their inability to perform judgment the way humans do (Agrawal et al 2018). In particular, **humans rely on causal reasoning to make decisions, but AI only observes patterns and correlations in the data without separating cause from effect**.

Evidence on the advances of these two types of technologies is not hard to find. Figure IRLESA 3 illustrates how the number of transistors on integrated circuit chips has doubled roughly every two years, a phenomenon known as Moore's Law. This is seen as a marker of technological progress since Moore's law predicts concomitant changes in computing power and processing speed. These advances have allowed both rules-based

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

algorithms as well as machine learning based prediction to perform tasks more quickly and cheaply.

This computer revolution has transformed workplaces. Firstly, digital technologies have replaced humans in some tasks and made them more productive in others – this is discussed in section 3.3. In the most direct sense, the prevalence of computer uses at work increased from one-quarter of the workforce in 1984 to nearly one-half in 1993 (Autor, Katz, and Krueger 1998). More recent business surveys across the OECD report show that around 59% of firms' employees regularly use a computer for work in 2018, compared to 47% in 2005: Figure IRLESA 4 depicts this upward trend, which is visible for firms of all different sizes, though most strongly for larger firms.

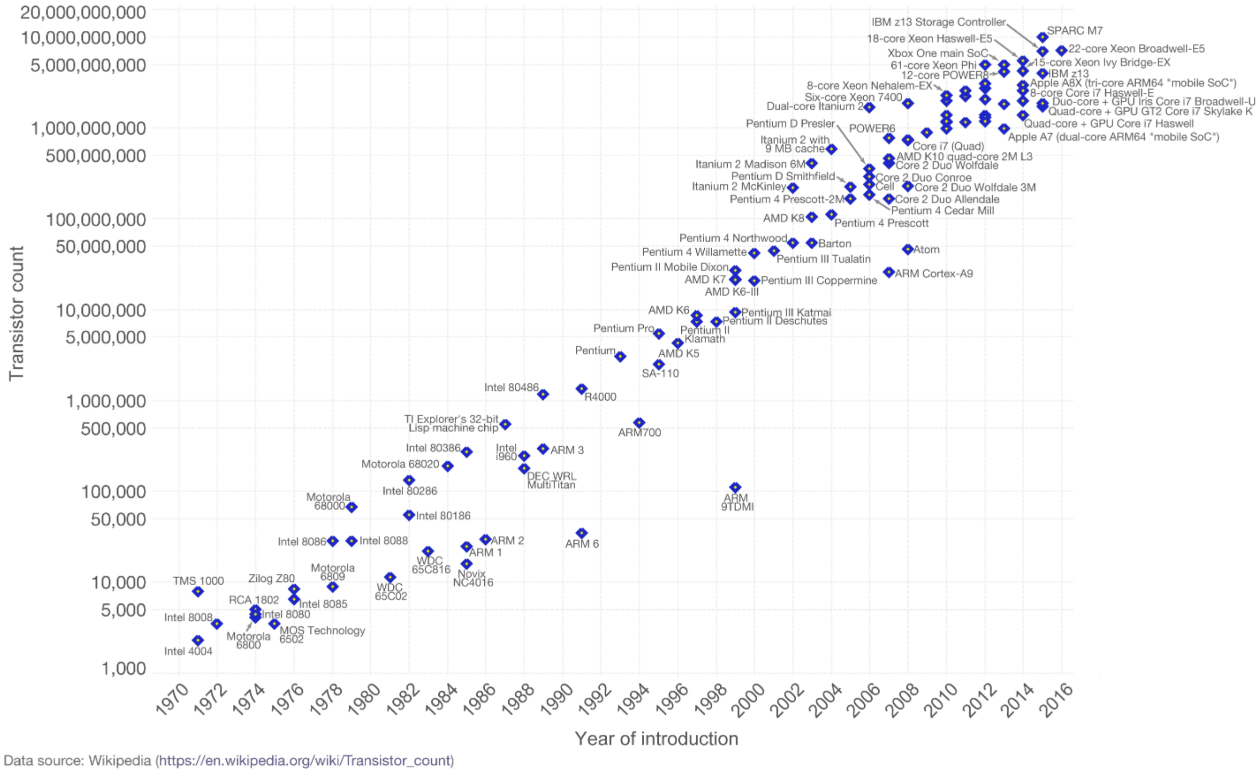
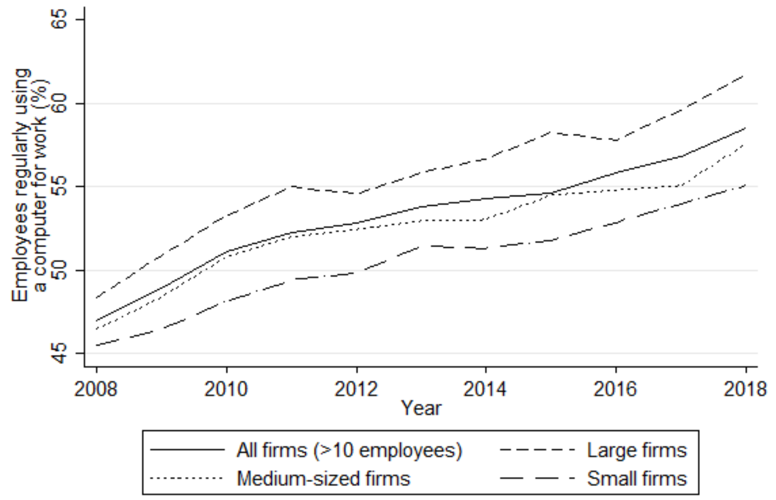


FIGURE IRLESA 3 MOORE'S LAW, 1971-2016

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073



Large: >250 employees; Medium-sized: 50 to 249 employees; Small: 10-49 employees
Source: Authors' calculation based on OECD Statline, ICT Access and Usage by Businesses.

FIGURE IRLESA 4 . COMPUTER USE AT WORK

Among recent technologies impacting work, robotics has garnered a disproportionate amount of public attention, as it arguably represents the leading edge of workplace automation. While most countries as yet only have a relatively small number of robots per 10,000 workers, as shown in Figure IRLESA 5; Figure IRLESA 6 illustrates that workplace robot adoption has grown markedly in recent years. Specifically, the International Federation of Robotics documents that average growth in robot adoption has been 16% per year over the past decade and predicts the operational stock of industrial robots to increase from 2,1 million in 2017 to around 3,8 million in 2021. Service robots are much less prevalent, though their adoption is also growing, particularly in the logistics sector, followed by healthcare applications (International Federation of Robotics 2018).

A final important advancing technology is Artificial Intelligence (AI): as shown in Figure IRLESA 7, patenting in AI is on the rise in recent years, and AI patent growth has regularly exceeded other patent growth. This boom is to a large extent due to a fast increase in patenting in machine learning, a type of AI that is based on algorithms performing pattern recognition and prediction tasks using large datasets.

Advances in computing and AI have decreased the price of computation and prediction, leading firms to adopt smart algorithms across many domains of business practice, including

Artificial Intelligence in human resource planning

While HR planning software usage is already high, its range of applications is expanding through the use of AI. For instance, the New York Times (2018) reports how a tech start-up called Humu has developed a data-driven approach to increase employee happiness. It analyses employee surveys to identify behavioral changes that are likely to make the biggest impact on elevating workforce happiness, and then uses emails and text messages to "nudge" individual employees into small actions that advance the larger goal. This nudging is based on economist Richard Thaler's Nobel-prize winning research on how to incite people to make better choices. Similarly, IBM has patented a "predictive attrition program" which was developed with its AI-based supercomputer Watson to predict employee flight risk and prescribe actions for managers to engage employees (CNBC 2019).

This project has received funding
esearch and innovation program

enterprise resource planning and customer service relations, as shown in Figure IRLESA 8**¡Error! No se encuentra el origen de la referencia.**. Enterprise resource planning has increased especially markedly over the past decade: some 35% of firms across OECD countries report using this in 2017, as opposed to 20% some 10 years ago.

While enterprise resources are a broad term that covers a large set of potential software applications, one that is of particular relevance is the use of algorithms in planning human resources. **Digital technologies do not only change the nature of jobs and the skills they require, but also impact the functioning of labour markets directly.** For example, online job search is the norm – in 2011, around three-quarters of the unemployed searched for jobs online, compared to only one-quarter in 2000[172]. More recently, algorithms have been employed to screen resumes, match workers to jobs (including providing unemployed jobseekers with algorithmic search recommendations), aid firms in recruitment and hiring, provide input for job performance and promotion decisions, predict employee turnover, and nudge workers towards behaviours that are estimated to increase employee motivation and happiness (see inset). Such developments would not only increase employee welfare, but also have the potential to raise worker productivity – research has demonstrated a causal link between happiness and productivity[173].

For instance, automated resume-screening is estimated to occur in over 90% of large companies[174]. While is this is in part a response to increases in application volumes which have occurred with the rise of online job boards (as many companies receive 100 or more applications for a single position); it also reflects HR cost-saving. Even at small companies, only 19% of hiring managers.

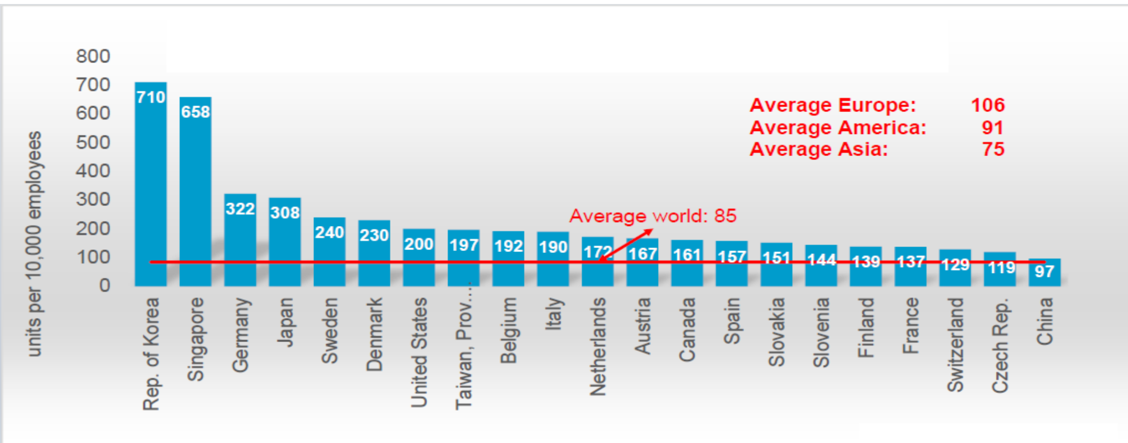


Figure IRLESA 5 Industrial robots per 10,000 manufacturing workers

Notes: Data are for 2017. Source: International Federation of Robotics (2018).

---

[172] Faberman, J. and Kudlyak, M.. "What Does Online Job Search Tell Us About the labour Market?" *Economic Perspectives, Federal Reserve Bank of Chicago*, (2016).

[173] Oswald, A.J. et al., "Happiness and Productivity," Journal of labour Economics 33 (2015): 789.
[174] The Wall Street Journal. 2012. "Your Resumé Vs. Oblivion,"
https://www.wsj.com/articles/SB10001424052970204624204577178941034941330

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

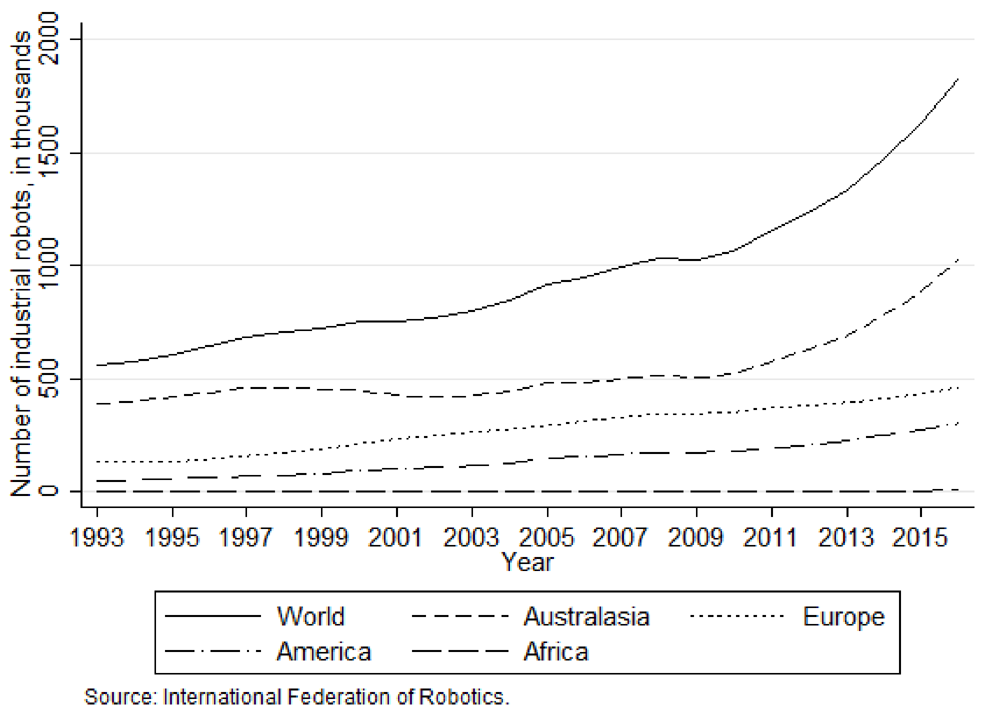


FIGURE IRLESA 6 . STOCK OF INDUSTRIAL ROBOTS OVER TIME

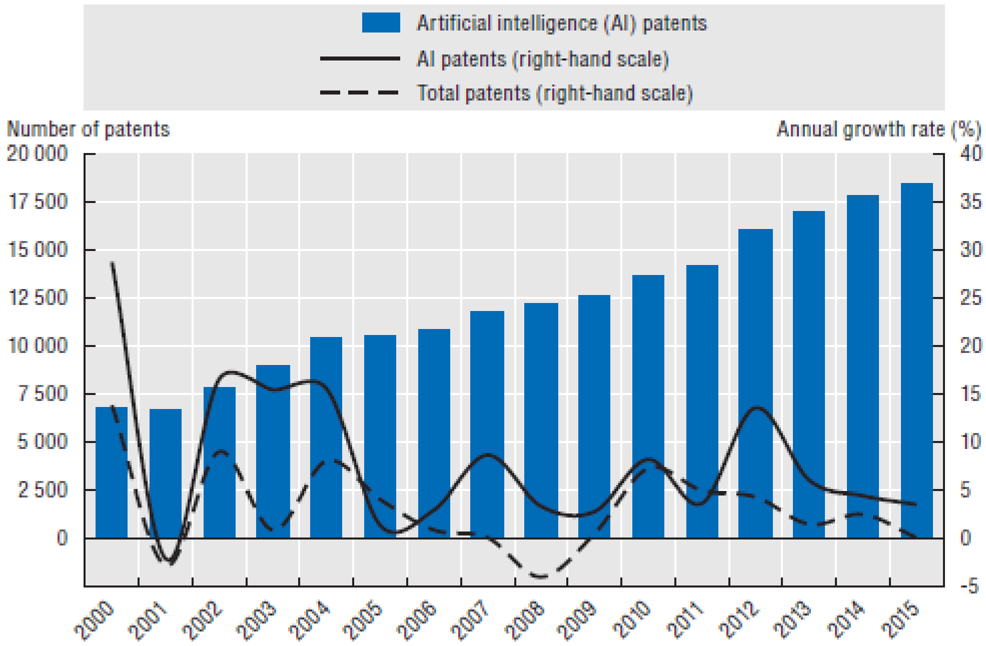


FIGURE IRLESA 7 PATENTS IN ARTIFICIAL INTELLIGENCE TECHNOLOGIES, 2000-2015

Source: OECD (2017a).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

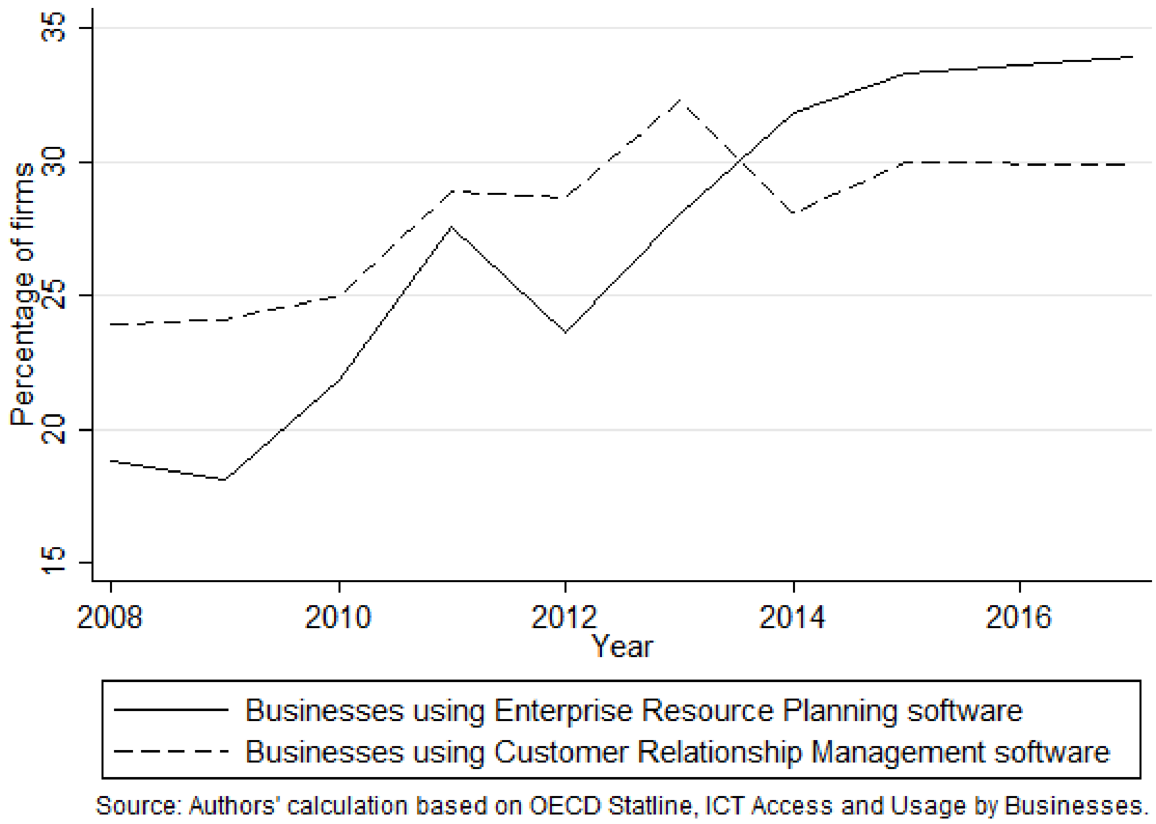


FIGURE IRLESA 8 EXAMPLES OF FIRMS' USE OF BUSINESS SOFTWARE

state they look at a majority of the resumes they receive, and 47% say they review just a few, according to a recent survey by Information Strategies Inc., publisher of Your HR Digest.

As with other AI-based technologies, applications are widening besides recommending CVs from an applicant pool for an interview, advanced HR software can suggest applicants by scraping online networks, as well as conduct initial interviews. Indeed, 96% of senior HR professionals believe AI has the potential to greatly enhance talent acquisition and retention [175]. For example, **chat-bots and physical robots have been developed that performs in-" person" HR interviews[176]. These technologies aim to offer a human-like experience without the unconscious biases that humans are prone to.** To that end, all questions are posed in an identical way, in the same tone, and typically, in the same order – and human recruiters or managers are then given text transcripts of each interview to help them decide which candidates should move to the next stage of the process, based on the answers alone. Other companies are going even

[175] Randstad. 2018b. "How Will Artificial Intelligence Affect Your Talent Acquisition Strategy?" Workforce Insights, Future of Work. https://www.randstad.com/workforce-insights/hr-tech/how-will-artificial-intelligence-affect-your-talent-acquisition-strategy/
[176] Randstad. 2018c. "Recruitbot: Why AI Will Hire Your Next Employee." Workforce Insights, Future of Work. https://www.randstad.com/workforce-insights/hr-tech/recruitbot-why-ai-will-hire-your-next-employee/

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

further and are also using AI to **recognize and categorize applicant emotions, a practice that has raised concern**[177].

b. Globalization and urbanization

Another key labour market trend is the changing geography of work, both internationally and within countries: here, we discuss these in turn.

Globalization causes international shifts in production, leading to a reallocation of work across borders: it is therefore an important force shaping the future of labour markets. Indeed, across the OECD, around 40% of business sector workers are estimated to be engaged in production to meet final demand originating in other countries[178] (OECD 2017a, OECD 2019), including as supply chain (sub)contractors. While 40% is the average, this percentage is of course higher for small open economies than for larger countries which have sizable home markets.

Globalization in part results from advancing technologies allowing production processes to be unbundled and work product to be delivered electronically, and partly as from declining man-made barriers to trade. One of the most significant recent changes in the global economy that has been scientifically studied is the rapid emergence of China from a technologically backward and largely closed economy to the world's third largest manufacturing producer. This shift occurred over just two decades: as shown in Figure IRLESA 10, the share of world manufacturing exports originating in China grew from 2% in 1991 to 16% in 2011[179]. While the resulting import competition from such trade integration produces large gains for consumers in the form of lower prices, it also leads to adjustments in home country labour markets.

[177] The Guardian. 2019. https://www.theguardian.com/technology/2019/mar/06/facial-recognition-software-emotional-science
[178] OECD-International Transport Forum (2018): Safer Roads With Automated Vehicles?, 2018 [Documento electrónico disponible en: https://www.itf-oecd.org/safer-roads-automated-vehicles-0]
[179] Autor, D.H. et al., "The China Shock: Learning from labour Market Adjustment to Large Changes in Trade." Annual Review of Economics, (2016).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

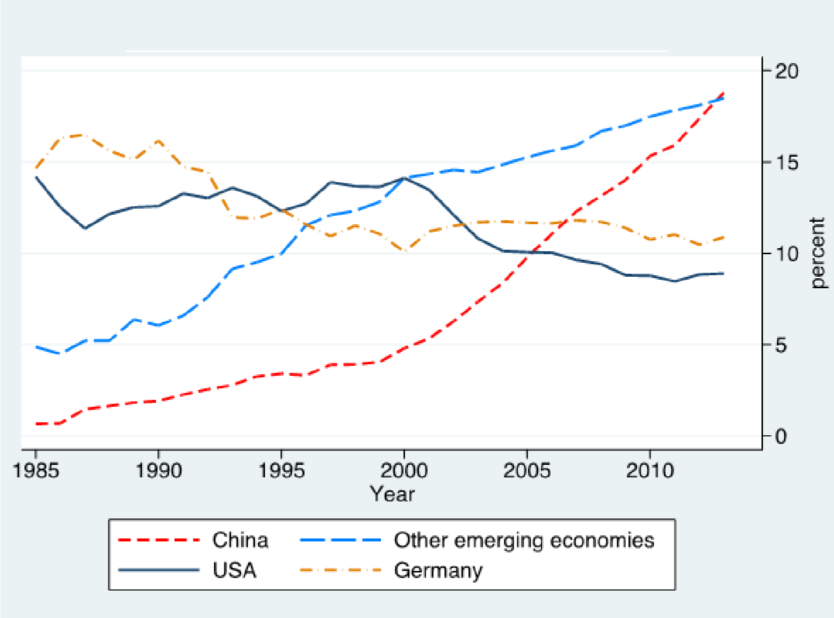


FIGURE IRLESA 9 SHARE OF WORLD MANUFACTURING EXPORTS

Source: Autor, Dorn, and Hanson (2016)

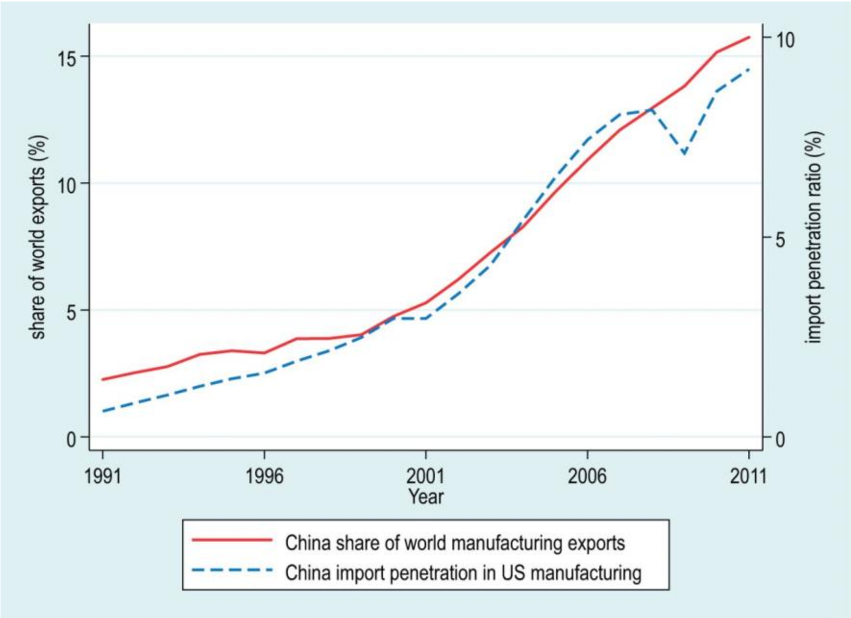


FIGURE IRLESA 10 RISING TRADE WITH CHINA

Source: Autor, Dorn, Hanson and Song (2014)

Lastly, some commentators have argued that certain types of emerging technologies may reduce the importance of global value chains in the future. For example, **advances in robotics and artificial intelligence are seen as a way to re-shore labor-intensive**

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

production from low labour cost countries back to developed countries[180] – however, while actual data on this phenomenon is as yet very scarce and unrepresentative, current reports suggest re-shoring is **only a small phenomenon** compared to offshoring[181].

Besides international changes in the allocation of work, the regional distribution and nature of jobs has also shifted *within* countries' borders: a key trend here is urbanization (Randstad 2018a). A large body of research documents how high-skilled workers have strongly moved to more densely populated (i.e. urban) areas over the last 35 years. Figure IRLESA 11**¡Error! No se encuentra el origen de la referencia.** illustrates the rising urban gradient in college degree holding that has been widely documented in the research literature. In 1970, working-age adults in the most densely populated regions were approximately 5 percentage points more likely to hold a college degree than those in the least densely populated regions. This gap rose to 15 percentage points between 1970 and 1990, and by, 2015, it had risen further to approximately 25 percentage points. No such urban-rural divergence is found in the location of the least-educated adults, high-school dropouts. As a result, the educational distribution in urban areas has become increasingly skewed towards higher educated workers. As shown in Autor (2019), this pattern is not driven by immigration patterns of foreign workers but is observed equally among natives.

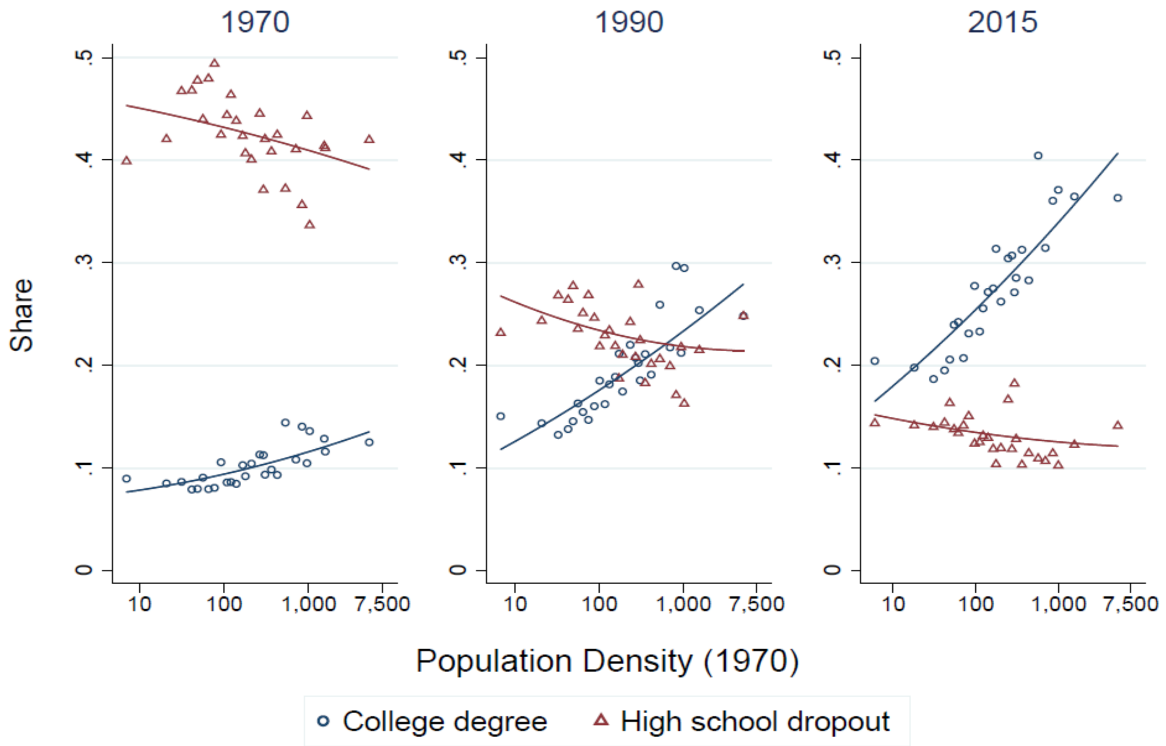


Figure IRLESA 11 The changing geography of work

[180] McKinsey Global Institute - MGI (2017). "Jobs lost, jobs gained: workforce transitions in a time of automation". Technical Report.
[181] Baker, G. and Hubbard, T. "Make Versus Buy in Trucking: Asset Ownership, Job Design, and Information", The American Economic Review, 93 (2003): 551.

Notes: Figure plots the US share of working-age adult residents by region who have either four-plus years of college or less than a high school degree. Each plotted point represents approximately 5 per cent of the working-age population in the relevant year. Source: Autor (2019).

Indeed, Autor and Salomons (2019)[182] show that this regional trend is further reinforced by the emergence of new work – those jobs which arise as the direct result of advancing technologies. Figure IRLESA 12**¡Error! No se encuentra el origen de la referencia.** documents how the set of high-paying new jobs which make use of novel technologies **(so-called "frontier work") has increasingly shifted towards urban centres over the past decades**. Current examples of such jobs are robotics developers, artificial intelligence researchers, wind turbine technicians, and search engine optimization experts. Examples from previous decades are word-processing supervisors, and controllers of remotely piloted vehicles (1980s); circuit layout designers, and robotic machine operators (1990s); and echo-cardiographers, molecular physicists, and programmer-analysts (2000s). The striking change is that these jobs were almost equally likely to be found in rural and urban areas in 1980, but by 2015 frontier jobs are strongly overrepresented in urban centres.

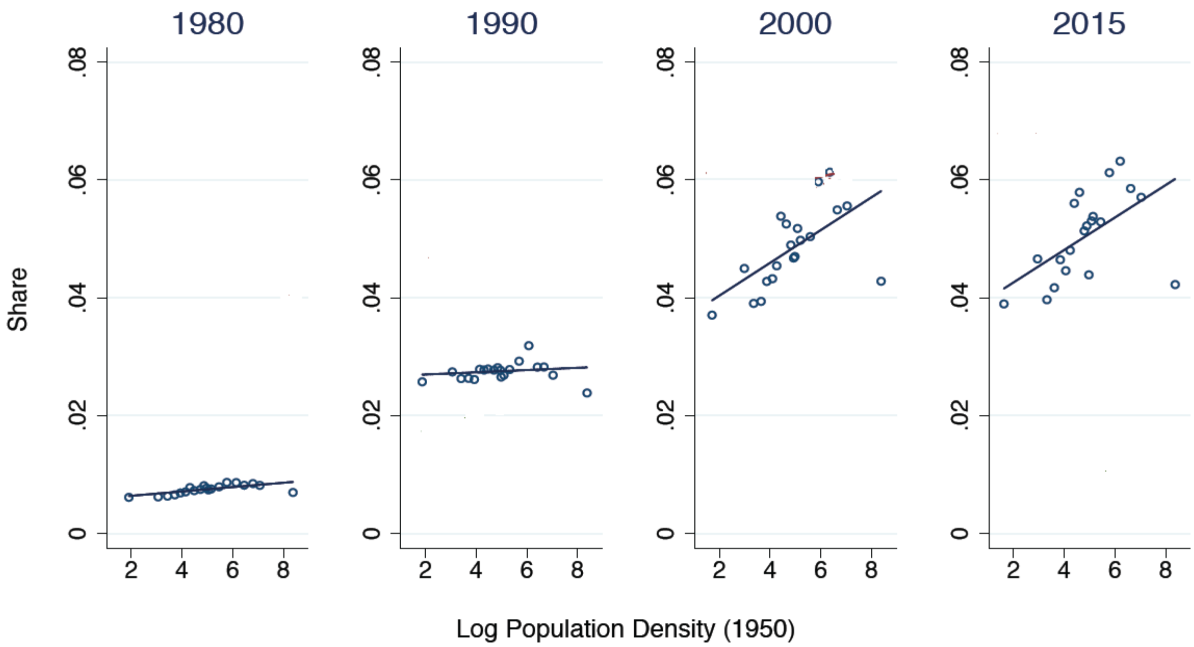


FIGURE IRLESA 12 NEW HIGH-TECH JOBS BY POPULATION DENSITY AND DECADE

Notes: The figure plots cumulate employment shares in new jobs for US Census occupations, among working age adults. Source: Autor & Salomons (2019).

182

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

c. The rise of new work forms

Much evidence indicates that there is an increase in new forms of work that differ from the traditionally large group of full-time workers with permanent contracts. These new work forms include:

- *Part-time workers* work less than full-time and have a permanent or temporary contract.
- *Agency workers* have a contract with a temporary work agency, and they are supplied to a company for a set period of time.
- *Contract workers* allow employers to hire staff with no guarantee of a regular work times or a set number of work hours. Contract work comes under different contractual forms. An example are on-call workers, who are expected to be available at any time, usually with short notice. Another example are zero-hours contracts for which the employer has no obligation to provide a set number of hours of work.
- *Self-employed workers* or *freelancers* are contract workers who are also the owners of a business and work for different companies at different times.
- The *gig economy* is a labour market characterized by the prevalence of contract workers. It is often referred to in the context of two-sided digital platforms that bring together supply and demand for set tasks or projects (which can take place online or offline). In this case, it is called the 'online gig economy'.

The online gig economy in Europe

The online gig economy in Europe is increasing and is now the main source of income for as many as 2% of adults across 14 EU member states, according to European Commission survey data (Pesole et al. 2018). This includes transport, delivery, care, and other on-location work mediated by 'gig economy' apps, as well as software development, translation, data entry, and other knowledge work delivered remotely via the platform. Although in absolute terms European employers are not yet hiring very actively on online labour platforms, their use of such platforms is growing (Kässi and Lehdonvirta 2018).

New work forms have been on the rise in much of the developed world. In the US, new work forms rose from 10.7% in 2005 to around 15% in late 2015[183]. This documented increase is attributable to growth in contract work and in the self-employed, while agency work remained stable with a share of around 2% of the working age population[184]. For Europe, the changes are qualitatively similar: part-time and temporary employment have increased from 12.5 to 15.8% in the

[183] Katz, L.F., and Krueger, A.B. "The Rise and Nature of Alternative Work Arrangements in the United States, 1995–2015." Industrial & labour Relations Review 72 (2019), 382
[184] World Employment Confederation. 2018. "Economic Report," https://www.wecglobal.org/fileadmin/media/pdf/WEC_Economic_Report_2018_Edition.pdf

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

European Union since 2002[185]; and self-employment grew significantly over the same period in some EU member states such as the Netherlands and UK[186].

To illustrate this further, Figure IRLESA 13 shows the prevalence of "standard" (employees working >30 hours holding a permanent contract) and new work forms ("self-employed" and the all other new forms of work captured by "atypical") in the total population aged 15-64 for the EU as well as several individual countries. The figure illustrates the importance and pervasive rise of new work forms. At the same time, it can be seen that on average there has been no decline in the importance of the traditional employment relationship. Rather, its share is still close to 40% of the working-age population in the EU-28. As such, the increase in new work forms has not always come at the expense of traditional work, but sometimes also at the expense of non-work (i.e. unemployment and inactivity) and informal work.

A particularly salient type of new work form has been reflected in the rise of online work platforms, which act as a clearing place for employers and workers to meet. In the online gig economy, there is a distinction between platforms acting merely as online intermediaries; and platforms where work is performed remotely and delivered digitally, such that employers and workers need not meet in person. Examples of the former are online intermediaries such as Monster.com; as well as online platforms specializing in specific in-person services, including Uber, Lyft, Deliveroo, and Helpling. Examples of types of platforms where work is performed at a distance are Upwork, TaskRabbit, Toptal, and Catalant.

As yet, the importance of platform-mediated work in developed economies appears to be small: in the US, a 2017 Bureau of labour Statistics survey suggests some 1% of jobs fall into this category[187] (see Figure IRLESA 14**¡Error! No se encuentra el origen de la referencia.**), and European Commission survey data estimates these jobs are the main source of income for around 2% of adults across 14 EU member states[188]. However, as illustrated in Figure IRLESA 15**¡Error! No se encuentra el origen de la referencia.**, which presents an index of online platform job postings, the usage of such platforms has risen sharply in recent years, and more so than other types of new work forms[189]. Specifically, it shows these job postings have increased by about 30% world-wide from May 2016 to January 2019, and that this growth is driven by developed countries.

---

[185] Rhein, T. and Walwei, U., Forms of Employment in European Comparison, https://www.iab-forum.de/en/forms-of-employment-in-european-comparison/
[186] Chiarli, T., Marzucchi, A., Salgado, E., Savona, M. 2018. "The Effect of R&D Growth on Employment and Self-Employment in Local Labour Markets." SPRU Working Paper Series 2018-08, SPRU - Science Policy Research Unit, University of Sussex Business School.
[187] Appelbaum, Kalleberg, A., and Hye, J.R. 2019. "Nonstandard Work Arrangements and Older Americans, 2005–2017." https://www.epi.org/publication/nonstandard-work-arrangements-and-older-americans-2005-2017/
[188] Pesole, A., Urzí Brancati, M.C, Fernández-Macías, E., Biagi, F., González Vázquez, I. 2018. Platform Workers in Europe, Publications Office of the European Union, https://bit.ly/2N2TciX
[189] Spreitzer, G.M., Cameron, L., and Garrett, L. Alternative Work Arrangements: Two Images of the New World of Work. Annual Review of Organizational Psychology and Organizational Behavior 4 (2017), 473

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

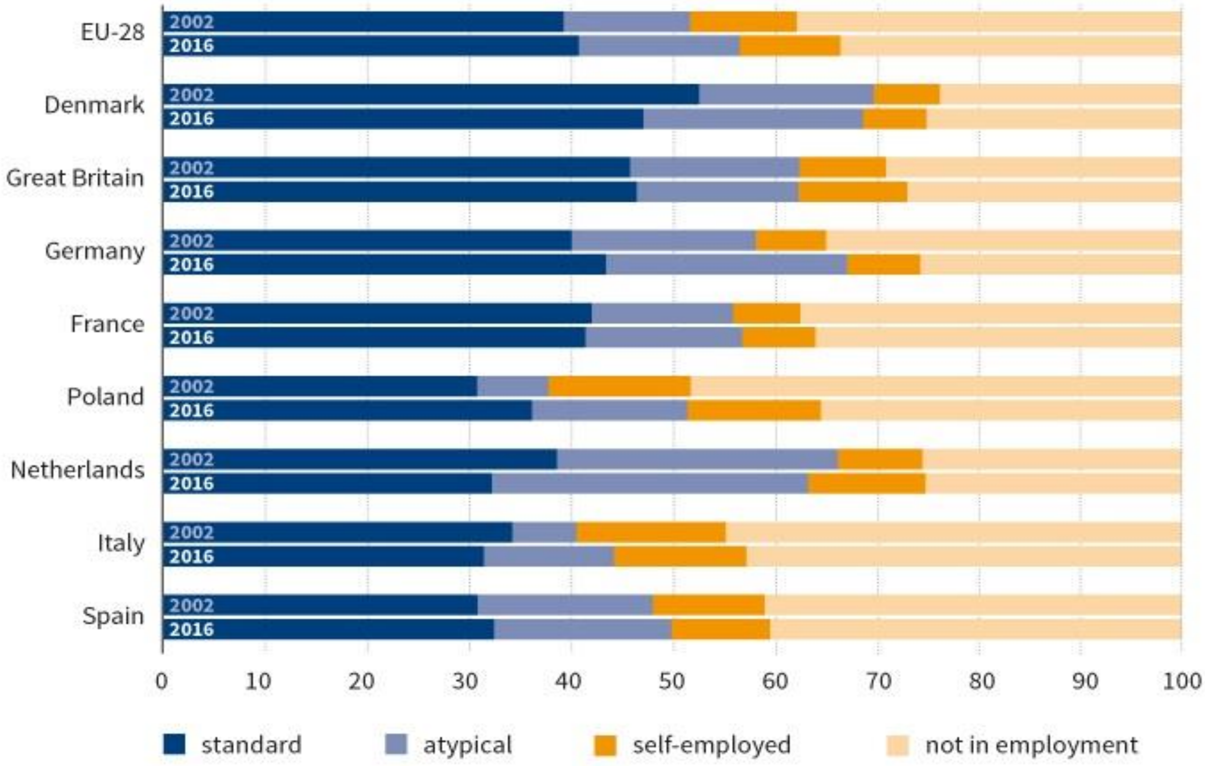


FIGURE IRLESA 13 NEW WORK FORMS IN EUROPE

Notes: Working-age population by labour force participation and type of main employment. Employees are defined as being in "standard" employment when they are working full-time or nearly full-time (> 30 hours) and hold a permanent contract. Source: Rhein and Walwei (2018).

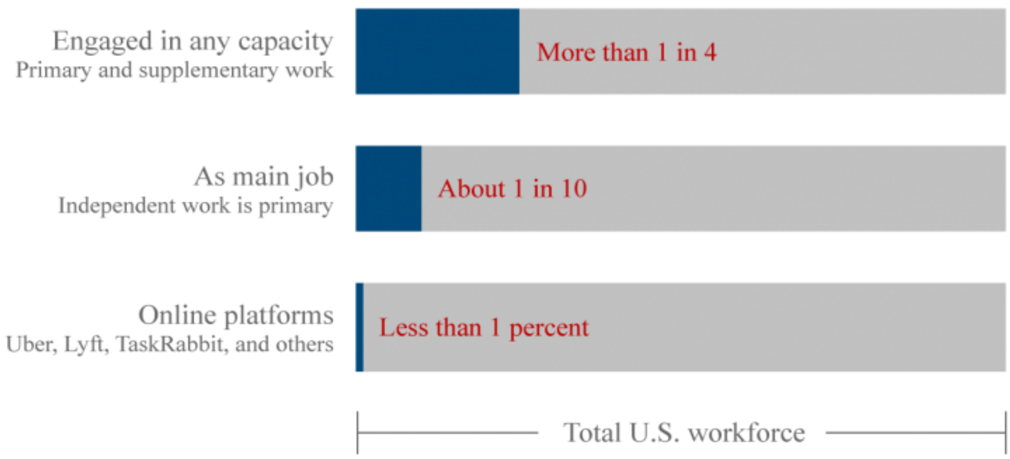


FIGURE IRLESA 14 HOW MANY WORKERS ARE EMPLOYED IN NEW WORK FORMS?

Source: ILR School & Aspen Institute.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

As documented by Kässi and Lehdonvirta (2018)[190], not all jobs are equally represented on five major online platforms (Upwork.com, Freelancer.com, Peopleperhour.com, Mturk.com, and Guru.com). Indeed, **the highest demand is for software development and technology skills, with roughly one third of online platform vacancies belonging to that category.** Other jobs frequently offered on these platforms are creative and multimedia work, as well as clerical and data entry work. Not coincidentally, these are jobs that are also typically subject to (domestic and international) outsourcing: they can be relatively easily described as stand-alone tasks and require relatively little tacit communication and knowledge of the client's local institutional environment, which distant online service providers may not possess. As a result, professional services such as human resources, project management, consulting, and legal services are so far only rarely being contracted on online platforms (corresponding to only 2% of the total online platform market).

While the term "new work forms" is used to describe a diverse set of alternatives to full-time employee work, many of these forms have been around for some time – for example, flexible staffing arrangements were used in 90% of firms in the 1980s, already (Abraham 1988). In surveys, the large majority of firms indicate such arrangements are critical for absorbing workflow fluctuations.

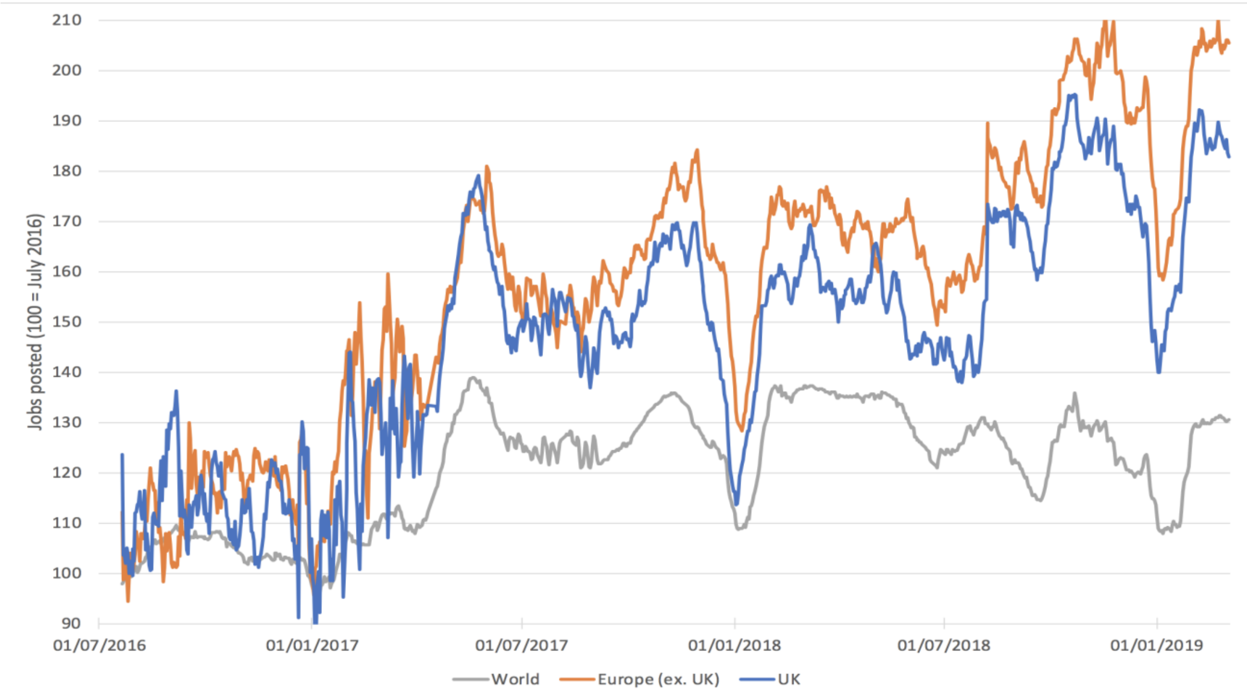


FIGURE IRLESA 15 EMPLOYERS' USE OF ONLINE PLATFORMS

Source: Kässi and Lehdonvirta (2018)

A range of explanations have been posited for the rise of new work forms, including technological changes and globalization that have standardized work and reduced

[190] Kässi, O., and Lehdonvirta, V. "Online labour index: Measuring the online gig economy for policy and research." Technological Forecasting and Social Change 137 (2018): 241–248. https://ilabour.oii.ox.ac.uk/online-labour-index/

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

monitoring and supervisory costs; a demographic shift toward an older workforce with older workers more likely to be self-employed; and a weak labour market in the wake of the Great Recession leaving workers with little bargaining power. As it turns out there is some truth to all of these: for example, the rise of online work platforms is a combination of advancing technologies for digitally delivering and monitoring work as well as business outsourcing practices driven in part by globalization. However, none of these factors appear able explain a quantitatively significant portion of the rise of new work forms[191]. Some have interpreted the secular rise in these forms as a response by companies to external market increases in skill differentials and wage inequality that raise the costs of compensation compression within a single employer[192].

On the other hand, there is evidence that some workers prefer the flexibility and associated work-life balance (e.g. in terms of scheduling and working from home) offered by new work forms[193]. It is likely that the demand for such flexibility has increased as women entered the labour force in growing numbers and families have increasingly relied on more than one earner, as this requires scheduling care of children and elderly parents. Further, a rising group of workers is pursuing (part-time) education or training while also working. These secular changes all inspire a need for more flexibility in the workplace, which is in part offered by new work forms[194].

d. Workforce ageing

A further trend that is important to consider from the perspective of the future of work are changes in the supply of workers. In recent history, some of the most important labour supply changes have been driven by increases in the average education level of workers and by the increase of female labour participation. These changes have been extensively documented and studied and are shown to have had positive effects on productivity and economic growth. In the future, changes in terms of educational attainment and female participation are not expected to be as large as they have been in recent decades.

However, another pervasive and still ongoing change across the developed world is the changing age structure of our populations. Figure IRLESA 16**¡Error! No se encuentra el origen de la referencia.** shows OECD historical and projected old-age dependency ratios over 1950-2060, where this ratio is defined as the population aged over 65 years

[191] Katz, L.F., and Krueger, A.B. "The Role of Unemployment in the Rise in Alternative Work Arrangements." American Economic Review 107 (2017): 388; Katz, L.F., and Krueger, A.B. "The Rise and Nature of Alternative Work Arrangements in the United States, 1995–2015." Industrial & labour Relations Review 72 (2019), 382.
[192] Weil, D. 2014. The Fissured Workplace: Why Work Became So Bad for So Many and What Can Be Done to Improve It. Cambridge, MA: Harvard University Press.West, S.M., Whittaker, M. and Crawford, K. 2019. "Discriminating Systems: Gender, Race and Power in AI." AI Now Institute. https://ainowinstitute.org/discriminatingsystems.html.
[193] Mas, A. and Pallais, A. "Valuing Alternative Work Arrangements." American Economic Review, 107 (2017): 3722-59.
[194] Council of Economic Advisors. 2010. Work-Life Balance and the Economics of Workplace Flexibility. https://obamawhitehouse.archives.gov/files/documents/ 100331-cea-economics-workplace-flexibility.pdf

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

old as a percentage of the population aged between 15 and 64. The data illustrate that OECD countries have witnessed a marked increase in the size of older populations relative to working-age populations, and that this trend is even stronger in Europe than in the US.

An ageing workforce may have unexpected interactions with other key trends highlighted above: in particular, **countries with older workforces may be more rapid adopters of automation technologies**. Advances in automation technologies are often viewed as the result of the inexorable march of technology, yet the development and adoption of these technologies is in part being boosted by demographic changes throughout the world[195]. Figure IRLESA 15illustrates this point: there are typically more robots per 1,000 workers in countries with more rapidly ageing populations, such as Germany, Japan, and South Korea. For example, 25% of the Germany-US difference in the adoption of robots is accounted for by these countries' different rate of workforce ageing. Speculatively, **population ageing could also impact the direction of technological innovation through an increased demand for automation in healthcare services**: Figure IRLESA 17provides examples by documenting the top 10 medical areas where artificial intelligence innovations are being developed.

Another interaction occurs between demographics and urbanization: Autor and Fournier (2019) show that US urban regions are ageing much more slowly than less densely populated areas. Indeed, since 1950, rural areas have aged 12 years on average whereas cities have only aged 2. This striking pattern is entirely accounted for by a dramatic change in migration patterns: younger workers are moving to the city, and no longer moving out.

[195] Acemoglu, D. and Restrepo, P. 2019. "Demographics and Automation." MIT working paper.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

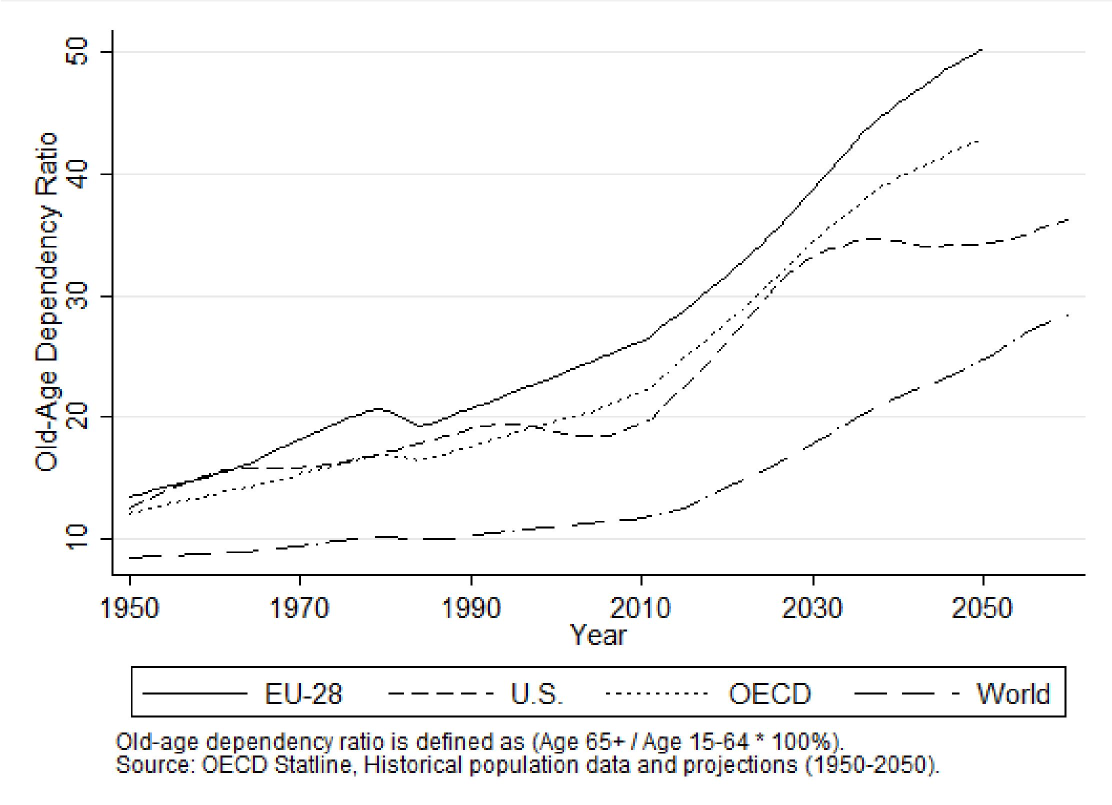


FIGURE IRLESA 16 HISTORICAL AND PROJECTED OLD-AGE DEPENDENCY RATIOS OVER TIME

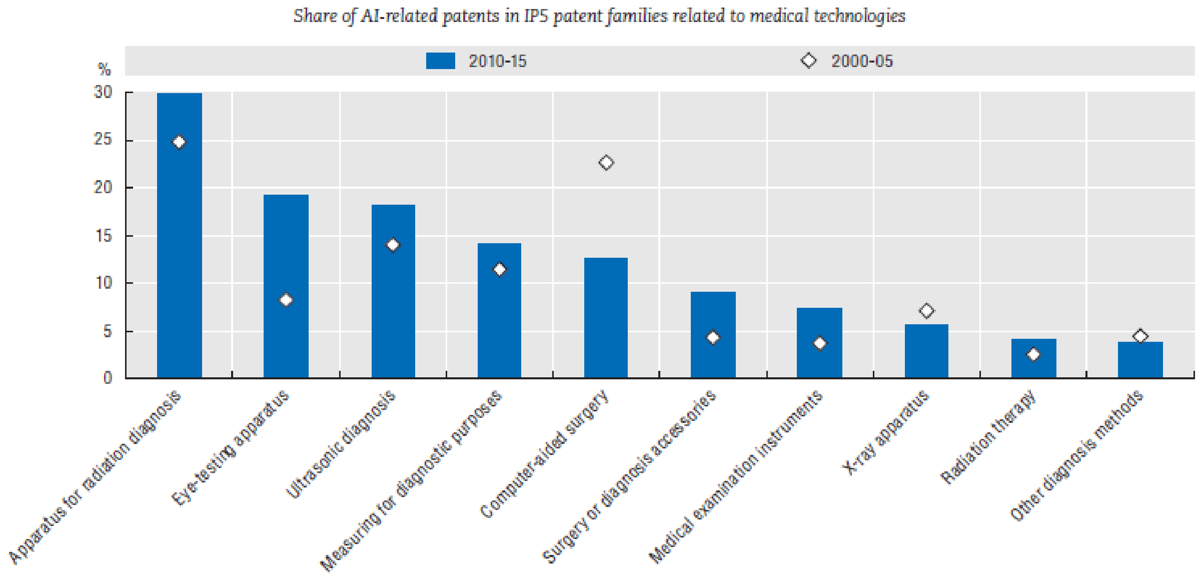


FIGURE IRLESA 17 TOP 10 ARTIFICIAL INTELLIGENCE-BASED MEDICAL TECHNOLOGIES

Source: OECD (2017a).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

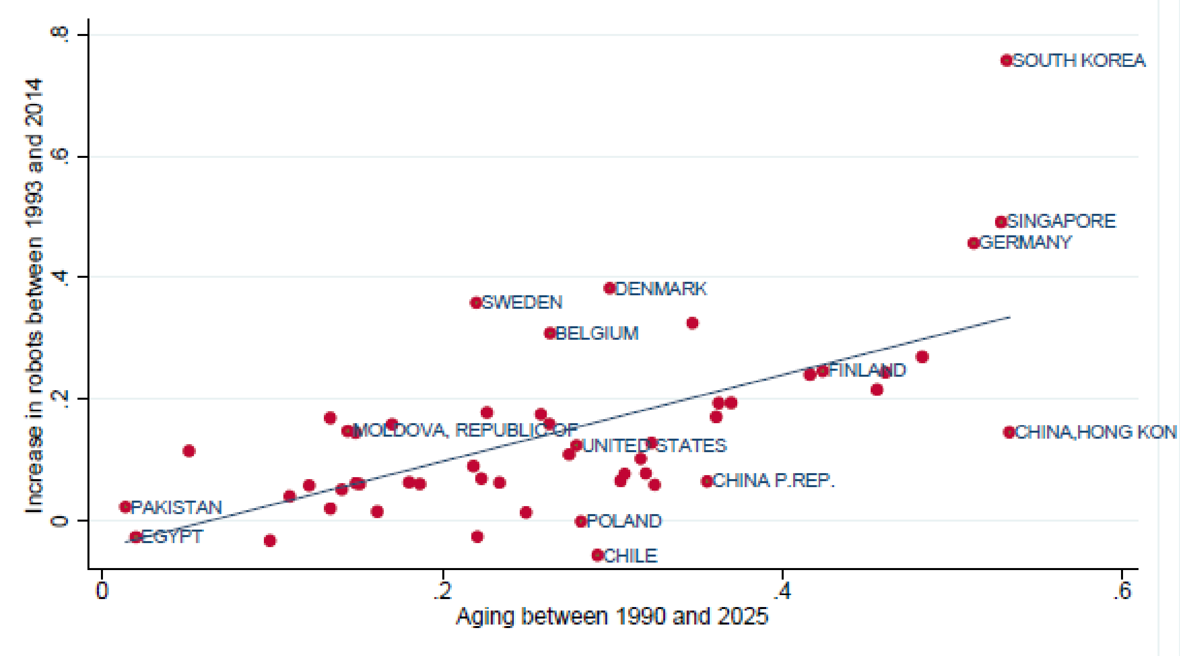


FIGURE IRLESA 18 WORKFORCE AGEING AND ROBOT ADOPTION

Source: Acemoglu and Restrepo (2019). The relationship plotted in the figure is based on a model controlling for other factors.

3. Chances and challenges for the future of work

The key trends outlined in the previous section are technological progress and automation; globalization and urbanization; a rising diversity of work forms; and demographic change. To a large extent, these trends have fuelled economic growth and brought unprecedented prosperity to our societies. As such, they do not pose a challenge to the future of our economies: rather, this future critically depends on our ability to continue innovating production processes and developing new goods and services, finding national and international markets for selling them, and reshaping and reallocating jobs to be their most productive.

However, these trends do pose challenges for the future of work. These societal challenges are not related to a shortage of jobs, as we explain in section 3.3.2, but rather a matter of distribution: in particular, rising inequality (section 3.3.1.); job reallocation and polarization (section 3.3.2.); and skill gaps and skill depreciation (section 3.3.3.).

a. Rising inequality

One of the key features of the digital age has been a tendency towards increasing labour market inequality in many OECD countries – albeit from different initial levels. **FIGURE IRLESA 19**shows how striking these patterns are for the US, a country which has one of the highest levels of inequality as well. For both men and women, real wage growth was broadly shared across different educational groups in the 1960s and early 70s. However, especially since 1980, wage growth for high-educated workers (those with a college degree) has raced ahead of that experienced by their lower-educated counterparts – as a

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

result, inequality rose. For women, the lowest-educated groups still experienced positive wage growth even as they fell behind more educated workers; but for non-college educated men, there have been striking real wage declines since 1980. While these changes are not equally strong or striking in all OECD countries, there has been a tendency to more inequality on average, as shown in **Figure IRLESA 20** Indeed, the top 90% earners have witnessed stronger wage growth than the bottom 40% or 10% earners, resulting in increasing wage inequality.

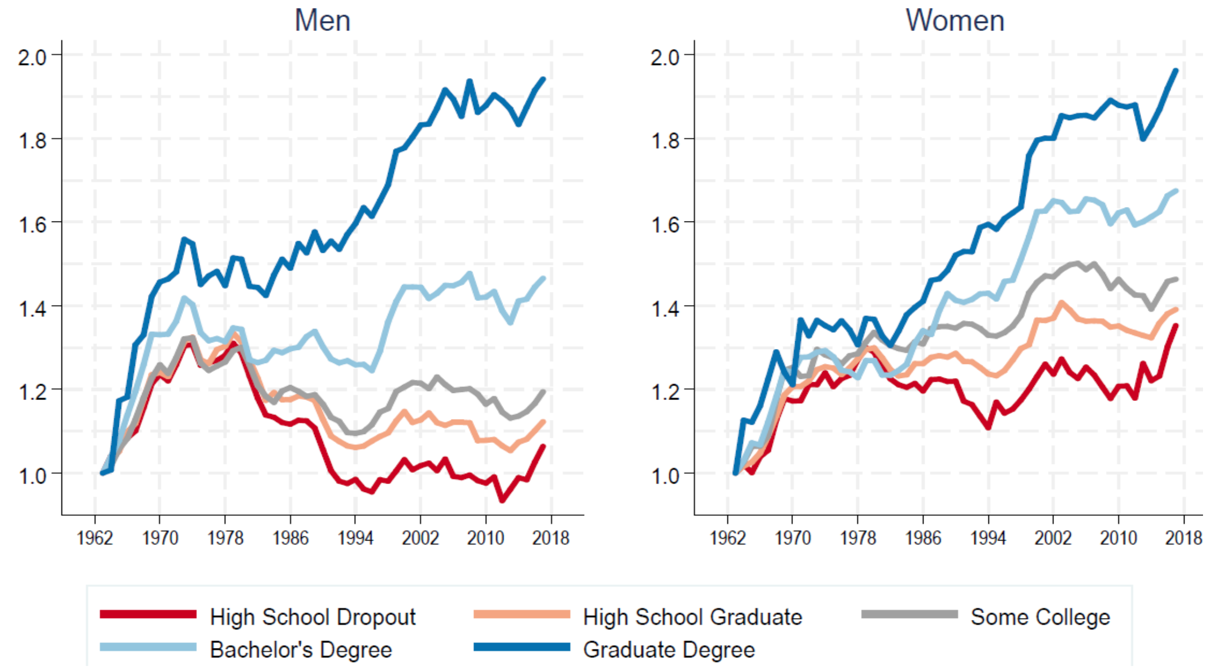


Figure IRLESA 19 Rising U.S. inequality, 1963-2017

Notes: Cumulative change in real weekly earnings for adults aged 18-64. Source: Autor (2019).

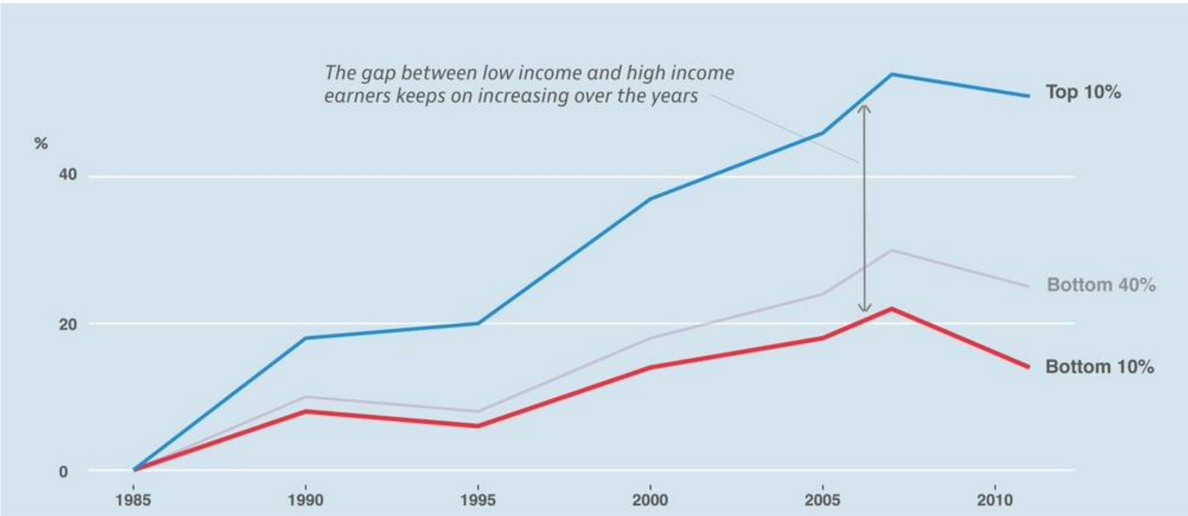


Figure IRLESA 20 Increasing inequality across OECD countries

Notes: Change in real household income relative to 1985, OECD average. Source: OECD (2015).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

The main driver behind this increasing dispersion in wage earnings has been technological progress (along with globalization): **new technologies have strongly increased the demand for high-skilled workers. This is because technologies are enabling skilled workers to be more productive and produce new goods and services.** The skill premium has risen whenever the increase in the demand for skill has outstripped its increase in supply – this is known as Jan Tinbergen's famous race between education and technology.

So far, the focus has been on inequality in labour earnings. However, there have also been changes in the distribution of earnings between capital and labour: in particular, labour's share of national income has decreased in many countries. **FIGURE IRLESA 21**shows this pattern for the world's four largest economies: the US, China, Germany, and Japan. This trend appears to be partially driven by advancing technologies as well [196]. It goes without saying that capital income is also owned by people, and therefore the declining labour share does not reflect a decrease in incomes – however, capital income is much more unevenly distributed than labour earnings, reinforcing the trend towards more income inequality.

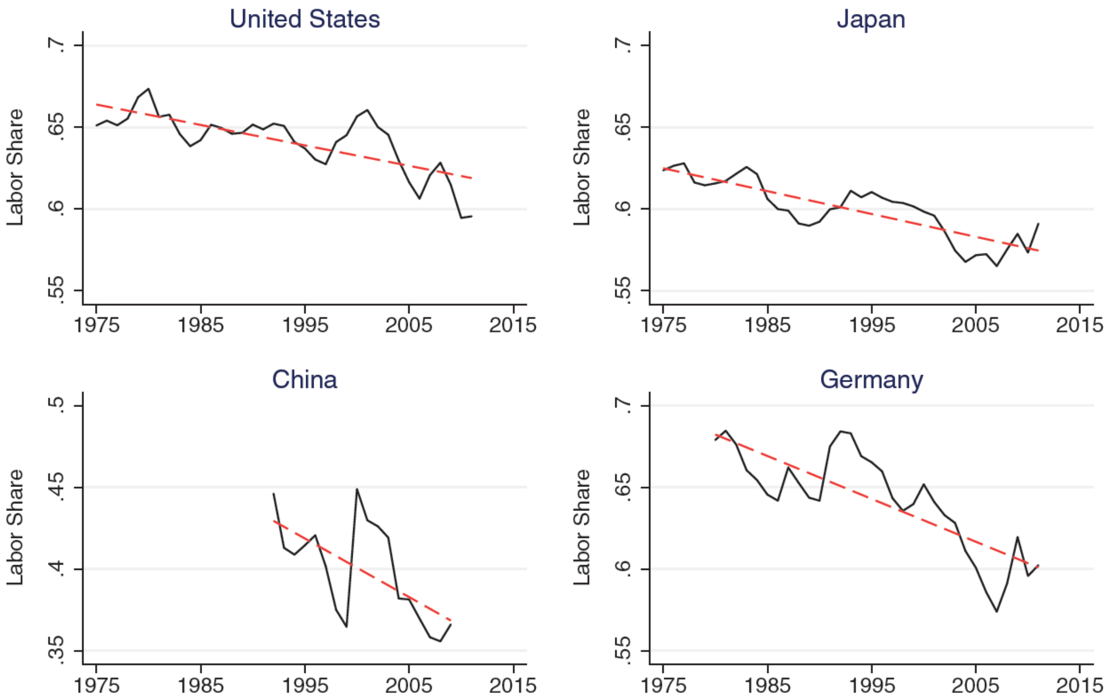


FIGURE IRLESA 21 DECLINING LABOUR SHARES IN THE WORLD'S FOUR BIGGEST ECONOMIES

Source: Karabarbounis and Neiman (2014).

[196] Karabarbounis, L., and Neiman, B. "The Global Decline of the labour Share," The Quarterly Journal of Economics, 129 (2014): 61; Autor, D.H., and Salomons, A. 2019. "New Frontiers: The Evolving Content and Geography of New Work in the 20th Century." Working Paper

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

b. Job reallocation and polarization

A commonly heard public concern for the future of work is that jobs are being eliminated on net: that is, that we are heading for a future of mass technological unemployment. While such fears are by no means new[197], they have found no empirical support despite widespread workplace automation.

To the contrary, the past decades of automation have led to net job creation[198]. **While direct replacement of labour by machines does reduce the demand for labour in industries where automation takes place, there are important countervailing mechanisms.** Firstly, automation reduces prices, leading to additional demand for the goods produced by automating industries. This increased product demand leads to more labour in industries linked to the automating industry through the supply chain, as either suppliers or customers. For example, when the car manufacturing industry becomes more productive because of automation and therefore sells more cars, steel industries (which supply inputs to car manufacturers) see an increased demand as well, and thereby rising employment. And lastly, technological advances increase consumers' income, leading to increases in output and employment in all industries: this can be termed the final demand effect.

The sizes of these effects are illustrated in **FIGURE IRLESA 22**: while automation has a negative direct effect on employment, there is a positive net effect on jobs because of large positive countervailing mechanisms. Overall, estimates indicate that advancing technologies *increase* employment by around 0.5% annually across developed countries.

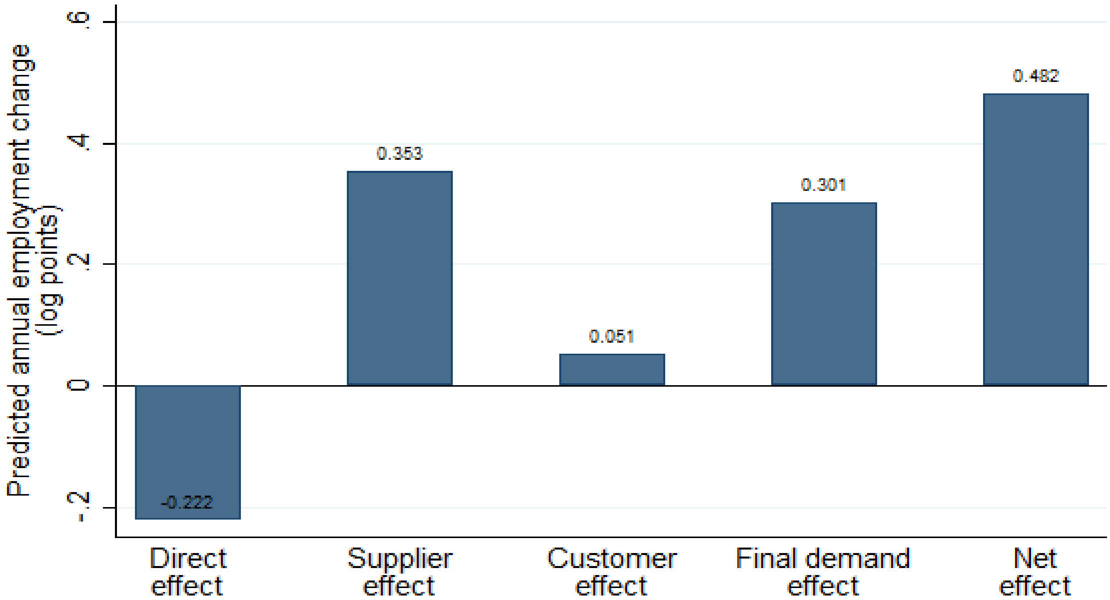


FIGURE IRLESA 22 EMPLOYMENT EFFECTS OF ADVANCING TECHNOLOGY

[197] Mokyr, J. et al., "The History of Technological Anxiety and the Future of Economic Growth: Is This Time Different?" Journal of Economic Perspectives, 29 (2015): 31
[198] Autor, D.H., and Salomons, A. 2019. "New Frontiers: The Evolving Content and Geography of New Work in the 20th Century." Working Paper.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Notes: Graph shows the predicted annual employment change from productivity growth on average across 17 developed economies. Source: Autor and Salomons (2018).

However, the absence of job destruction in aggregate does not preclude job displacement for individual workers who are facing direct automation or offshoring of their job tasks. Indeed, job reallocation has been an enduring feature of past waves of technological progress, also: a prime historical example of this is the transition from agriculture to manufacturing.

Such ongoing reallocation is clearly seen from changes in the job structure – that is, the changing employment shares of different job types. Over the last decades, many advanced countries have seen a process known as job polarization[199]. This means that there has been a decline in employment shares of jobs in the middle of the wage distribution, such as clerical and production work, while employment shares of high-skilled professional jobs have increased. Employment shares of low-skilled jobs (such as construction labourers, childcare workers, waiters, and a range of personal care and service workers) have also risen, albeit to a lesser extent.

Job polarization arises because many of the tasks performed by medium-skilled workers can be automated using digital technologies. High-skilled workers, on the other hand, have been made more productive by these technologies, but without being replaced: these technologies complement their tasks. At the same time, many tasks that do not require high levels of human skill (e.g. cleaning or hairdressing) are as yet very difficult to automate this is why low-skilled work has not declined in the same way as middle-skilled work.

This pattern of polarization is reinforced by international trade, as it is mostly middle-skilled production work that is exposed to offshoring and import competition: low-skilled services need to be delivered in-person, and developed countries have a strong comparative advantage in high-skilled professional work. Job polarization is illustrated for OECD countries in **FIGURE IRLESA 23**: in all countries, employment shares of middle-skilled jobs have decreased, and for some, markedly so.

---

[199] Autor, D.H. et al., "The Polarization of the U.S. labour Market," American Economic Review Papers and Proceedings , 96 (2006), 189; Goos, M. et al., "Job Polarization in Europe." American Economic Review Papers & Proceedings: 99 (2009), 58

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

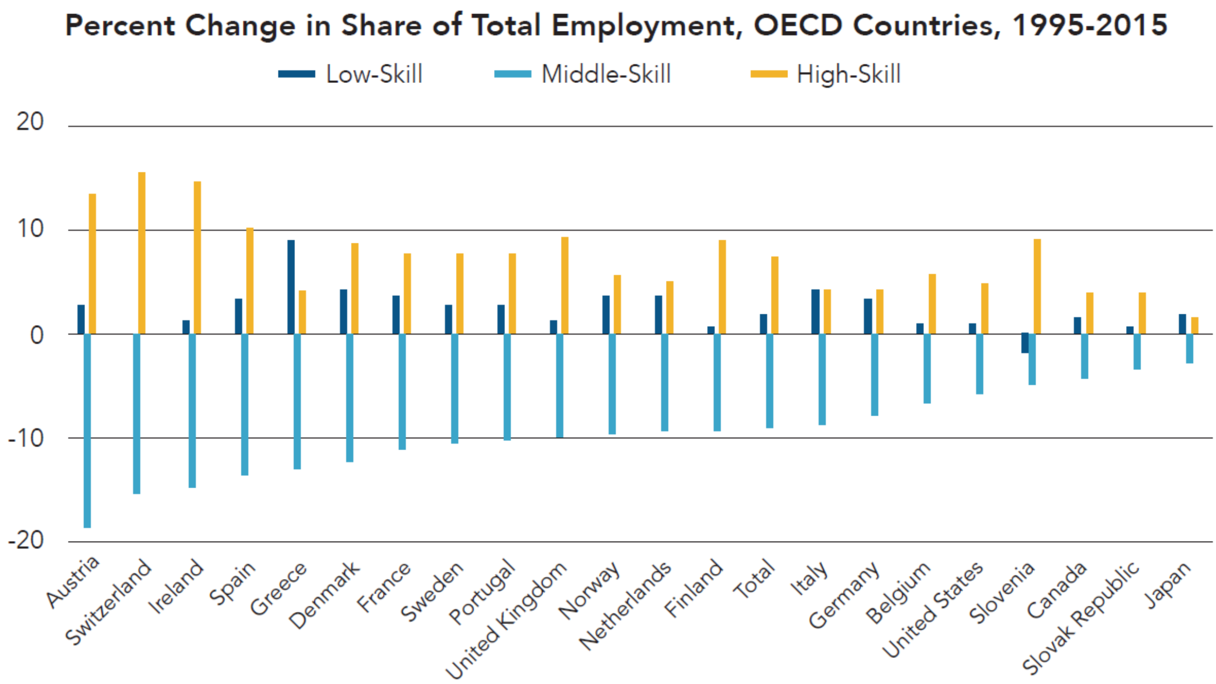


FIGURE IRLESA 23 JOB POLARIZATION

Source: OECD (2017b).

Polarization also has an important regional component. **FIGURE IRLESA 24** shows US employment shares for three broad occupation groups (low-paid service work; middling production and clerical work; and high-skilled professional work) across regions with different population densities over 1970-2015. This reveals that in the 1970s, denser (i.e. more urban) regions had far more middle-skill work than did suburban and rural regions. But this feature attenuated and subsequently reversed sign over the next four decades. That is, while middle-skill work was differentially present in urban areas in the 1970s, it was differentially absent from those same places 45 years later, in addition to less prevalent everywhere in absolute terms.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

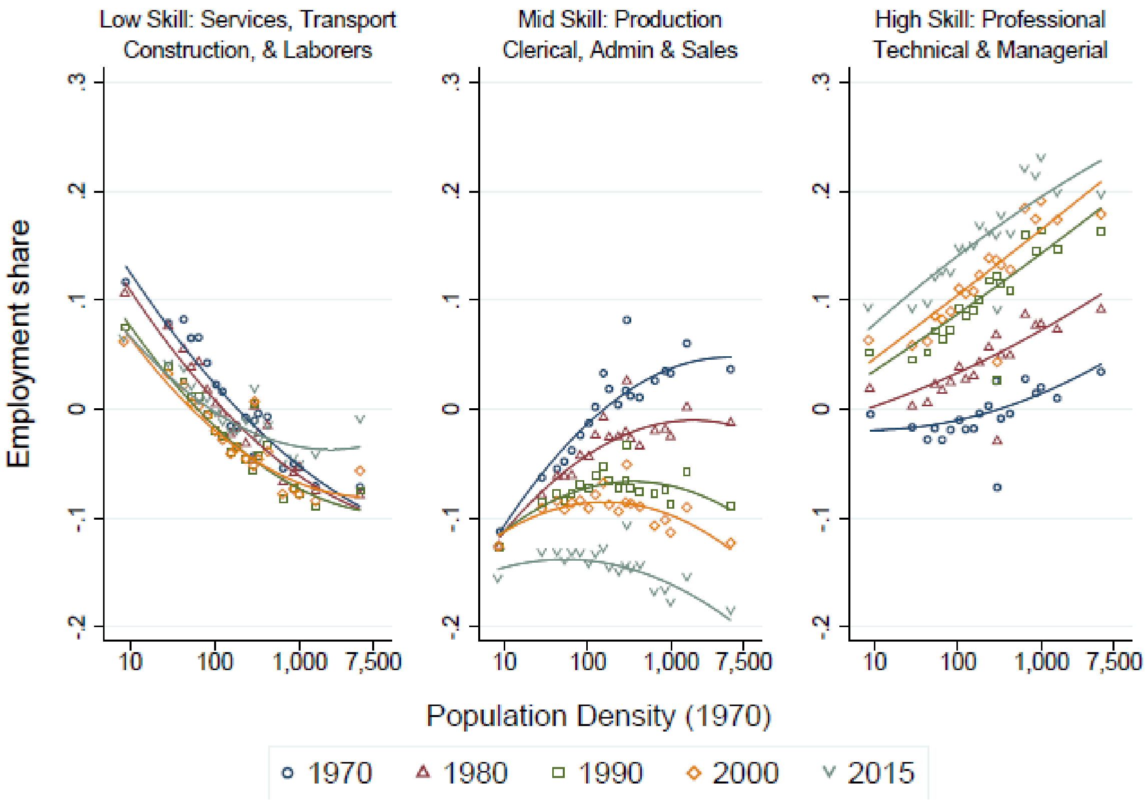


FIGURE IRLESA 24 MIDDLING JOBS ARE DISAPPEARING FASTER IN URBAN AREAS

Notes: The figure plots occupational employment shares among working-age adults by commuting zone population density, 1970 - 2015: level relative to 1970 mean. Each plotted point represents approximately 5 percent of the working-age population in the relevant year. Source: Autor (2019).

This trend of job polarization is to some extent reinforced by new job creation[200]. This is because advancing technologies appear to lead to three broad classes of new jobs: frontier work, wealth work, and last-mile work.

Frontier work involves directly producing, installing, maintaining, and deploying novel technologies. Recent examples are jobs in robot integration, search engine optimization, and radiological medicine – these are the jobs that have increasingly moved to cities, as documented in section 3.2.2. Frontier jobs are highly paid, and predominantly performed by college-educated men.

Last-mile work, on the other hand, involves carrying out nearly automated tasks that retain only a residual set of human components. Last-mile tasks typically do not require high levels of technology-specific expertise. Historical examples include call-center operators, order fulfillment workers, machine offbearers, and data entry clerks, whereas current examples include content taggers and facial recognition verifiers. These jobs are

200 Autor, D.H., and Salomons, A. 2019. "New Frontiers: The Evolving Content and Geography of New Work in the 20th Century." Working Paper.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

generally low-paid, can be performed with minimal training, and are geographically spread-out as they do not require in-person interaction.

The final class of new jobs, wealth work, appears to arise as novel consumer luxuries driven by increased incomes. As such, wealth work occupations perform in-person services for affluent consumers: nail technicians, dog groomers, and many forms of personal training and counseling. Most wealth work is neither technologically novel nor broadly demanding of technical skills. It is also not highly paid. Women make up a disproportionate share of workers in wealth work occupations. These jobs have always been most prevalent in densely populated areas.

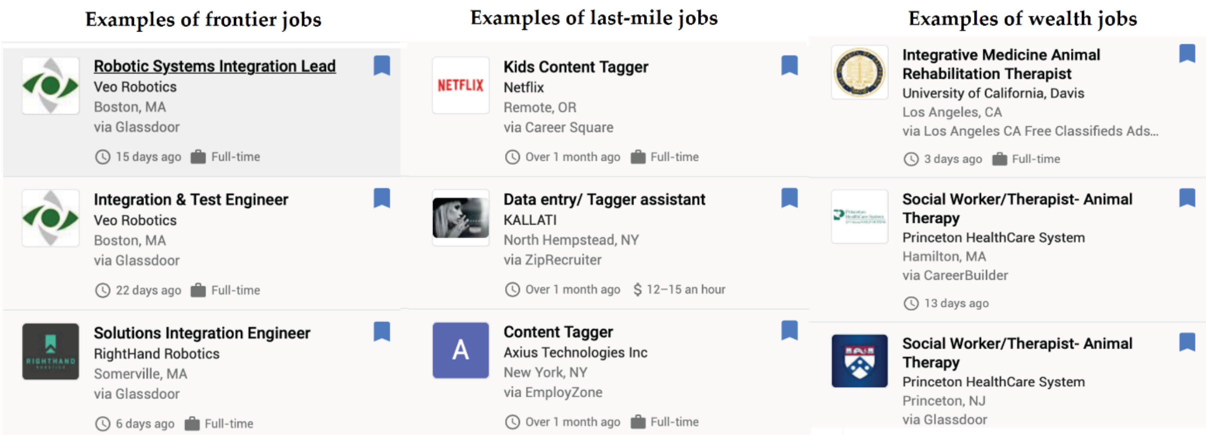


FIGURE IRLESA 25 . EXAMPLES OF NEW JOB TYPES

Source: Autor and Salomons (2019)

Examples of recent vacancies for these three types are shown in Figure IRLESA 25. While such newly emerging jobs make up only a small share of total employment, they do appear to reinforce the overall pattern of polarization. After all, in net, a disproportionate share of such `new work' is generated within dense urban labour markets. Moreover, the bifurcated structure of new work does not suggest that a technology-driven 'reinstatement' of middle-skill, non-college jobs is underway.

Recent work has quantified the adjustment costs from job reallocation for individual workers, finding a rise in the chance of firm separation for workers whose firms make major automation investments[201]. As shown in **FIGURE IRLESA 26**, workers affected by automation at their firm have up to 7 percentage point higher firm exit rates after 5 years. For workers with longer firm tenure ("incumbents"), these effects are sizable relative to their firm separation chance in the absence of automation. Indeed, because of automation, these workers are 24% more likely to leave their firm. Research shows that workers displaced by automation do find re-employment, but this takes some time: in

[201] Bessen, J. et al., "Automatic Reaction: What Happens to Workers at Firms that Automate?" Boston University School of Law, Law and Economics Research Paper, 2019, https://ssrn.com/abstract=3328877

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

total over 5 years, they lose around 10% of one annual salary from the resulting unemployment spells[335].

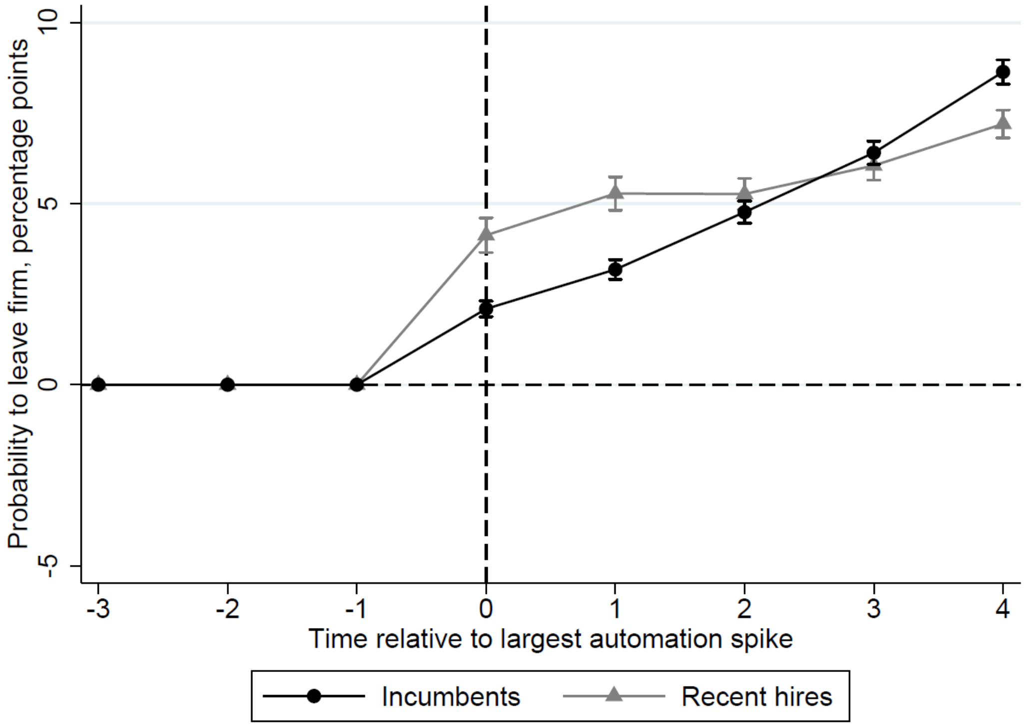


Figure IRLESA 26 Job displacement increases after automation

Notes: The graph shows the probability of workers whose firm automates (in t=0) separating from their job, compared to control group workers. Incumbents are workers who have been employed at the firm for at least 3 years prior to the automation event, recent hires have been hired less than 3 years prior to the automation event. Source: Bessen, Goos, Salomons, and Vandenberge (2019).

Similar displacement effects have been documented for workers adversely affected by increased international trade. As an example, **Figure IRLESA 27** shows how import competition from China has driven down US manufacturing employment.

All in all, this points to job reallocation and displacement as a likely feature of the future of work: while this is not new, it does impose costs on individual workers who see their jobs automated or offshored. Further, the types of jobs and workers affected may change over time along with the nature of international competition and automation capabilities.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

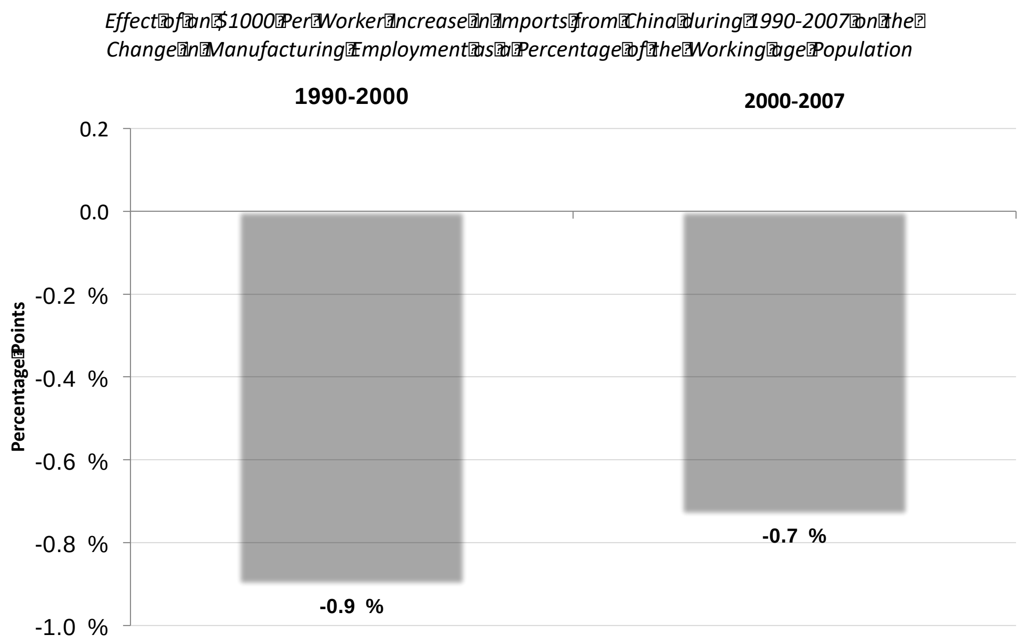


FIGURE IRLESA 27 TRADE-DRIVEN JOB DISPLACEMENT

Source: Autor, Dorn, and Hanson (2016).

c. Skill gaps and skill depreciation

Key trends such as automation and globalization lead to a change in the skills demanded in the labour market. This occurs through two broad channels: changes in the occupational structure, and changes in the task content of existing jobs. This first (between-occupational) effect occurs as jobs where routine tasks are performed decrease in importance in the economy: job polarization is one driver of such change. The creation of entirely new jobs also contributes to changes in the occupational structure. The second (within-occupational) effect arises when existing jobs change their task content: for example, secretaries currently perform a very different set of tasks than before the advent of the computer. These within-occupational changes in task content are estimated to be an important source of labour market adaptation to change[202]. Together, these two sources of changing skill needs imply **workers have to be able to acquire new skills for switching to other job types as well as adapt to changing task competence requirements in their existing work.**

When skill needs in the economy are not met, skill gaps occur: these gaps are known to be especially large in the fields of Science, Technology, Engineering and Math (STEM). **FIGURE IRLESA 28** illustrates that many firms across the OECD report having ICT specialist vacancies that are difficult to fill, and that this difficulty has increased since the end of the crisis. Similarly, analysis of vacancy data shows that **the median duration of advertising for a STEM vacancy is more than twice as long as for a non-STEM**

[202] Arntz, M. et al., "Revisiting the risk of automation," Economics Letters, 159 (2017): 157

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

vacancy (Rothwell 2014). These indicators signal that there is a shortage of supply of STEM skills in the labour market, relative to demand.

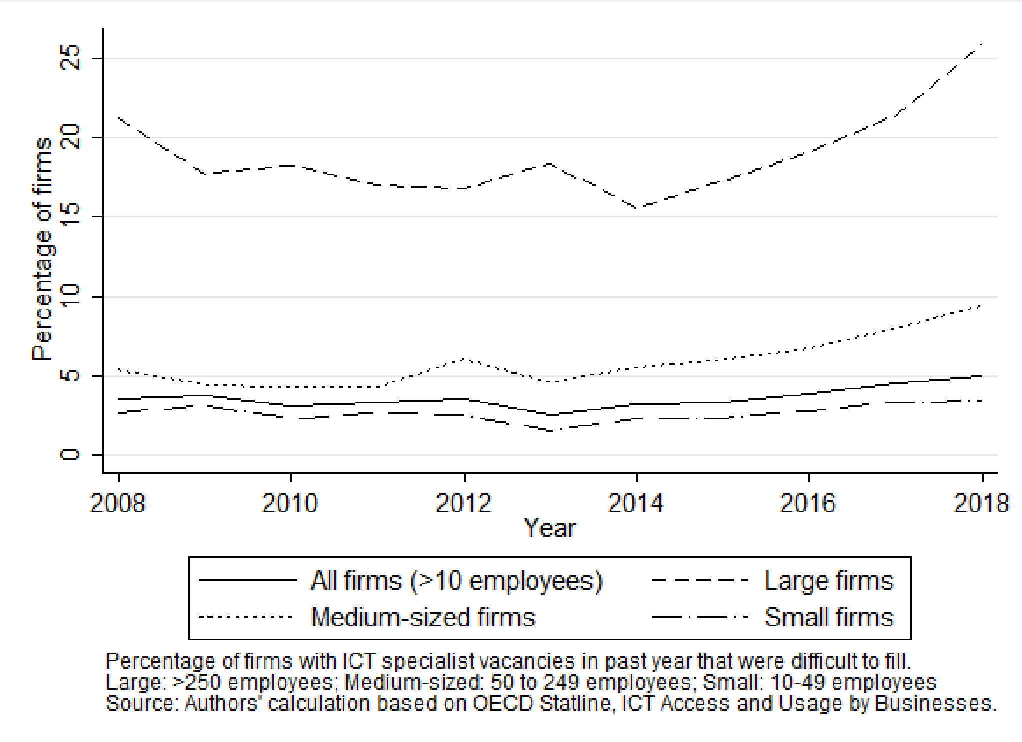


FIGURE IRLESA 28 ICT SPECIALIST VACANCIES ARE DIFFICULT TO FILL

Along with the rising demand for STEM and other technical skills there has also been a lesser well-known rise in the demand for social skills, and even more so when *combined* with mathematical skills[203]. This is illustrated in **FIGURE IRLESA 29** and **FIGURE IRLESA 30**, which respectively show the rising employment and wages of jobs involving a combination of social and maths

Women in Science, Technology, Engineering, & Math (STEM)

It is well-documented that women are underrepresented in STEM fields in terms of education and jobs. Only 30% of STEM graduates are women, and female representation in careers in these fields is even lower, as well as decreasing in job levels (OECD 2017a, Randstad 2017). Further, across the OECD, fewer than 10% of patents are filed by female innovators. Recent work shows that very little of these sizable differences can be attributed to differences in STEM ability or achievement between boys and girls at the time of course specialization choice (Delaney and Devereux 2019). This suggests these differences could be influenced by cultural factors including preferences, socialization, role model and peer effects, as well as expectations of future workplace discrimination. A fair representation of women as well as

[203] Deming, D.J. 2017. The Growing Importance of Social Skills in the labour Market. Quarterly Journal of Economics, 132 (2017) :1593

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

skills. **Social skills are argued to matter because they help workers cooperate and work in teams, which is increasingly important in today's workplaces.** These overall patterns are consistent with the OECD Skills for Jobs Database, which documents skill shortages in not only mathematical skills, but also judgment and decision-making as well as social perceptiveness.

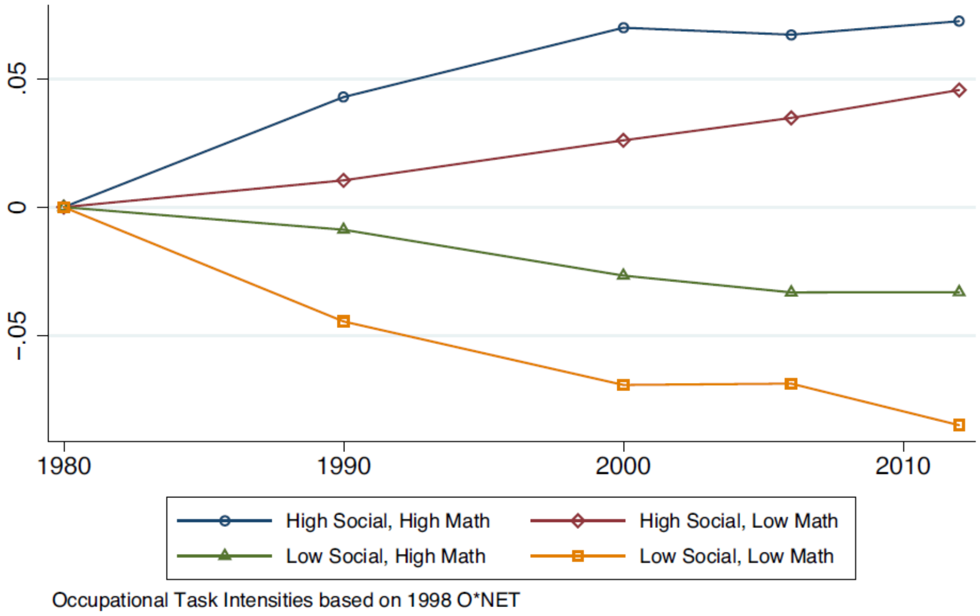


FIGURE IRLESA 29 RISING EMPLOYMENT FOR JOBS WITH SOCIAL SKILLS

Notes: Cumulative changes in employment share by occupation task intensity, 1980–2012. Source: Deming (2017).

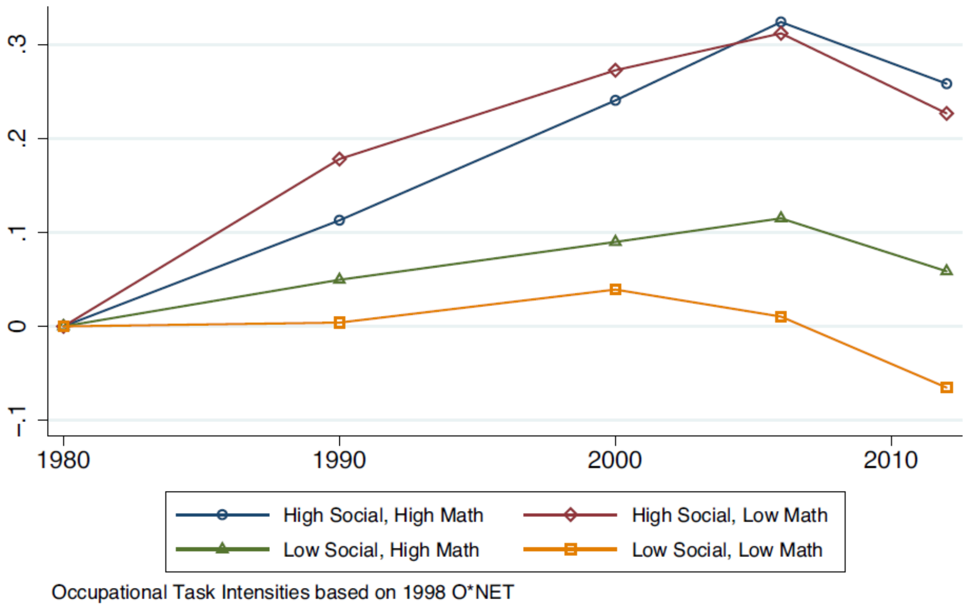


FIGURE IRLESA 30 RISING WAGES FOR JOBS WITH SOCIAL SKILLS

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Notes: Cumulative changes in real hourly wages by occupation task intensity, 1980–2012. Source: Deming (2017).

Women in high-skilled jobs

Since the 1980s, high-educated women have fared well relative to high-educated men. One key channel driving this seems to be the greater increase in the demand for female-oriented skills in cognitive/high-wage occupations relative to other occupations. Research shows that this relative increase in the demand for female workers is due to an increasing importance of social skills within such high-skilled occupations (Cortés et al 2018).

For female workers, these changing skill demands have led to two contrasting patterns: while women are still strongly underrepresented in STEM occupations, high-skilled women overall have fared at least as well or even better than high-skilled men because there has been such a strong increase in the demand for social skills within high-skilled jobs (see insets).

Lastly, skill depreciation is also a challenge whenever technology advances: this is especially relevant given the increasing average age of our workforces. An analysis of vacancy data shows that technology-intense STEM jobs have changed especially quickly over the last decade, leading to flatter age-earnings profiles as the skills of older cohorts become obsolete[204]. Further, **older workers experience larger earnings losses following automation in their firm, since they have more trouble transitioning to new jobs**[205]. In some countries such as the US, **entry to a different occupation may be further restricted because of occupational licensing**. Strikingly, 5% of jobs required an occupational license in the 1950 versus nearly 25% today[206].

All in all, **skill demands change both within and between jobs as work is partially automated** or offshored, or as new types of production is enabled as a result of the key trends outlined in section 1. These trends therefore present ongoing skill-related challenges for the future of work.

4. Expected job loss and job creation from digitalization

McKinsey Global Institute (Bughin et al. 2018) quantifies net job changes resulting from automation and other macroeconomic trends. They identify four drivers of job loss and gain from both AI and non-AI factors:

1. Job losses due to automation
2. Job loss due to non-AI productivity gains
3. Direct job gains from automation

---

[204] Deming, D.J. 2017. The Growing Importance of Social Skills in the labour Market. Quarterly Journal of Economics, 132 (2017) :1593

[205] Bessen, J. et al., "Automatic Reaction: What Happens to Workers at Firms that Automate?" Boston University School of Law, Law and Economics Research Paper, (2019), https://ssrn.com/abstract=3328877

[206] Kleiner, M. M. and Krueger, A. "Analyzing the Extent and Influence of Occupational Licensing on the labour Market." Journal of labour Economics, 31 (2013): S173

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

4. Job gains due to macroeconomic drivers, including indirect effects from automation.

In order to be able to identify the number of jobs that will be eliminated and the new jobs that will be created, due to the effect of automation, we have had to resort to different databases. The hours or percentages of each of the skills have been obtained from McKinsey Global Institute (MGI)[207] (pages 8 and 14). Data on hours worked per worker in each of the countries have been downloaded from the OECD (2019). Average annual hours worked is defined as the total number of hours actually worked per year divided by the average number of people in employment per year. Actual hours worked include regular work hours of full-time, part-time and part-year workers, paid and unpaid overtime, hours worked in additional jobs, and exclude time not worked because of public holidays, annual paid leave, own illness, injury and temporary disability, maternity leave, parental leave, schooling or training, slack work for technical or economic reasons, strike or labour dispute, bad weather, compensation leave and other reasons (OECD, 2019).

From the International Labour Organization (ILO) (2019) we use data referring to the employed population, both in the year 2016 and the latest available estimate (year 2022). In **Table IRLESA 1**, we can see the employed population in each of the countries analysed. If we multiply this population by the average number of hours worked per worker, we obtain the total hours worked in each country. We can observe that the United States, with a smaller occupied population than Western Europe has the same total hours worked. This productivity differential is motivated by the fact that a US worker works more hours than a European worker. This fact has positive effects in the face of less job loss.

In **Table IRLESA 2** (United States) and **Table IRLESA 3** (Western Europe), we have calculated, from MGI estimates, the hours that change in the period 2016-2030 due to the effect of automation according to each of the demanded skills. In order to extrapolate the number of jobs affected, we have divided the total number of hours by the average number of hours worked per worker. The result is the number of jobs lost and jobs gained.

We can see in **Table IRLESA 2** that the United States would lose 9,730,337 jobs but would need 21,179,775 new jobs (Higher cognitive skills, Social and emotional skills, and Technological skills). The net balance would be an increase of 11,449,438 jobs.

Table IRLESA 3 shows that Western Europe (Austria, Belgium, Denmark, Finland, France, Germany, Greece, Italy, Netherlands, Norway, Spain, Sweden, Switzerland, and the United Kingdom) would lose 18,792,969 jobs but would need 27,605,046 new jobs. The net balance would be an increase of 8,812,077 jobs.

From the MGI data analysed we can make a more exhaustive analysis of five European countries with the largest working population: Germany, United Kingdom, France, Italy and Spain. To do this, we had to calculate the number of workers according to the percentage of hours they dedicate to each of the skills. MGI provides the data as a

[207] Bughin, J. et al., "Skill Shift Automation and The Future Of The Workforce", McKinsey Global Institute, May 2018. https://www.mckinsey.com/featured-insights/future-of-work/skill-shift-automation-and-the-future-of-the-workforce

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

percentage of total hours worked. As we know the employed population and the average number of hours worked per worker, we can calculate the number of average workers affected.

TABLE IRLESA 1 : LABOUR FORCE, HOURS WORKED PER WORKER, TOTAL HOURS WORKED (2016 AND 2022) BY COUNTRIES

Country	Labour Force 2016 **(thousands) ILO** **(A)**	Hours worked **Hours/worker (OECD) (B)**	Labour Force 2022 **(thousands) ILO** **(C)**	Hours worked Western Europe and USA 2016 **(AxB)**	Hours worked Western Europe and USA 2022 **(CxB)**
Austria	4.250	1.487,00	4.269	6.319.750.000	6.348.003.000
Belgium	4.624	1.546,00	4.752	7.148.704.000	7.346.592.000
Denmark	2.830	1.408,00	2.871	3.984.640.000	4.042.368.000
Finland	2.454	1.531,00	2.514	3.757.074.000	3.848.934.000
France	27.158	1.514,00	27.817	41.117.212.000	42.114.938.000
Germany	41.291	1.356,00	41.784	55.990.596.000	56.659.104.000
Greece	3.894	1.906,00	4.133	7.421.964.000	7.877.498.000
Italy	22.148	1.722,60	22.296	38.152.144.800	38.407.089.600
Netherlands	8.485	1.433,00	8.742	12.159.005.000	12.527.286.000
Norway	2.679	1.419,10	2.766	3.801.768.900	3.925.230.600
Spain	18.443	1.686,50	19.223	31.104.119.500	32.419.589.500
Switzerland	4.665	1.570,00	4.786	7.324.050.000	7.514.020.000
Sweden	4.839	1.453,00	5.027	7.031.067.000	7.304.231.000
UK	32.166	1.514,00	33.348	48.699.324.000	50.488.872.000
TOTAL **Western Europe**	179.926	**1.522,91**	184.328	**274.011.419.200**	**280.715.276.715**
USA	154.274	**1.780,00**	159.886	**274.607.720.000**	**284.597.080.000**

Source: Own elaboration from the data provided by International Labour Organization (2019); OECD (2019)

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

TABLE IRLESA 2 JOBS LOST AND GAINED UNITED STATES (ESTIMATION2016-2030)

United States, **all sectors (1.780 hours/worker)**						
Skills	Hours worked 2016 (billions)	Change in hours worked by 2030 (%)	Hours worked 2030 (billions)	Change in hours worked by 2030	Change in Worker by 2030	
Physical and manual skills	90	-11,00	80,1	-9,90	-5.561.798	
Basic cognitive skills	53	-14,00	45,58	-7,42	-4.168.539	**-9.730.337**
Higher cognitive skills	62	9,00	67,58	5,58	3.134.831	
Social and emotional skills	52	26,00	65,52	13,52	7.595.506	
Technological skills	31	60,00	49,6	18,60	10.449.438	**21.179.775**
TOTAL	288		308,38	20,38	**11.449.438**	

Source: Own elaboration from the data provided by MCI in Bughin et al (2018); International Labour Organization (2019); OECD (2019)

TABLE IRLESA 3 JOBS LOST AND GAINED WESTERN EUROPE (ESTIMATION 2016-2030)

Western Europe, **all sectors (1.522,91 hours/worker)**						
Skills	Hours worked 2016 (billions)	Change in hours worked by 2030 (%)	Hours worked 2030 (billions)	Change in hours worked 2030	Change in Worker by 2030	
Physical and manual skills	113	-16,00	94,92	-18,08	-11.872.008	
Basic cognitive skills	62	-17,00	51,46	-10,54	-6.920.961	**-18.792.969**
Higher cognitive skills	78	7,00	83,46	5,46	3.585.241	
Social and emotional skills	67	22,00	81,74	14,74	9.678.839	
Technological skills	42	52,00	63,84	21,84	14.340.966	**27.605.046**
TOTAL	362		375,42	13,42	**8.812.077**	

Source: Own elaboration from the data provided by MGI in Bughin et al (2018); International Labour Organization (2019); OECD (2019)

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Table 1: Workers by Skills per Country, 2016

TABLE IRLESA 4 WORKERS BY SKILLS PER COUNTRY, 2016

Country	2016 Hours (%) by Skills					2016 Workers by Skills					
	A	B	C	D	E	A	B	C	D	E	TOTAL
Spain	38	17	19	17	9	7008340	3135310	3504170	3135310	1659870	18443000
Italy	38	17	19	15	10	8416240	3765160	4208120	3322200	2214800	21926520
Germany	30	17	23	16	14	12387300	7019470	9496930	6606560	5780740	41291000
France	28	17	22	21	11	7604240	4616860	5974760	5703180	2987380	26886420
UK	25	17	22	23	12	8041500	5468220	7076520	7398180	3859920	31844340

Source: Own elaboration from the data provided by Bughin et al (2018); International Labour Organization (2019); OECD (2019). Skills: (A) Physical and manual skills; (B) Basic cognitive skills; (C) Higher cognitive skills; (D) Social and emotional skills; (E) Technological skills

In

TABLE IRLESA 4 we can see the total number of workers in the five countries in each of the skills in the year 2016. The hour percentage adds up to 100. Table IRLESA 5**¡Error! No se encuentra el origen de la referencia.** shows the forecast of total workers, in each of the different skills, in each country. The percentages of hours do not add up to 100. This is due to the fact that the variations (increase or decrease in hours) have occurred in each of the skills and not in the total.

We can see in Table IRLESA 6**¡Error! No se encuentra el origen de la referencia.** the jobs lost and won, in each of the skills, for each of the countries. All the countries analysed need more new jobs than the ones they are going to lose. This allows us to conclude that automation (robots and Artificial Intelligence) will not have a negative impact on jobs. Another issue is that countries may have workers with the skills required for the new jobs.

Those countries that are able to train their workers in the new skills and retain talented workers can shorten the transition period. This will enable them to become more competitive and take advantage of the automation benefits.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Table IRLESA 5 Workers by Skills per Country, estimate 2030

Country	2030 Estimates Hours (%) by Skills					2030 Estimated Workers by Skills					
											TOTAL
Spain	30,4	13,94	19,76	20,4	14,94	5843792	2679686,2	3798464,8	3921492	2871916,2	19115351,2
Italy	32,3	12,75	20,71	19,05	16,1	7201608	2842740	4617501,6	4247388	3589656	22498893,6
Germany	23,4	13,26	24,15	24,2	19,74	9777456	5540558,4	10090836	10111728	8248161,6	43768740
France	23,8	13,14	23,32	25,2	16,83	6620446	3655153,8	6486924,4	7009884	4681601,1	28454009,3
UK	22	14,79	23,98	27,83	15	7336560	4932169,2	7996850,4	9280748,4	5002200	72222749,3

Source: Own elaboration from the data provided by Bughin et al (2018); International Labour Organization (2019); OECD (2019). Skills: (A) Physical and manual skills; (B) Basic cognitive skills; (C) Higher cognitive skills; (D) Social and emotional skills; (E) Technological skills

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Table IRLESA 6 Jobs lost and gained by Skills per Country (Estimation 2016-2030)

Country	A	B	C	D	E	TOTAL	Lost	Gained
Spain	-1.164.548	-455.624	294.295	786.182	1.212.046	**672.351**	-1.620.172	**2.292.523**
Italy	-1.214.632	-922.420	409.382	925.188	1.374.856	**572.374**	-2.137.052	**2.709.426**
Germany	-2.609.844	-1.478.912	593.906	3.505.168	2.467.422	**2.477.740**	-4.088.756	**6.566.496**
France	-983.794	-961.706	512.164	1.306.704	1.694.221	**1.567.589**	-1.945.500	**3.513.090**
UK	-704.940	-536.051	920.330	1.882.568	1.142.280	**2.704.188**	-1.240.991	**3.945.179**

Source: Own elaboration from the data provided by Bughin et al (2018); International Labour Organization (2019); OECD (2019). Skills: (A) Physical and manual skills; (B) Basic cognitive skills; (C) Higher cognitive skills; (D) Social and emotional skills; (E) Technological skills

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

5. Themes for the impact of digitalization on the future of work

Here we outline three themes for thinking about the future of work. Together, these themes represent a yardstick along which developments in the world of work can be measured: do they safeguard decent work; do they promote inclusion of various groups in the labour market; and do they foster skill investment and acquisition?

a. Decent work

Several of the challenges outlined in the previous section touch on a key concern about the future of labour markets: the availability of decent work. The International Labour Organisation defines this as "opportunities for work that are productive and deliver a fair income, security in the workplace and social protection for families, better prospects for personal development and social integration, freedom for people to express their concerns, organize and participate in the decisions that affect their lives and equality of opportunity and treatment for all women and men". Broadly considered, decent work therefore comprises job availability, job pay, as well as non-wage job characteristics[208].

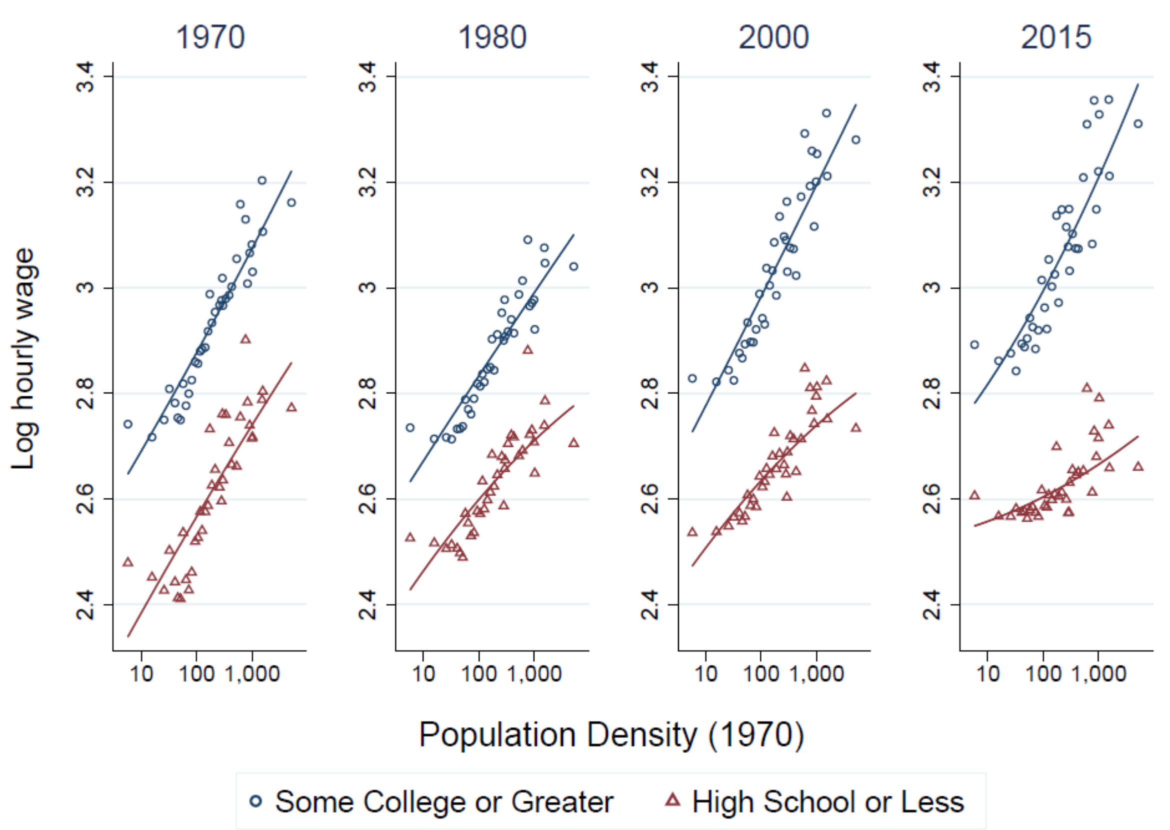


FIGURE IRLESA 31 DECLINING URBAN WAGE PREMIUM AMONG NON-COLLEGE WORKERS

[208] The equality aspects that the ILO partially classes under decent work are discussed separately, as inclusion, in section 3.5.2.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Notes: Figure plots real mean log hourly earnings among college and non-college workers in 1970, 1980, 2000, and 2015. Each plotted point represents approximately 2.5 percent of the working age population in the relevant year. Source: Autor (2019)

For a long time, advancing technologies were thought to only increase the prevalence of high-skilled, highly paid, work. However, the trend of job polarization, as well as the emergence of new wealth and last-mile work, shows that **there is likely to be a substantial fraction jobs with relatively low human skill requirements which can nevertheless not (yet) be automated. Safeguarding the quality of these jobs in terms of wages and non-wage characteristics such as autonomy is a key societal challenge**.

This is all the more important given how the trend of urbanization has impacted workers of different skill types, as urban areas have become more abundant in high-skilled jobs, while the availability of middle-skill work has declined. Although this skill-upgrading of the occupational structure is good news from a macro-economic perspective, workers without college degrees are shown to have much more limited access to these higher-paying jobs, and they are instead increasingly observed in low-skilled rather than medium-skilled occupations[209]. This shift is more pronounced in urban areas, highlighting the importance of regional disparities. Figure IRLESA 31 depicts how the urban wage premium has declined for workers without a college degree. That is, while urbanization has benefited highly educated workers as their wages are higher in more densely populated regions, lower-educated workers have witnessed their urban wage premium erode over time. This implies opportunities for decent work are reduced for this group.

Non-wage job characteristics, such as autonomy, trust, and work stress, have been shown to matter greatly for work satisfaction, which in turn matters for job retention[210]. These non-monetary job aspects may also be affected by the trends outlined above. While research about these aspects is still in its infancy, some patterns emerge. For one, automation has eliminated the need for human intervention in many dangerous and tedious tasks, relieving workers of these workplace disamenities (OECD 2019). Indeed, **there is recent survey evidence suggesting the implementation of ICT and AI technologies increases job satisfaction in many jobs[211]. However, workers interacting with these technologies also report higher levels of stress.** This is arguably the case because as routine work is automated, there is more room for creative and problem-solving tasks – and these complex tasks typically bring higher mental loads. At the same time, many last-mile jobs which emerge as a result of partial automation processes are not stressful because of an abundance of challenging tasks

Public concern about the gig economy

Some food delivery companies operating in multiple countries have been criticized in recent years for classifying its couriers as self-employed, circumventing worker rights such as minimum hourly wages and paid holiday leave (e.g. see Reuters 2018). A 2018 UK parliamentary inquiry found that this employment model fosters a polarized labour market that works well for some and poorly for others, including large earnings gaps between the platform's riders for identical work. Similar concerns have been brought up with respect to other gig economy companies, illustrating a growing societal concern to ensure those employed in new work forms have decent work.

---

209 Autor, D.H., and Salomons, A. "New Frontie
20th Century." (2019), Working Paper.
210 Clark, A.E. "What Really Matters in a Job? H
(2001): 223; Helliwell, J. and Huang, H. "Well-I
Studies, 12 (2011): 747.
211 Yamamoto, I. 2019. "The Impact of AI and I
https://voxeu.org/article/impact-ai-and-informa

This project has received funding from the
esearch and innovation program under gra

– rather, they are anecdotally known to lack autonomy and may be emotionally stressful in some cases. For instance, warehouse order fillers have documented how their work pace is determined and monitored by technology; whereas human content taggers are being exposed to emotionally taxing material. **More research is needed to determine how the introduction of new workplace technologies affects worker health and wellbeing**: it seems highly likely that there are different outcomes across different job types, different technologies, and different institutional environments.

The emergence of new work forms, including online platform-based work, has also spurred some discussion about decent work. In part, this is because these contracts sometimes fall outside of standard regulatory frameworks, possibly providing less protection for workers – see insert. However, workers in new work forms earn similar hourly wages to comparably skilled workers in traditional contracts[212]. Further, some 80% of workers who are independent contractors or freelancers value the flexibility and independence that comes with being their own boss and report they prefer working for themselves to be an employee. Indeed, new work forms can provide a valuable option for workers. For example, recent studies on Uber drivers show that the option to work through the platform rather than buying a taxi medallion is highly valued[213] , although there also remain disputes over working conditions and pay.

b. Inclusion

Beyond decent jobs being available, challenges for the future of work can arise if certain groups of workers do not have equal access to such jobs: inclusion is therefore an important second labour market consideration.

Rising inequality and polarization threaten inclusion for a number of reasons. For one, **polarized job opportunities imply the middle rungs of the job ladder are less commonly available**: this matters since job mobility directly from the lowest to the highest rungs is much less likely. Further, **increased inequality has been shown to impact intergenerational mobility**: this reflects how likely parents are to transmit their wage income rank to their offspring. If this mobility is high, offspring of low-earning and high-earning parents have similar chances of becoming high-earning; whereas if it is low, offspring of high-earning parents have a strong advantage.

FIGURE IRLESA 32 shows the so-called Great Gatsby curve, using cross-country evidence to show that intergenerational mobility is lower in countries with higher inequality. This can for example occur because a rise in the return to education leads to both a rise in income inequality at any one point in time and a decline in intergenerational mobility because educational attainment is positively correlated across generations. There is an active academic debate on whether intergenerational mobility has declined in recent decades: such long-run patterns are often hard to distinguish because of a lack of comparable data across multiple generations. Yet recent research using high-quality data for Norway has shown that

[212] Katz, L.F., and Krueger, A.B. "The Rise and Nature of Alternative Work Arrangements in the United States, 1995–2015." *Industrial & labour Relations Review* 72 (2019): 382
[213] Chen, M.K. et al., "The Value of Flexible Work: Evidence from Uber Drivers." Journal of Political Economy, forthcoming, (2019).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

intergenerational mobility has decreased, and especially for the very lowest-income workers[214]. As the authors remark, "[we] may suspect that standard measures of intergenerational mobility [...] have failed to pick up trends toward lower rank mobility also in other countries" and that as Norway has "[a] large and ambitious welfare state explicitly designed to ensure equality of opportunities; this may not bode well for countries with less ambitious social and educational policies".

Beyond their tendency to raise labour market inequality, **advancing digital technologies may also have direct impacts on labour market inclusion**. In particular, the fairness consequences of the rising use of algorithmic prediction for screening and evaluation purposes is an area of active research. Standard machine learning has been shown to acquire stereotyped biases from textual data, propagating cultural stereotypes to artificial intelligence technologies that are already in widespread use[215]. While research efforts are being made to debias such algorithms, this is far from standard practice, and **not all experts agree that current debiasing efforts are effective** (Kleinberg et al 2018). Indeed, algorithmic bias has been shown in a number of labour market settings. One example documented how setting an otherwise identical worker's gender to female instead of male resulted in getting fewer instances of an ad related to high paying jobs[216].

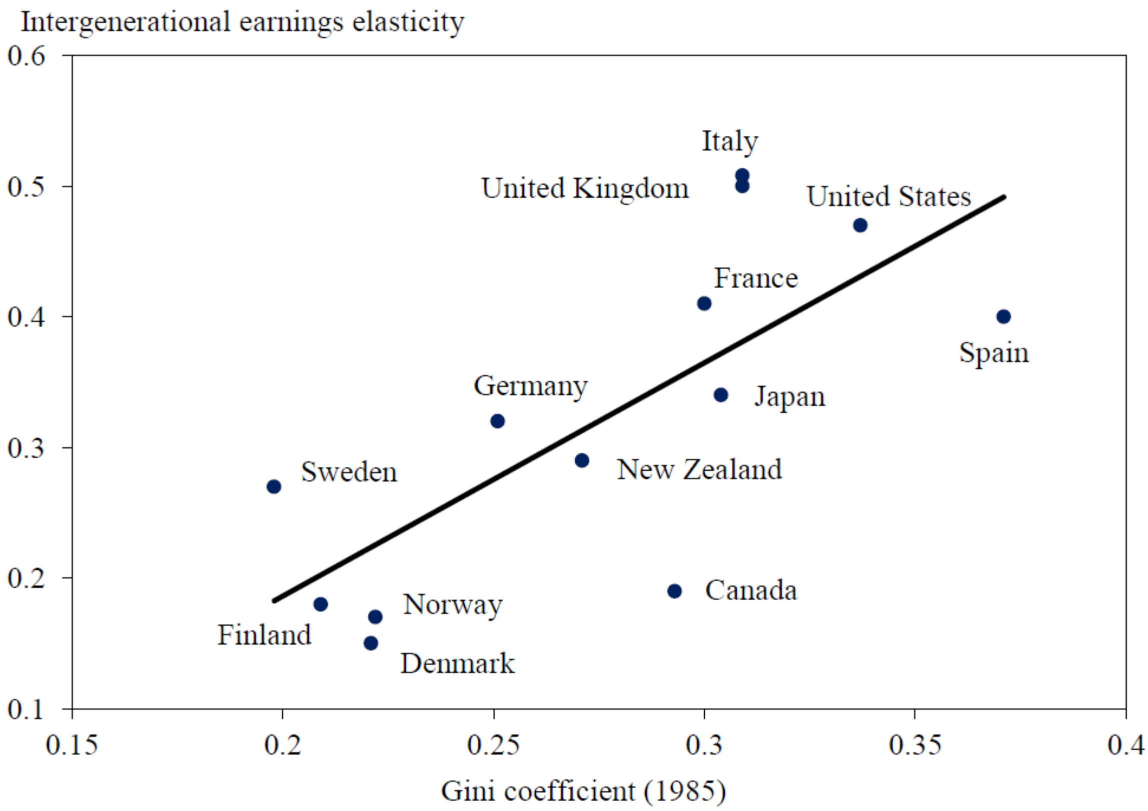


FIGURE IRLESA 32 THE GREAT GATSBY CURVE

---

[214] Markussen, S. and Road, K. "Economic Mobility Under Pressure." *Journal of the European Economic Association, forthcoming,* (2019).
[215] Bolukbasi, T. et al., "Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings." *Adv. Neural Inf. Process. Syst.:* (2016): 4349
[216] Datta, A., Tschantz, M.C. and Datta, A. "Automated Experiments on Ad Privacy Settings," *Proceedings on Privacy Enhancing Technologies*, 1 (2015): 92

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Notes: The Gini coefficient is a measure of income inequality – the higher the Gini coefficient, the more inequality. The intergenerational earnings elasticity is a measure of mobility: the higher, the less mobile people are across generations (i.e. the more the current generation's income rank is determined by their parents' income rank). Source: Krueger (2012).

Further, bias may arise from other causes than biased training data, leading particular worker groups to be disadvantaged. For example, advertising STEM jobs on platforms such as Facebook can lead to the exclusion of women, purely because women are less likely to be shown these ads[217]. This lower exposure to STEM job ads was shown not to be due to women being less qualified or less likely to apply – rather, **it is the algorithms of the advertising market that are biased**. In particular, young women are a valuable demographic on Facebook, making it more expensive to show them ads: ad algorithms then automatically show ads to men, as this leads to higher ad exposure for the same advertising outlay. As such, even neutral ads may have discriminatory impacts. This suggests that algorithmic transparency is not sufficient to address all types of bias – one has to understand the underlying economic mechanisms leading to discriminatory outcomes.

On the other hand, such algorithmic applications should be judged relative to the counterfactual of human judgment, which itself is far from unbiased (see insert). For example, human recruitment often focuses on people already in one's network. But since these networks have been shown to be homogeneous in terms of gender and ethnicity, they can generate more male or white applicants in jobs that are already dominated by males or whites. **Where they exist, comparisons between machine and human judgement suggest machines may be less biased than humans, even when trained on historical data**[218].

> "A lot of people are saying this is showing that AI is prejudiced. No. This is showing we're prejudiced, and that AI is learning it."
>
> *Joanna Bryson, computer scientist at the University of Bath*

For example, research studying hiring for white-collar jobs found that the introduction of machine learning technology yielded candidates that are substantially more likely to pass interviews and receive as well as accept a job offer, as well as more productive once hired as employees[219] .These results were driven by candidates who were evaluated in a biased way when humans had made job offer decisions. In particular, **the candidates suggested by the artificial intelligence were broadly non-traditional**: candidates who graduated from non-elite colleges, who lacked job referrals and prior experience, whose credentials are atypical and who had strong non-cognitive soft skills. In other words, the algorithm was better at picking excellent candidates from among those would not normally be hired.

Other work on the introduction of job testing technologies in low-skilled service sectors has found that this raised productivity and the quality of job matches without harming minority

---

[217] Lambrecht, A. and Tucker, C.E. "Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads." Working paper, (2016)

[218] Kleinberg, J. et al., "Human decisions and machine predictions," National Bureau of Economic Research, (2017); Cowgill, B. "Bias and Productivity in Humans and Algorithms: Theory and Evidence from Resumé Screening." Working paper, (2018)

[219] Cowgill, B. "Bias and Productivity in Humans and Algorithms: Theory and Evidence from Resumé Screening." Working paper, (2018)

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

hiring[220] When faced with similar applicant pools, managers who hire against machine-based recommendations end up with worse average hires. This emphasizes the potential of machine-based algorithms to mitigate errors and biases in human judgment across a variety of labour market domains.

Further, there is an emerging field aimed at auditing algorithms for bias (also known as "AI neuroscience" because it aims to understand AI decisions). This is intended to deal with AI's weakness of being a black box in terms of decision-making: without such auditing there is no way to know what caused the bias. For example, the company Pymetrics, who is using AI to recruit job applicants, explicitly states that they regularly audit their algorithm. **The aim of AI auditing is to increase transparency by highlighting which groups are (dis)advantaged by the algorithm.** However, the case of STEM job ads shows that such audits aimed at algorithmic transparency may not always be sufficient.

All in all, there is reason to be optimistic about machine learning and other AI technologies to help identify and overcome human biases, and thereby improve decision-making and inclusion in labour markets.

The rise of new work forms is a double-edged sword in terms of inclusion. On the one hand, it increases inclusion in a multitude of ways. Although the literature is small, the best available evidence suggests that adoption of more flexible practices can boost productivity, improve morale and work-life balance, and benefit advanced economies[221]. Workplace flexibility, such as part-time work or job sharing, can also facilitate a phased retirement that helps older workers transition slowly out of the workforce, allowing them to take care of health needs and maintain economic security while moving toward retirement. More generally, EU Collaborative Economy and Employment survey data show that workers cite a preference for flexibility as a reason for working in these forms (OECD 2019).

However, new work forms also raise inclusion concerns to the extent that they do not always offer the same institutional protections such as parental leave, paid sick days, and skills training budgets as do full-time employee contracts. This is compounded because the rise of new work forms is disproportionately seen among women and minorities[222]. Lastly, while the platform economy can be a valuable way for older workers to continue earning money in semi-retirement[223] there is evidence that these workers may not always reap the same return to experience as they would outside of the gig economy[224].

To illustrate this further for European countries, experts in 12 European countries were each asked to score the degree of rights' protection workers had in various forms of employment

---

[220] Autor, D.H., and Scarborough, D.. "Does Job Testing Harm Minority Workers? Evidence from Retail Establishments," Quarterly Journal of Economics, 123 (2008): 219; Hoffman, M. et al., "Discretion in Hiring." The Quarterly Journal of Economics, 133(2018): 765
[221] Council of Economic Advisors. 2010. Work-Life Balance and the Economics of Workplace Flexibility. https://obamawhitehouse.archives.gov/files/documents/ 100331-cea-economics-workplace-flexibility.pdf
[222] Katz, L.F., and Krueger, A.B. "The Rise and Nature of Alternative Work Arrangements in the United States, 1995–2015." Industrial & labour Relations Review 72 (2019): 382–416.
[223] Chen, M.K. et al., The Value of Flexible Work: Evidence from Uber Drivers." Journal of Political Economy, forthcoming, (2019)
[224] Cook, C., Diamond, R., and Oyer, P. "Older Workers and the Gig Economy." American Economic Review Papers and Proceedings, (2019).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

relationship[225] . The ratings the experts gave were based on their expertise in relation to the legal and collective bargaining outcomes in their country. They rated forms of employment relationship between 1 (no employment protection) and 5 (total protection) and the extent of legal protection from the principal risks attached to different groups of workers between 1 (lowest level of rights) and 5 (highest level of rights).

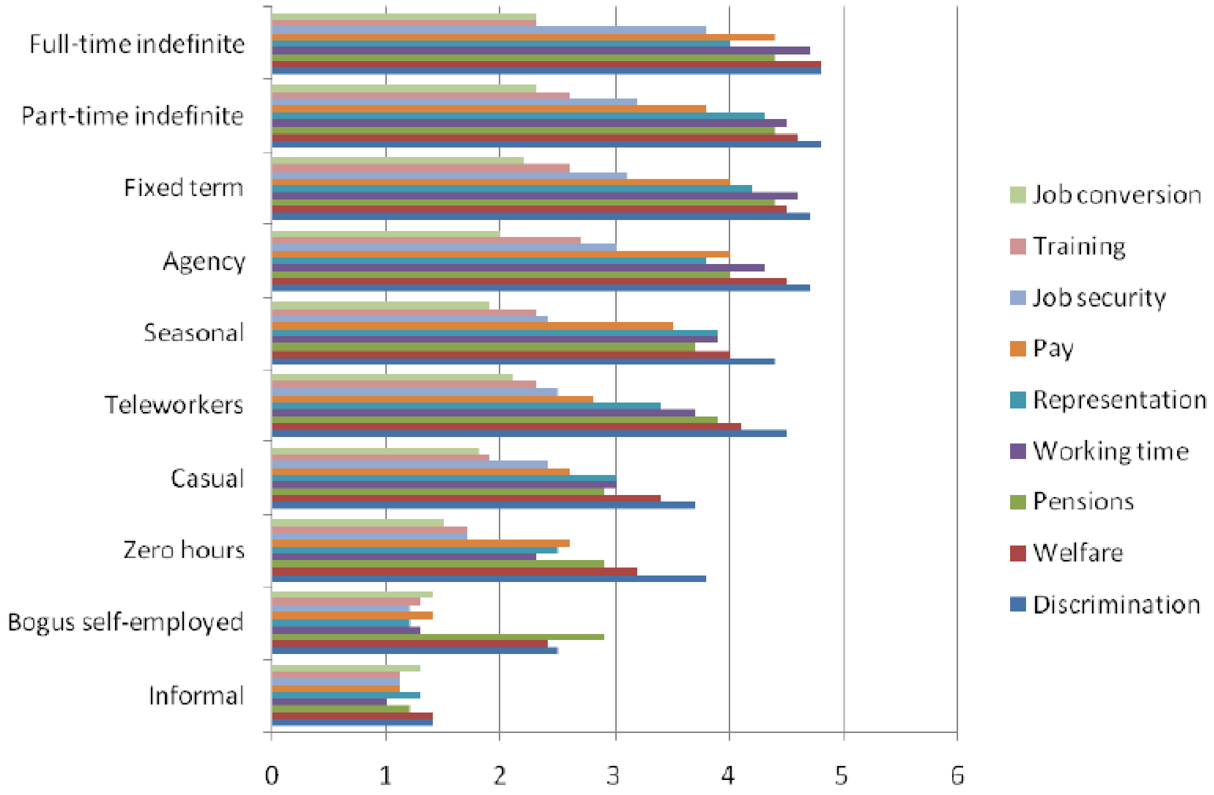


FIGURE IRLESA 33 EMPLOYMENT RELATIONSHIPS AND RIGHTS IN 12 EU MEMBER STATES

Notes. Average ratings are between 1 (no rights) and 5 (full rights). Source: PWSR Ratings (2011)

FIGURE IRLESA 33 displays the average ratings for each category of employment relationship and each set of employment rights. A clear distinction exists in these experts' perceptions between work that is full-time, part-time, fixed term, or even agency, seasonal or telework, and other forms of employment relationship such as casual, zero hours, informal and bogus self-employment. Workers in the informal economy and in bogus self-employment are perceived as having least access to all nine employment rights provisions.

c. Skills

Nobel Prize winner Wassily Leontief famously remarked that "*Computers and robots replace humans in the exercise of mental functions in the same way as mechanical power replaced them in the performance of physical tasks. […] This means that the role of humans as the most important factor of production is bound to diminish—in the same way that the role of horses in*

225 McKay, S. et al., 2012. 'Study on precarious work and social rights', European Commission, VT/2010/084, p. 78, section 2.6, https://ec.europa.eu/social/BlobServlet?docId=7925&langId=en

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

agricultural production was first diminished and then eliminated by the introduction of tractors." (Leontief 1983).

This humans-as-horses analogy is frequently cited as a concern for the future of work, sparking fears of a robotic apocalypse (or "robocalypse") where humans have been made obsolete. In this context, it has also been remarked that "the future of work requires the presence of work". And horses have certainly been almost fully replaced by cars and tractors. However, as shown above, *human* employment is not declining as a result of automation.

A key reason for this is because there is an important difference between humans and horses: namely, that humans are not one-trick ponies. Whereas horse abilities have remained constant over time, human skills have adapted: if we could transport the labour force from 1900 to today, most workers would be utterly unable to perform in today's jobs. This is exactly because human skills have changed in line with what is demanded in the labour market. Therefore, a prime concern for providing access to work is measuring how the skills demanded from workers are changing and investing in their acquisition.

As documented above, **automation and other technological advances have been shown to increase as well as change skill demands in important ways. Further, skilled workers are known to be better able to adapt to changing skill demands over their working lives. As a result, more highly skilled workers suffer shorter unemployment spells and lower income losses from automation following displacement**[226].

Further, research has documented substantial differences in skill requirements of jobs between firms, even within the same sector and occupation, and controlling for the job's formal educational and experience requirements[227]. These skill differences are shown to correlate to differences in wages between regions and between firms. For example, computer programmers in Washington DC earn about 25% higher in Washington, DC, than in Manchester-Nashua, New Hampshire. At the same time, 35% of computer programmer vacancies in Washington, DC, require social skills, compared with only 21% in Manchester-Nashua, and firms with computer programmer vacancies in DC have about 10% higher revenue per worker than their New Hampshire counterparts. This is consistent with computer programmers in DC performing more complex functions, such as strategizing with clients or overseeing co-workers, raising their productivity.

As such, increasing and adapting the workforce's skill level is an important consideration for the future of work – and achieving this also matters for guaranteeing inclusive access to decent work.

6. Policy insights

This section translates the challenges identified in the previous section into policy insights for businesses, governments, social partners, or other actors on labour markets. The objective is not to provide an exhaustive list of policy options or recommendations, but to focus on some of those that could fundamentally change current thought leadership and have a significant impact

[226] Bessen, J. et al., "Automatic Reaction: What Happens to Workers at Firms that Automate?" Boston University School of Law, Law and Economics Research Paper, (2019), https://ssrn.com/abstract=3328877
[227] Deming, D.J. and Kahn, L. B. "Skill Requirements across Firms and labour Markets: Evidence from Job Postings for Professionals," Journal of labour Economics 36 (2018): S337

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

on the future of work. This implies that more traditional policy recommendations regarding educational investments and income redistribution are not reiterated here: while these remain of critical importance, we choose to focus attention on newer areas of thought to raise awareness of the possibilities. The policy insights discussed here are grouped into three main categories:

- New work relations: adapting relationships to today's realities (section 3.6.1.);
- An inclusive society: upgrading the social fabric of our labour markets (section 3.6.2.);
- A skilled workforce: ready to contribute to tomorrow's world of work (section 3.6.3.).

a. New work relations – adapting relationships to today's realities

Digitalization and globalization, together with changes in labour market institutions, regulations and policies, are drastically changing work relations. This has resulted in a reduced need for static hierarchies, fixed desks, and long-term contracts. These are being replaced by flat management and temporary cross-functional teams, virtual workplaces, and shorter contracts (Baldwin 2019). These changes pose several important challenges to how workers, employers, intermediaries, and other actors on the labour market interact, and this section discusses two of these challenges: increasing worker mobility between types of work forms; and investing in inclusive technologies.

i. Increasing worker mobility between work forms

Across OECD countries, new forms of employment are emerging that differ significantly from traditional employment relationships. These new work forms are providing more flexibility for both employers and employees in some sectors. For example, early studies suggest that using the Internet to recruit and to search for jobs is cheaper than doing so by conventional means[228], has a small positive effect on wages[229] and may reduce structural unemployment[230]. Studies have also found Internet use to increase worker mobility[363] and employer-to-employer worker flows[231]. More recently, social media services such as LinkedIn have become means for workers in some occupations to market themselves to potential employers beyond their local markets. In a survey of European freelancers, almost a third said that they found work via social media platforms (EFIP & Malt 2019).

However, in other sectors workers and employers might both prefer to be in a more permanent relationship for several reasons[232] .First, the investment required to set up a business may be too large for a single worker, or even a group of workers. Even if the investments are affordable, some people prefer not to put up with the risk and stress of running a business,

---

228 Freeman, B."The Labour Market in the New Information Economy." *Oxford Review of Economic Policy*, 18 (2002): 288
229 Bagues, M. F. and Labini, M. S. "Do Online labour Market Intermediaries Matter?". In Studies of labour Market Intermediation. Chicago: University of Chicago Press, 2009, 127
230 Kuhn, P. and Skuterud, M., "Internet Job Search and Unemployment Durations." American Economic Review,94 (2004): 218
231 Stevenson, B. "The Internet and Job Search," NBER Working Papers 13886, National Bureau of Economic Research, Inc., 2008.
232 Tirole, J. Economics for the Common Good. Princeton University Press, (2017): 419-420.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

such as doctors or dentists who choose to be employees of a medical clinic rather than set up on their own. Second, from the perspective of a business owner, having someone work for other people may be undesirable too. If the worker has access to confidential information at work, an employer is likely to insist that people work for the one firm exclusively. When the work involves teams, and the productivity of each individual worker cannot be measured objectively (unlike that of a craftsman who works alone), the worker is not always free to organize work as he or she likes. In this case, having several employers could generate significant conflicts over the allocation and pace of work. Third, it may be the case that individual reputations based on ratings do not function well. For example, the quality of individual consultants may be hard to monitor, at least immediately, by their clients, whereas a traditional consultancy employing many consultants may be more efficient at 'guaranteeing' quality.

In short, new work forms are emerging rapidly but it is also unlikely that they will replace all traditional work relationships. That is, **diverse work forms will co-exist in future labour markets. The challenge for business leaders, policy makers and social partners therefore is to ensure that a diversity of work relationships can exist and that workers can easily move between them**.

This mobility of workers between forms of employment can be fostered in several ways, and we briefly discuss three examples. The first example suggests social security that is neutral to and transferrable between different work forms. The second example is to reduce the many institutional hurdles that workers face due to being in a new work arrangement or due to mobility between work forms. The third example is an illustration of how current regulations can reduce worker mobility between work forms, and it proposes one possible solution: It considers shifting the costs for employers of worker turnover from severance pay for workers to a contribution paid into the Unemployment Insurance (UI) system. It explains how this shift could reduce overall costs of job turnover for employers, thereby incentivizing them to create more jobs. At the same time, unemployment benefits ensure the income of dismissed workers despite an increase in labour market flexibility, thereby encouraging mobility of workers between work forms.

On-Board Computers in Trucks and Driver-Owner Relationships

Although digitalization is likely to have contributed to the rise in new work arrangements, it can sometimes also have the opposite effect and favor more traditional employment. One example is the use of onboard computers in trucks (Baker and Hubbard 2003). Many truck drivers work for themselves, which causes some problems. The driver owns their own truck, which is a substantial investment. Drivers are investing their savings in the same sector as their labor, which is risky. In a recession, income from work and the resale value of the vehicle decrease at the same time. In addition, owner-drivers have to pay for repairs, during which time their only source of income is unavailable. If so, why aren't truckdrivers employees of a company that buys and maintains a fleet of trucks? One answer could be that, without the proper technology to monitor the behavior of truckdrivers, an employer needs to worry about the driver not being careful with the vehicle, whereas the independent trucker has every incentive to take good care of it. However, digitalization can alleviate this problem. The trucking company can monitor the driver's behavior using onboard computers, thereby favoring a more traditional

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Neutral and transferable social security

Social protection in case of unemployment, sickness, accident, old age, becoming a parent, and other life circumstances is a fundamental part of advanced economies' welfare systems. Protection is provided through means such as social insurance and social assistance, with the details of these schemes varying between countries. However, these schemes generally tend to assume that a person is either in traditional employment or unemployed. As a result, people engaged in new work forms are often left behind. For instance, self-employed workers are typically individually responsible for enrolling in and paying for sufficient unemployment insurance, disability insurance, and pensions.

To the greatest extent possible, this suggests that social insurance should be neutral to the work arrangement. This could involve portable rights and benefits between different work forms, especially when combining more than one job, as current gaps in transferability may discourage individuals from moving between different work forms.

Reducing institutional hurdles for people in new work forms

Even if social security were made neutral, several administrative hurdles for people engaged in new work forms remain. For instance, registering as a taxpayer, filing taxes, and getting insurance is often much more complicated for workers in new work forms than it is for employees. Self-employed workers are often treated as firms, even though they lack the dedicated administrative resources and know-how of a firm. People engaged in new work forms also face other hurdles such as obtaining a mortgage, as they are unable to present standard salary slips. Platform workers with significant work experience may struggle to prove their experience to a regular employer or educational institution, because they are unable to provide a conventional reference from a line manager.

All this is likely to cause people in new work forms to face penalties consisting of costs and hurdles, reduced access to government services and credit, and reduced mobility to more traditional employment. These penalties should be addressed by ensuring that governments, financial institutions, and employers provide equally accessible services to all workers regardless of their form of employment. For example, temporary work agency Randstad is collaborating with financial institutions to help agency workers get access to mortgage loans. To do this, Randstad issues a statement similar to an employer statement for workers with permanent contracts. This way, worker mobility between work forms will be increased.

ii. Investing in inclusive technologies

As documented in section 3.5.2, emerging Artificial Intelligence technologies have the potential to raise labour market inclusion by reducing reliance bias in human decision-making across a wide range of domains, including many areas of human resources. Further, as a general-purpose technology, AI has the potential to increase productivity, leading to economic growth and rising prosperity.

Investing in AI technologies which are proven to raise inclusion would therefore be good policy across many labour market domains. Examples are educational choice, recruitment and hiring, job search recommendations, job retention prediction, and skill (re)training. At the same time, such investments should be accompanied by careful evaluation, research, and algorithm auditing, to ensure their rollouts do not have unintended

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

adverse consequences for workers or firms. Here, the role of human judgment is likely to be pivotal: in a world where prediction becomes ever cheaper and more accurate, making the right choices and targeting the right outcomes becomes even more critical.

b. An inclusive society – upgrading the social fabric of our labour markets

Though digitalization, globalization, and the rising diversity of work forms have generally been sources of economic growth, they also pose challenges for making sure these gains are fairly shared among all citizens in society. This section discusses two pathways for building a prosperous and inclusive society.

In most advanced economies, labour unions were at their peaks in the 1950s, 1960s and 1970s. Union activity has generally declined since 1980s. One reason for this decline is the process of employment de-industrialization, and the difficulties of trade unions to organize workers in new establishments in particular. The gig economy has also emerged as a small but mostly non-unionized sector, despite significant efforts and some successes by unions to organize gig economy workers. Structural changes in labour markets are thus challenging the social dialogue in many advanced economies. **Social partners should (and in many cases have already started to) review and revise their organizational models and participatory processes in line with what is possible and necessary in today's environment**.

Workers of the Internet Unite?

Coworker.org is a platform that allows people who work for a given company to form a 'network' together. Some networks have tens of thousands of members. Any member can start a campaign to advocate for changes in their workplace, and others can sign their petition. Workers use the platform to campaign on diverse issues, from corporation-wide pay policies to improvements to the local break room. The companies range from large multinational corporations to local firms and gig economy platforms, and participating workers range from standard employees to gig workers. Many campaigns have been successful in starting a dialogue with the

For instance, some workers have self-organized into informal 'social worknets': social media groups in which they support each other and formulate collective responses to workplace issues. These **social worknets are more direct, real-time and provide a basis for a granular social dialogue at the company level**, compared to the wider institutionalized social dialogue.

Thanks to its digital and informal nature, social dialogue can also cross regional and sectoral boundaries, facilitating collaboration between divisions within multisectoral and multinational companies. Furthermore, to realize the gains from data as labour, data workers will need some organization to vet them, ensure they provide high-quality data, and help them navigate the complexities of digital systems without overburdening their time.

c. A skilled workforce – ready to contribute to tomorrow's world of work

Digitalization and globalization are rapidly changing the demand for workers' skills and task competencies. This way, they are contributing to skill mismatch and shortages that require investments in employee training. In light of these challenges, several actors, including the OECD, have focused on **the question of how to achieve a better alignment of skill**

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

supply and demand, with a focus on: i) understanding how countries collect and use information on skill needs; ii) investigating cost-effective training and labour market policies to tackle skill mismatch and shortages; iii) studying the incentives of training providers and participants to respond to changing skill needs; iv) setting up a database of skill needs (OECD 2016).

Despite these efforts, several important challenges remain. We briefly discuss three of those: Developing 'A high-resolution framework of workplace skills'; Incentivizing 'Intermediaries to increase investment in worker skills'; and 'Skills for displaced workers'.

i. A high-resolution framework of workplace skills

One possible definition of a worker's skills is the formal education that s/he received during formal schooling when young. One can then think of how digitalization and globalization have changed the demand for workers with more relative to less formal schooling (see for example [233]).

However, a more precise view would be that digitalization and globalization are changing the demand for tasks that workers do on-the-job because some tasks can be automated or offshored, but others cannot. Consequently, **digitalization and globalization will change the demand for workers with different levels of formal schooling only indirectly through changes in on-the-job task requirements for workers**.

This decoupling between workers' formal schooling levels and their task competencies poses the question of how to properly define skills. Consequently, several different classifications of skills (e.g. years of formal schooling, occupational or sector experience, tasks done in an occupation or soft skills such as personality traits) have been developed recently, and some studies have successfully used them to examine changes in skill usage in the labour market as a whole due to economy-wide digitalization, globalization, migration or aging.

Although these recent classifications go beyond formal schooling as a measure of skill, they are not always informative for individual workers, each with their specific formal schooling, task experiences and other competencies, as well as for individual employers, each with their specific workplaces and related skill needs.

For example, Civil Engineers and Medical Doctors are both professions that fall into the same conventional labour categories: They both have high educational requirements, make high wages, and require cognitive non-routine labour. Yet, their skill sets are largely non-transferable. To explain why Civil Engineers are unlikely to become Medical Doctors—and to explain where skill sets might limit other

Constructing the 'Skillscape'

An illustration of such a high-resolution framework for workplace skills is given by Alabdulkareem et al. (2018). They use high-resolution occupational skill surveys carried out by the US Department of labour to construct a 'Skillscape'. By examining how pairs of detailed skills co-vary across occupations, they identify pairs of skills that tend to be bundled together. For example, Spatial Orientation and Peripheral Vision are both required in occupations such as Bus Drivers, Light Truck or Delivery Service Drivers, Taxi Drivers, or Parking Lot Attendants. On the other hand, Mathematics and Programming have high complementarity in a very different set of occupations. They then employ a data-driven approach to map how each detailed skill is required in combination with each other detailed skill in an occupation, workplace, sector or region.

233 Goldin, C. and Katz, L.F. 2009. The Race Between

This project has received funding from the Europe
research and innovation program under grant agr

workers' career mobility—a higher-resolution framework for specific workplace skills is needed.

An improved resolution of workplace skills and their complementarities can **shed new light on where bottlenecks limit career mobility due to skill mismatch. Or, how workers can leverage their existing skills to grow their skill set and open-up new career opportunities**. A high-resolution framework for workplace skills also has the potential to inform worker re-training programs aimed at maintaining employment opportunities in an increasingly competitive economy due to digitalization and globalization.

ii. Intermediaries to increase investment in worker skills

A second challenge is to better understand why there is underinvestment in on-the-job training. Underinvestment in on-the-job training is most often explained by temporary misalignments between demand and supply, in large part driven by the business cycle. However, given the persistency of skill gaps across countries and over time, it is likely that there are other and more fundamental reasons why there is underinvestment in training.

Consider the rise of new work forms with more flexible contract durations, such as agency workers, on-call workers, contract workers, and independent contractors or freelancers. For these types of contracts, it becomes riskier for workers and their employers (for whom these workers actually provide their labour, not the agency that matches these workers to employers) to credibly commit ex-ante to share in both the initial costs as well as the later benefits of training. **For the worker the risk is that his or her current contract is not extended by his or her employer after she has invested in training but before she can earn a return on his or her investment. The same is true for the employer. Consequently, both the worker and the employer will be reluctant to invest in on-the-job training**.

In other words, coordination problems between workers and firms lead to a market failure in skill acquisition, and **policies involving third-party intermediaries that share in the costs and benefits of training are required to increase training and reduce skill gaps**.

For example, temporary work agencies have an incentive to invest in on-the-job training if they can recoup their training costs from employers by charging a wage premium for trained workers. Employers are willing to pay this wage premium because they are no longer faced with the uncertainty about a worker's skills (given that they know exactly know what training these workers received from the temporary work agency). Finally, workers do not have to invest in their own training and receive part of their increased productivity in terms of higher wages.

There are several ways in which private or public intermediaries can provide training in practice. Examples include training provided by Public Employment Services (PES), outplacement offices that assist displaced workers in finding new jobs (and that are funded by companies that mass lay-off workers), or temporary help agencies. Typically, these intermediaries focus on the re-employment of job seekers who are currently unemployed or without permanent jobs.

But intermediaries can also train workers who have a permanent contract with their employer, especially when the labour market is tight such that the costs for companies to search for new workers is high. Many companies may not have a clear view of their own employees' talents. Specialized intermediaries can help workers find logical, reasonable career paths within the organization, while also boosting worker productivity for employers. In doing this, intermediaries can use companies' underused repositories of data on a person's skills, internal

This project has received funding from the European Union's Horizon 2020
research and innovation program under grant agreement No 780073

reputation, learning capacity, ambitions and interests. For some of the training, intermediaries could **partner with institutions specialized in formal schooling to hand out short-term badges, nanodegrees of less than a year, one-year master's degrees, or even a 'skill passport'**.

iii. Skills for displaced workers

Changes in the job structure and in the task content of existing work lead to job reallocation and displacement. For some time, the academic consensus was that job training programs were not very effective at mediating these problems. However, a meta-analysis of 97 job training program evaluations from 1995 to 2007 that more frequently employ experimental variation and higher quality data has led to the opposite conclusion: **training programs are associated with positive medium-term impacts, although in the short term they often appear ineffective**[234].

Recent work has studied the impact of Trade Adjustment Assistance (TAA)—a large social insurance program that couples retraining incentives with extended unemployment insurance (UI) for workers displaced by trade shocks[235]. This program is found to be effective: TAA-recipients have $50,000 greater cumulative earnings ten years after displacement. These effects are driven by both higher incomes and greater labour force participation immediately after training. This suggests that training can play an important role in overcoming adjustment for displaced workers. Further, **returns are concentrated in the most adversely affected regions**, where workers are more likely to switch industries and move in response to TAA training.

This work also indicates that training policies are potentially underused: a higher **availability of such programs and awareness among workers of their eligibility may help alleviate some of the transition costs** associated with structural labour market change and reduce labour market mismatch across skills as well as regions in the process.

7. Concluding remarks

Several key trends, including automation, globalization, urbanization, a rising diversity of work forms and demographic change are generating significant uncertainty and anxiety about the

Key take-aways from 'Chances and Challenges':

Key take-aways from 'Policy Insights':

- Technological progress, international trade, and diverse work arrangements have increased productivity, economic wealth and opportunity.
- However, these forces are also accompanied by several challenges related to the future of work: rising inequality, job reallocation, and skill gaps.
- While technological advances have not decreased total employment, they have led to increased wage inequality, displaced individual workers from their jobs and hollowed out the skill distribution, leading to job polarization.
- International trade and resulting import competition have had similar distributional impacts on workers, as well as affected regions unevenly.
- Economic opportunity has been increasingly concentrated in cities, and in favor of skilled workers.
- Science, Math, Engineering and Technology (STEM) skills as well as social skills are in high demand in our labour markets.

234 ((20

235 Adj

future of work. Here, we have outlined current scientific thinking on these changes and the chances and challenges they pose for our societies, with the aim of channelling the sense of collective concern into a discussion about how to harness these changes for social benefit.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

5. Providing an overview of financial and tax law, labour and social security law

1. Introduction: Ethics, Law and Inclusive Robotics

Law needs Ethics as a foundation for its rules and Ethics needs Law to give more force to its conclusions. Roboethics, understood as the set of criteria or theories, which are formulated as a response to the ethical problems arising from the design, creation, development and use of robots[236], warns us that Robotics gives rise to unique issues that go beyond those which are common to all the so-called 'emerging technologies'. These include the relationship or interaction between human beings and machines (uses and limits of robotics) and the moral status of robots, i.e. their possible consideration as moral agents. There is also debate as to whether, when robots possess certain characteristics, which make them similar to humans, they cease to be an object and can be considered a subject[237]. Robotics presents several issues that affect its political and legal implications. On the one hand, these implications should guide public authorities in their job to oversee the introduction of robots in a socially responsible manner; in such a way that society perceives that they are necessary and useful to people and are accepted by them. On the other hand, they should facilitate a situation whereby those authorities and the many groups with a social and economic stake in the matter can work towards legislation, which addresses not only the specific characteristics intrinsic to robots but the type of problems they pose.

As part of the process of eliminating existing barriers, the need to legally regulate robotics is obvious, although there is no unanimous position as regards the manner and scope of that regulation. **Whether we follow the path of regulation, which we can call *hard law*, or whether we opt for the path of *soft law,* in both cases we need to ensure transparency and accountability regarding the social and economic costs and benefits**.

Should the path of mandatory regulation be the one finally chosen, it will be necessary to bear in mind not only the multiplicity of technological applications but also the range of legal problems they generate and the difficulty of fitting them all into a uniform paradigm. That is why approaches which until recently were different, those of the United States[238] and Europe

---

[236] Veruggio, G.: *The EURON Roboethics Roadmap*, 2006, available at: http://www3.nd.edu/~rbarger/ethics-roadmap.pdf.
Also, Veruggio, G. and Operto, F.: 'Roboethics: Social and Ethical Implications of Robotics', in *Handbook of Robotics*, Siciliano, B. and Khatib, O. (eds.), 2008. By the same authors: 'Roboethics: a bottom-up interdisciplinary discourse in the field of applied ethics in robotics', *International Review of Information Ethics*, Vol. 6, No. 12, 2006, pp.2-9.
[237] De Asís, R.: *Una mirada a la Robótica desde los Derechos Humanos*, Edit. Dykinson, 2014, pp. 41-43, 74, 75.
[238] Calo, R., Froomkin, A.M. and Kerr, I. (Eds): *Robot Law*, EE Elgar, 2016; CALO, R. described Robots some time ago as 'entities' and consider them to be artificially intelligent devices with 'cognitive' faculties, in Calo, R.: 'La robótica y las lecciones del Derecho cibernético', *Review Privacidad y Derecho Digital*, núm.2, 2016, pp.155-157.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

[239], can now come together around regulations or minimum mandatory rules that establish the necessary balance between facilitating robotic technological development and protecting the values desired by humans. National reports in various countries, some in the general context of digitalization, others related more specifically to Artificial Intelligence, and others relating even more specifically to Robotics in connection with AI (e.g. Economic and Social Council in Spain; Italy; France; United States; Great Britain; and Japan) stress the interaction between robotics and human being and the social impact thereof by reminding us that 'we are creating systems to help us; we are not creating life'[240].

In order for the Law to be able to adopt measures in this regard, we need to clearly define the problem and the challenges that have to be addressed, based on the set of common general principles in the acquis of the European Union (which help to construct what has started to be called 'fully-fledged digital citizenship'[241]). Today the framework of the European Pillar of Social Rights pays special attention to human abilities. Those abilities that are affected by robots, and those that, though it may seem paradoxical, can be fostered by robots so that people can devote themselves to performing inherently human activities, i.e. those related to emotions, awareness, reflection, abstract processing, personality and free will. It is a question of **moving forward on the basis of 'the principle of caution'[242] as applied to freedom of scientific investigation and, beyond rules of 'technological neutrality'[243] (which cannot become an end in itself), to implement the principle of socially and legally responsible technological innovation.**

2. Labour Law implications of Robotics

posed by technology in general, and automation in particular, in the world of work[244].

From this perspective, studies and reports have adopted varying approaches. Two expressions

[239] Palmerini, E.: 'Robótica y Derecho: sugerencias, confluencias, evoluciones en el marco de una investigación europea', *Review de Derecho Privado,* Universidad Externado de Colombia, n.° 32, January-June de 2017, p. 80; Leenes, R., Palmerini, E., Koops, B.J., Bertolini, A., Salvini, P. and Lucivero, F.: 'Regulatory challenges of robotics: some guidelines for addressing legal and ethical issues' , Journal Law, Innovation and Technology,, 2017, Vol. 9, No. 1, 1-44, p.12.
240 As discussed in Nisa Ávila, J.A.: .: 'Robótica e Inteligencia Artificial ¿legislación social o nuevo ordenamiento jurídico', El Derecho.com, Lefebvre, 2016, http://tecnologia.elderecho.com/tecnologia/internet_y_tecnologia/Robotica-Inteligencia-Artificial-legislacion-social-nuevo-ordenamiento_11_935305005.html, (consulted in March 2018).
241 Cortina, A.: 'Ciudadanía digital y dignidad humana', opinion article in El País, 26 March 2018, which considers it to be a fair and essential requirement that digital citizenship be at the service of autonomous and vulnerable people.
242 'Which supports the adoption of protective measures with regarding to certain products or technologies which are suspected of posing a serious risk even though there is no scientific proof of this', De Asís, R.: Una mirada a la Robótica…, cit., p. 68.
243 Leenes R., et al.: 'Regulatory challenges of robotics…', cit., p.12.
244 From ILO, Nieto, J.: 'El futuro del trabajo que queremos y el Derecho del Trabajo', in Ius labour 3/2017, https://www.upf.edu/documents/3885005/140470042/1.+Editorial.pdf/406c3008-6ef9-7ed6-f4ee-8f754c9adc31 (last access March 2018).
The ILO Initiative for the Future of Work can be consulted in http://www.ilo.org/global/topics/future-of-work/lang--es/index.htm; And the national syntheses in http://www.ilo.org/global/topics/future-of-work/WCMS_591507/lang--es/index.htm.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

have become generalized: techno-pessimism[245], which identifies the risks posed by robotics to employment and concludes with a drastic prediction of job disruption or destruction; and techno-optimism[246], which identifies the challenges and opportunities presented by the robotization of the labour market.

In any event, we should bear in mind that the studies are carried out in a socio-economic context of employment precariousness (we speak of 'The Precariat' in sociological and economic terms[247] as an emerging social class which lives in a state of economic and professional insecurity), together with a high level of unemployment, and in a demographic context of an ageing population and high levels of life expectancy.

In spite of all this, the fundamental question is whether the increase in company productivity and competitiveness (which does not appear to be in doubt) will also be accompanied by an increase in the quantity and quality of human employment (for human and environmental well-being). And in this respect, apart from the 'replacement' effect, problems arise in relation to the transitional period we are still going through, including those which affect the working conditions of humans: pressure on salaries, particularly on those of less skilled workers; the move towards decentralized production; the reallocation of jobs and tasks; and the effect of technological unemployment.

There is no snapshot available that covers all sectors, all kinds of work, every skill level, all markets or all countries[248]. Therefore:

- The analysis requires a temporal perspective but should avoid making long-term projections exclusively. We need to promote a situation in which the absence of constraints on technological innovation goes hand in hand with the principle that automation and robotics must permit employment to be focused on 'jobs that add greater value'; this means committing ourselves now to the development of technological competences and balancing the two needs, one which stems from growth and competitiveness, and consequently the adoption of technology, and the other which minimizes disruption in the labour market to prevent social inequalities.

---

245 For example, from USA, Acemoglu, D. and Restrepo, P.: 'Robots and Jobs: Evidence from US labour Market', NBER Working Paper No. 23285, 2017, in http://www.nber.org/papers/w23285 (last access March 2018); also, Frey, C.B. and Osborne, M.A.: 'The Future of Employment: How susceptible are Jobs to computerisation', Technological Forecasting and Social Change, Vol. 114, No. C, 2013, pp. 254-280, which analyses 702 jobs and claims that 47% of jobs in the U.S. are at risk of being computerized or robotized. The projections of these authors are used in the 2018 BBVA report; Doménech, R. et al.: How vulnerable is employment in Spain to the digital revolution? https://www.bbvaresearch.com/wp-content/uploads/2018/03/Cuan-vulnerable-es-el-empleo-en-Espana-a-la-revolucion-digital.pdf (last access March 2018)

[246] McKinsey Global Institute. 2019. Globalization in Transition: The Future of Trade and Value Chains. www.mckinsey.com/mgi; PwC (2018): Will robots really steal our jobs? An international analysis of the potential long term impact of automation, available at https://www.pwc.es/es/publicaciones/tecnologia/assets/international-impact-of-automation-2018.pdf

[247] Standing, G.: *The Precarious: a new social class*, Edit. Pasado y Presente, 2013, Barcelona, 2013. Founder of the Basic Income Earth Network that postulates universal basic income.

[248] United Nations: *Trade and Development Report*, 2017; in particular, Chapter III, *Robots, Industrialization and Inclusive Growth*: 'This discussion shows that disruptive technologies always bring a mix of benefits and risks. But whatever the impacts, the final outcomes for employment and inclusiveness are shaped by policies'; p.60.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

- The digital breach and the associated social cost must be dealt with, paying special attention to the vulnerability of certain groups (based, among other factors, on gender and age).

- A 'safety net' for the transition period is required, and it demands the proposal of related political and legislative measures.

To solve these problems, is it necessary to make changes to the labour Law (and also the Social Security Law), i.e. in the regulatory framework of labour relations in the robotized neo-technological context? The answer must take into account the multi-functionality of the legal definition of robots (autonomy, physical configuration and ability to interact with workers[249]), and the involvement of entrepreneurs and workers' representatives, especially trade unions, in order to take decisions, which are based on consensus between them and accepted by the public authorities. Lastly, **a minimum, albeit necessary, mandatory legal intervention is needed to ensure a balance between entrepreneurial freedom and the function and purpose of labour Law, especially with regard to the protection and guarantee of human work**. This last part affects two areas: the concept of 'worker' and forms of employment in the robotized labour market; and the working conditions. And, for both these, we need to reflect urgently on some of the problems and the possible legal solutions in the transition phase.

a. Concept of worker and forms of employment in the robotized labour market

Let us pose these issues in question form and briefly indicate a solution:

1. Impact on the concept of salaried employee? There is no direct reconsideration; however (especially as regards the solution to the problem of liability for damages) it could indirectly affect 'Labor Law evasion'.

2. Reconsideration of the concept of disabled worker? The concept of disability may change in the future and this means it will be necessary to **rethink the legal concept of occupational integration based on the distinction between therapy and the improvement of capabilities**.

3. Due to the replacement effect, will there be a change in the traditional forms of work, whether the person has an employment contract or is self-employed? Contractual diversity should be maintained, although it should be committed to guaranteeing social, economic and labour rights, i.e. by adopting **measures to prevent and correct regressive segregation**. It is precisely this, which is, or should be, a priority of the regulatory framework of labour relations in this phase of robotization for the purpose of preventing worker polarization, or digital polarization as a synonym for precarious employment, even if it is necessary to accept a new division of labour, between digital labour and human labour.

[249] Del Rey Guanter, S.: *Robótica y su impacto en los Recursos Humanos y en el marco regulatorio de las Relaciones Laborales*, Edit.Wolter Kluwer, 2018, Chapter 3, describes these elements in the following way: autonomy, acquired by means of sensors and/or data sharing and analysis (interconnectivity); physical configuration, i.e. a minimum physical materialization, which entails the possibility of physical movement applied to the work, with the possibility of total or partial displacement; interaction with the environment with reponse capability by means of the appropriate programming.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

4. As far as '*robot workers*' are concerned, electronic personality or Impersonal Capable Entities (the original Spanish term is *Entes Capaces No Personales*, ECNP)[250]labour law allows us to take a **critical position regarding a robotic personality or ECNP, i.e. regarding legal recognition of a new category of capable subjects or entities, and even more so when it comes to defining a relationship, such as a labour relationship**, which is both proprietary and personal in nature, an exchange which is permanent and stable, and above all, voluntary. Voluntariness and awareness are something that robots are unlikely to ever achieve, even if they are endowed with AI.

5. If this were the case, could they have employment status? For this to be possible we would have to stop considering them as a mere instrument of work and start to consider them as technologically complex tools. They would not, however, be subject to labour rights and obligations due to the absence of the characteristic of 'voluntariness'.

6. With regard to the change in the model of work and the polarization of workers, we would counteract the digital breach (and the effects thereof on specific groups of workers) by means of actions and measures geared towards safeguarding the distribution of work and the protection of the most vulnerable workers within the framework of the principle of equality and non-discrimination in employment. Would it be necessary to go further in **the principle of *equality and robotic non-discrimination*, which would also require, as preventive and corrective measures, the implementation of rules on positive or affirmative action in favour of human workers**? This would require us to address the following issues, among others:

- A 'human **quota'** in companies? A measure which could be implemented in the transition period on an exceptional, extraordinary and temporary basis.
- Financial incentives for entrepreneurs? **Incentives** for retraining and relocating workers could be adopted.
- Technological requirements during the selection process, i.e. can these conditions or the **technological adaptability (robotics) of the workers** be used **as selection criteria**? This means considering whether the assessment of such requirements in the framework of the right to equality and non-discrimination should be applied restrictively during the transition period and progressively extended for companies which exhibit a high level of robotization in their production processes or in those which are immersed in a process of robotization.

b. Working conditions and the impact of Robotics

There are two aspects to be analysed here: one is related to the 'replacement effect' of robotics; the other is related to collaborative or cooperative work with robots.

If we **analyse the impact of robotics from a negative perspective, i.e. the replacement of the human workforce by robots (botsourcing[251])**, and if we maintain

[250] García Mexía, P.: 'Entes Capaces No Personales. ¿Hacia una personalidad para los robots', in https://www.automatas.tech/pablo-garcia-mexia-colaboracion.html (last access March 2018).
[251]Waytz, A. and Norton, M.I."Botsourcing and Outsourcing: Robot, British, Chinese, and German Workers Are for Thinking - Not Feeling – Jobs" , Emotion, 14 (2014): 434-444, available at https://www.hbs.edu/faculty/Publication%20Files/waytz%20norton_a358958c-3b94-4f25-bb8c-7d10605738d8.pdf

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

that this effect will be inevitable in the short term, it is worth considering whether or not robotics is a technological phenomenon which impacts on the right to non-discrimination in terms of working conditions, the main legal 'safety net' for workers. And this, in our opinion, allows us to assess not only arbitrary behaviour on the part of entrepreneurs but also those forms of behaviour which can, 'aseptically', be considered as linked to an entrepreneurial right to technological-robotic innovation in the company. If this impacts on certain groups of workers for reasons which include gender and age, robotization could be considered as a cause of indirect discrimination (with robotization being identified as a neutral criterion which has or can have an adverse impact on, or prove detrimental to, one of the vulnerable groups identified for one of the causes where discrimination is prohibited) and would activate the guarantee process provided for in EU regulations, particularly Directive 2000/78.

This leads us immediately to analyse the 'technical cause', arising from robotization, in such a way that measures are devised which are geared towards workers' remuneration (which is none other than the effective and real application of the right to equal pay for equal work; *human work as work of equal value to that of the robot*). And also, in relation to eliminating vacant positions (or dismissals or job changes) arising from robotization.

From this latter perspective, we need to analyse *botsourcing* and its labour implications, bearing in mind that the robot is, or can be considered to be, a 'technical improvement' in the company, which affects aspects of internal and external flexibility. This will require the legislator to specify, at the regulatory level, the technical reason arising from robotization (in the sense of a massive incorporation of robots and replacement of workers); the implementation of well-thought out and balanced measures which limit the impact of dismissal (need for preventive measures, prior to termination, and immediate corrective measures, need to relocate the worker affected). As far as the preventive measures are concerned, one of the issues posed in this respect is the definition of 'reasonable adjustment' as regards the necessary readjustment of the job of the worker who has been replaced by robots with the subsequent problem of whether these adjustments can be considered an 'excessive burden' for the entrepreneur. In our opinion, it would be necessary to assess the reasonableness or otherwise of the adjustment in relation to tax incentives and subsidies for innovation which have been granted to entrepreneurs. And, lastly, a highly controversial aspect is the issue of 'auxiliary aids', particularly the use exoskeletons which able-bodied people could request of entrepreneurs in order to improve their personal skills[252].

If we **analyse the perspective of collaboration, cooperation, interaction of human workers and robots**, a number of issues arise, particularly in connection with:

- Health and safety at work, new psychosocial risks, professional retraining of workers and working time (assessment of the digital disconnection of the robots and its influence on pay and the calculation of working time).
    - The worker's right to privacy and, in particular, 'robotic' surveillance.
- Intellectual property rights/patent right when the worker 'trains' the robot.
    - Advanced perspective of entrepreneurial secrecy.

[252] Hoder, C. et al.: 'Robotics and law: Key legal and regulatory implications of the robotics age (part II of II)', *Computer law & Security Review*, 32 (2016): 557

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

3. Financial and Tax Law Implications of Robotics

a. General discussion

There is an obvious need to adapt the financial and tax rules to face the challenges posed by robotics[253]. This process of adaptation has just begun and will have to be maintained in the future. Apparently, as a machine includes new functions - similar to human capabilities, it is referred to as 'robot'; once we differentiate them from the ones attributable to humans, it is called again 'machine'[254]. Those functions may help to dynamically define a job.

In this changing environment, an interactive robot (with different levels of complexity) may supplement or substitute a human being for the development of some tasks[255]. In the coming years, the economy and the appearance of various jobs will possibly experience alterations, particularly regarding the remuneration and/or the number of required personnel in a specific field. Hence, **the design of new public policies will surely have an impact on the financial legal order; and conversely the latter will probably impose restrictions of the formers' feasibility**, in the light of the expected qualitative and quantitative transformations[256].

Undoubtedly the types of work have changed throughout history. From now on, the degree of robots penetration may be gradually increased in a given job (considering the total, partial or nil reservation of some tasks for human participation). As the use of robotics in our society evolves, the legislation will have to consider how individuals and companies use robots[257].

Nowadays, world trends already show a continued proliferation in the use of robots[258]. In this context, many questions arise: should taxation intervene to slow down the spread use of robots, or to finance new labour opportunities? Should the companies that invest in robots as substitutes for workers' pay taxes and Social Security contributions for the benefits obtained due to the increase of productivity?[259] Which would be the competent authority to receive those payments? In the case where several States are involved, because the robot interface may act in a multiple off-line real (on top of virtual) environment, these States will have to cooperate in the design of coherent measures to safeguard social protection, taking advantage of the pace and volume of information flowing across borders.

The main issue, in the middle of this digital revolution, is how to allocate rights and responsibilities among human beings for the actions of non-human beings, fighting inter-personal and inter-

[253] Grau Ruiz, M.A.: 'La adaptación de la fiscalidad ante los retos jurídicos, económicos, éticos y sociales planteados por la robótica', *Nueva fiscalidad*, 4 (2017): 35
[254] Professor Bernard Roth elaborated this variable concept at Stanford. García-Prieto Cuesta, J. '¿Qué es un robot?' (Chapter I), in Barrio Andrés, M. (dir.): *Derecho de los robots*, La Ley-Wolters Kluwer, 2018, p. 33.
[255] García-Prieto Cuesta, J.: '¿Qué es un robot?', cit., p. 39.
[256] Froomkin, A.M.: 'Prologue', in Barrio Andrés, M. (dir.): *Derecho de los robots*, La Ley-Wolters Kluwer, 2018, p. 22.
[257] Barrio Andrés, M.: 'Del Derecho de Internet al Derecho de los robots', in Barrio Andrés, M. (dir.): *Derecho de los robots*, La Ley-Wolters Kluwer, (2018): 71-73
[258]Available at https://ifr.org/news/world-robotics-survey-service-robots-are-conquering-the-world-/ (last access 1 April 2018)
[259] Segura Alastrué, M.: 'Los robots en el Derecho Financiero y Tributario' (Chapter VII), in Barrio Andrés, M. (dir.): *Derecho de los robots*, La Ley-Wolters Kluwer, 2018, p. 173.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

national inequality (as intellectual and financial capital providers are supposed to enjoy the greater benefits)[260].

The risks of job destruction and structural unemployment have propitiated the discussions around the universal basic income[261]. The International Monetary Fund (IMF) states that this instrument helps us face the acceleration of the decline in income and the uncertainty arising from the impact of the technological evolution on employment. However, there are doubts regarding its affordability and costs, particularly if they are displacing other priority expenditure programs that actually promote inclusive growth. Additionally, some critics argue that separating this income from the participation in the workforce would be problematic.

A number of experiments, considering distinct features of the **Universal Basic Income** (UBI), are being carried out by the private and the public sector, to determine its impact on the individual and the society[262]. This movement may be understood as an extension of the Social Security network, an alternative to bureaucracy and public intervention, or a manner to maintain social peace. Evidently, behind the supporters of universal basic income one may find completely different views, but they all face the same problem: how to finance it? [263] For instance, in Switzerland the proposal of universal basic income was rejected by citizenship because of its costs, among other reasons[264].

Clearly, the means of financing this measure will influence its net redistributive impact. The IMF has recently declared that the possibility of replacing the current system of social protection with a basic universal income will depend on the performance of this system, the governmental administrative capacity and the prospects to improve targeting. Presently, in developed countries, it seems preferable to reinforce the existing systems, directly eliminating the gaps in the coverage nets - caused by the participation rules or the incomplete take-up and focus on the proper design of salary subsidies to incentivise the work of low-income workers.

The IMF has recognised that a sound motivation to adopt a universal basic income could be 'enhancing income insurance in the context of rising job insecurity due to technological change and automation or building public and political support for structural reforms, such as eliminating food or energy subsidies and broadening the consumption tax base'[265].

---

[260] Gupta, S.; Keen, M.; Shah, A.; Verdier, G. (eds.): *Digital revolutions in public finance*, International Monetary Fund, Washington DC, November (2017): 11-12. Available at http://www.elibrary.imf.org/view/IMF071/24304-9781484315224/24304-9781484315224/Other_formats/Source_PDF/24304-9781484316719.pdf (last access 1 April 2018).
[261] No country has approved a UBI for the entire population to date. International Monetary Fund: 'Tackling Inequality', *IMF Fiscal Monitor*, October 2017, p. 3-4. Executive summary available at http://www.imf.org/en/Publications/FM/Issues/2017/10/05/fiscal-monitor-october-2017 (last access 1 April 2018).
[262] E.g. 'Y Combinator' in San Francisco, or Canada (Ontario), Finland and the Netherlands.
[263] A 70% of Finnish say that they are in favour of a basic income, however this percentage is 35% when they are told that taxes would be raised to finance it. As the cost of the experiment is limited to 20 million euros, it is impossible to guess how the taxpayers' incentives could change, if the model would become general. *El coste de mantener las prestaciones de un Estado como Noruega*, BBVA, June 2017.
[264] Segura Alastrué, M.: 'Los robots…', cit., p. 178.
[265] International Monetary Fund: 'Tackling Inequality', *IMF Fiscal Monitor*, October 2017, p. X. Executive summary available at http://www.imf.org/en/Publications/FM/Issues/2017/10/05/fiscal-monitor-october-2017 (last access 1 April 2018).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Other voices call for **modelling the technology instead, for policies to increase economic growth and improve jobs for all, by investing in education, research and development, and infrastructures**[266].

Noticeably, the effects of Universal Basic Income will vary depending on the country. The developing countries may use it to quickly reinforce their safety nets, but this would require efficient and equitable increases in taxes or cuts in spending. Whilst in developed countries, the UBI might result in a reduction of benefits for low-income households.

UBI, financed through the general budget, might become a disincentive to search for employment and also produce a call effect. Therefore, in reality it is likely to be transformed into a conditional basic income, limited subjectively (e.g. a specific group of beneficiaries depending on age), quantitatively or temporally. Others believe that a specific tax on robots could be a collateral effect and argue that it should be only a transitional tax with earmarked revenue[267]. An alternative could be a negative tax on income, to substitute the benefits not depending upon contributions, for all the citizens below the poverty line[268].

Legislators should not look for extreme solutions when the social panic threaten traditional legal institutions whose continued existence must be preserved. **It is necessary to reach consensus (based on the principle of solidarity) while shaping the relations between technology, social processes and regulation**.

b. Tax Incentives for Training in A Socially Responsible Transition

The International Labour Organization (ILO), Global Commission on the Future of Work, has just published its report Work for a brighter future[269]. It calls for the use of technology in **support of decent work and a "human-in-command" approach**. The discussion about technology in the future of work has mainly focused on the issues of job creation and destruction and the need for "reskilling". This human-centred agenda points out the broader role of technology in advancing decent work.

On the one hand, technology can free workers from arduous labor, from dirt, drudgery, danger and deprivation; and collaborative robots, or cobots, can reduce work-related stress and potential injuries. On the other hand, technology-driven processes can also render labour superfluous, ultimately alienating workers and stunting their development. Automation can reduce worker control and autonomy, as well as the richness of work content, resulting in a potential deskilling and decline in worker satisfaction.

---

266 Brynjolfsson, E.; Mcafee, A.: *The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies*, W. W. Norton & Company, New York, 2014. Taxing robots seems easier to include in an electoral programme, than an increase of investment in R+D+i and education. Segura Alastrué, M.: 'Los robots…', cit., p. 183.
267 This could lead to double taxation of capital and should not be admitted remaining unmoved. Alastrué, M.: 'Los robots…', cit., pp. 174, 176 and 184.
268 Rodríguez Márquez, J.: 'La justicia tributaria. ¿Cómo puede el sistema fiscal contribuir a disminuir la desigualdad?', *VI Encuentro de Derecho Financiero y Tributario* (on 'Tendencias y retos del Derecho Financiero y Tributario'), Instituto de Estudios Fiscales, Madrid, 27 February 2018.
269 Available at https://www.ilo.org/wcmsp5/groups/public/---dgreports/---cabinet/documents/publication/wcms_662410.pdf (last access 22 February 2019)

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

The ILO recommends that governments and employers' and workers' organizations monitor the impact of new technology on work, steer its development in a manner that respects the dignity of workers and consider the adoption of new regulations in this light. The final decisions affecting work should be taken by human beings, not algorithms. The exercise of algorithmic management, surveillance and control, through sensors, wearables and other forms of monitoring, needs to be regulated.

The European Parliament, in its latest Resolution of 12th of February 2019, on a "Comprehensive European industrial policy on artificial intelligence and robotics"[270], highlights the characteristics of the workforce in the era of artificial intelligence and robotics. It stresses that "as in previous technological revolutions, some jobs will be replaced but new jobs will also be created transforming lives and work practices". It recognizes that the "increased use of robotics and AI should also reduce human exposure to harmful and hazardous conditions and should also help to create more quality and decent jobs and improve productivity".

The European Parliament encourages the development of apprenticeships and vocational training priorities to help workers adapt to technological changes. Accordingly, it recommends that "**Member States, alongside private sector actors, identify the risks and develop strategies to ensure that relevant retraining and reskilling programmes are developed; underlines that companies themselves must invest in the training and reskilling of their existing workforce in order to meet their needs**". It clearly states that "advances in implementing AI in industry should be made with broad consultation of social partners, as the potential shift in the number of people working in the industry requires proactive policies to help workers adapt to the new demands and to ensure that the gains are broadly shared; notes that **this requires re-thinking and re-designing labour market policies, social security schemes and taxation**".

The Commission, for instance, in the transition to road freight transport without a driver, will talk with all interested parties, including the social partners, and possibly consider regulatory initiatives. It takes into account the proposal of the International Transport Forum of the Organisation for Economic Co-operation and Development, on the establishment of a system of **temporary permits, with a fee** to be paid by hauliers to obtain authorisation, which would make it possible to **manage the speed of automation and obtain funds for the retraining of workers and the provision of assistance to redundant workers**[271]. Given that automated and driverless vehicles are still in an experimental phase, it is recognised by the European Commission that the long-term effects of driverless mobility on the transport system, the economy, the environment and existing jobs remain largely unknown - in some sectors the effects will be positive; in others not so much; inevitably changing the profile of the jobs offered - but, in any case, significant efforts are already being made to preserve European values (such as accessibility, social inclusion and attention to the needs of vulnerable users)[272].

---

270 Available at http://www.europarl.europa.eu/sides/getDoc.do?pubRef=-//EP//TEXT+TA+P8-TA-2019-0081+0+DOC+XML+V0//EN
271 OECD-International Transport Forum: Safer Roads With Automated Vehicles?, 2018. [Document available at https://www.itf-oecd.org/safer-roads-automated-vehicles-0]
272 The economic impact of automated and connected mobility will bring benefits beyond the automotive sector, but may also be detrimental to certain sectors such as insurance, maintenance and repair. The impact of automated mobility will largely depend on the ability of European industry to keep pace with

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

When making proposals for an inclusive technological revolution in the labour market, with regard to training one should note that there are different situations:

a) The temporal lack of skills may lead to a reinforcement of the investment in training during the transition phase, while the worker cannot work.
b) When re-training is not feasible due to special circumstances of vulnerability, a safety net is necessary.
c) When compatibility between training and working is possible, e.g. collaboration with interactive robots (cobots) the human adaptation allows recycling or setting new contracts.

Whilst the two first situations are costly for employers and the public sector, the third one is the optimal, because the **human keeps his or her workplace, learns by doing continuously, generates income instead of demanding it**, avoiding botsourcing.

Bearing in mind the goal of adopting an inclusive approach and the embedded costs in necessary adaptations, several options for allocating them are discussed in the following section.

If a company invests in robotics, it may be placed in a best position to compete and subsequently, it may (or eventually not) reap **economic benefits**. These benefits **can be reinvested in the company and somehow flow to the worker**: either as a private **minimum basic income** received due to being displaced (in extreme cases), **or preferably as a package for education** in technological skills, while still keeping the workplace. Of course, in the personal income tax, the latter option should not be declared as a payment in kind by the worker, as ultimately it promotes a better performance for the employer.

If a company decides not to follow the previously referred behaviour (as a sort of direct procedure), then through the **payment of the corporate income tax, the State at some point would be in a position to use the tax revenue collected**, again, (indirectly) for both purposes: either by providing a minimum basic income, or devoting funds for education to improve technological skills. This scheme may be actuated through the company, **creating fiscal incentives** required for the program designed to improve technological skills in a tailored manner; or alternatively, allow the worker to enjoy subsidies or tax benefits to ensure his or her own training.

Additionally, there is a similarity with the classical debate on the need to place the workers elsewhere if a collective dismissal action may cause unemployment, or on the obligation for a worker to re-skill – not just the right, instead of opting for a minimum universal income. The life-long learning educational policies could help moving in this direction.

As long as a worker keeps working the solidarity in the system may run. Envisioning a future with massive unemployment and only a few capable of working and producing for them, alongside even fewer owners of the means of production, is not desirable at all, as it results in a highly polarized society and stigmatized individuals. It increases inequality. The solution is to be

---

international competitors (especially the IT sector). Social inclusion aspects and attention to the needs of vulnerable users will also be important in ensuring that there are benefits for society as a whole, including people, such as the elderly or people with disabilities, who may currently be excluded from mobility services. Particular attention will be paid to increasing the accessibility of remote areas and the provision of wider mobility services. EUROPEAN COMMISSION: "On the Road to Automated Mobility: EU Strategy for Future Mobility", COM/2018/283 final, Brussels, 17.5.2018.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

found through the equal opportunities and efforts to reallocate the gains. A flat minimum income, though attractive due to its simplicity, deepens the divide. The proportionality principle should guide the process of attributing financial help, in accordance with realistic data.

The equality principle, understood as non-discrimination, calls for a search of **legal actions in favour of vulnerable groups due to new forms of disability, such as the lack of technological skills**. The concept of "human" quota could be here implemented. Previous similar contexts have required a positive action by the policy makers (e.g. the gender quota and the disabled quota). However, this quota should be subject to conditions, and should relate to the age and proximity to retirement, limiting the costs. This quota should also consider a temporal perspective, in accordance with the foreseeable change in the professional qualification scale towards jobs that add greater value. Moving to a higher level will help to retain talented young generations and, to some extent, limit non-qualified workers' immigration. In addition, from a broader perspective, some changes are expected in the labour market structure (regarding in/dependent workers), therefore a growing share of technological entrepreneurs –already skilled, would ease the pressure on the employers' side to take in charge this quota. Legislators need to quickly address the situations of the new class of self-employed workers and adopt rules preserving equity and dignity.

As the predictions related to the risks of workers' displacement in a company may vary with the passage of time, due to the speed of the technological change and the improvements of robotics endowed with systems of artificial intelligence, a company should consider the possibility of transferring the probable risk of displaced workers in the future to a third party. Here, the **insurance companies could cover the technological obsolescence implications for workers, on top on the maintenance of robots**. The State should support this systemic risk. A mixed solution is required, enhancing public-private partnership in the light of the United Nations Sustainable Development Goal number 17. The experience of many countries with contributions to pension plans and funds by both employees and employers could provide some guidance in this sense. Public Finances should **promote these socially responsible behaviours through the tax expenditure budget, while establishing clear controllable conditions**.

c. Main considerations on tax capacity and robots

If an act were to say that robots must make tax payments, this would simply be an abbreviated formula for designating a person, such as the owner or usufructuary of the robot; or the person who has the right to use the robot[273].

Currently, in the tax field, robots are not and cannot be considered entities as they are without personality. They are always persons (physical or legal), although under a special regime of action and liability. These factual entities "are also, and fundamentally, subjects of private law relations. And if anything can be criticized of tax law, it is not the extension of

[273] A. Falcón Pulido, 'Taxing robots: clarifications on legal and economic capacity, capacity to act and representation?', J.L. Pons (Ed.), Inclusive Robotics for a Better Society. Selected papers from INBOTS Conference 2018, 16-18 October, 2018, Pisa, Italy, Springer (forthcoming 2019).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

legal capacity with respect to common law, which does not occur, but the lack of knowledge of the relationships that prevail in other sectors of the system"[274].

Nor is it even possible to advocate a special tax capacity for robots. It makes no sense to state that someone has special tax capacity (i.e. that may be obliged to pay taxes), when under ordinary law cannot be the holder of the money and goods necessary to pay those taxes, or be liable for them in the event of non-payment. Additionally, the concepts of person and legal capacity are characteristic of the general theory of law and, as such, have a univocal meaning in all branches of law.

Through the existing types of legal person, such as a trading corporation or foundation, it could be possible in practice to separate the activity of the robot and the goods necessary for that activity, including the robot, from their legal owners. But it is clear that the taxpayer would be the corporation or the foundation. And those who would act in the traffic would be their organs: a general meeting of shareholders and administrators or board of directors in a corporation; board of trustees in a foundation. Although, at least hypothetically, these organs could choose to follow the instructions given by the robot in some areas where artificial intelligence might be applied.

The question that immediately arises is whether robots have economic capacity or ability to pay. The answer is, in our opinion, negative. The businessman who uses robots will obtain more income thanks to the "work" of those robots, and their greater productivity, but that economic capacity is already taxed through the personal income tax or corporation tax[275].

If robots were attributed legal personality and they could accumulate a wealth in their own name, they would probably have to be taxed in order to avoid deferral or evasion. But as long as the general legal capacity of the robots is not recognised there are no fair reasons to justify the levy.

Those who defend the creation of a tax on robots do not do so by affirming the existence of (a special electronic) ability to pay[276], but as an extrafiscal measure, as a way of maintaining the collection despite the loss of jobs, which would even allow for the creation of a universal minimum income. This approach is debatable because in countries with more density of robots, such as Germany or Japan, the unemployment is relatively lower. In addition, there would be enormous problems in designing such a tax on the income of robots, given the difficulty of measuring the work that has been replaced by the robot[277].

---

[274] These entities without personality are usually a case of plurality of people who have common assets, on which the tax obligation must be effective. In other cases, there is a temporarily indeterminate owner of certain assets that make up a separate estate (e.g. the unsettled inheritance, until it is accepted). R. Falcón y Tella, Derecho financiero y tributario (parte general), Servicio de Publicaciones de la Facultad de Derecho, UCM, 7th ed,, Madrid, 2017, pp. 223-224.
[275] C. García Novoa, La tributación de los robots y el futurismo fiscal, Taxlandia, 17 April 2018.
[276] X. Oberson, Taxing Robots? From the Emergence of an Electronic Ability to Pay to a Tax on Robots or the use of Robots, World Tax Journal, May 2017. X. Oberson, Taxing Robots.
Helping the Economy to Adapt to the Use of Artificial Intelligence, Edward Elgar, Cheltenham, 2019.
[277] C. García Novoa, "Impuestos atípicos en la era post BEPS", in A. Cubero Truyo (Dir.): Tributos asistemáticos del ordenamiento vigente, Tirant lo Blanch, Valencia, 2018, pp. 223-224.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

d. The recent experience with the Californian tax on autonomous vehicles

The Transportation Assistance Funding Act -passed in California in 2018, allows the City and County of San Francisco to impose a tax on each ride originating there, when provided by an autonomous vehicle, among others. Several controversial issues - raised by different stakeholders along the legislative procedure, are here discussed[278].

On 21 September 2018, chapter 644 was published in the Statutes of 2018, which is the result of Assembly Bill No.1184 (hereinafter AB 1184, Ting), entitled "City and Country of San Francisco: local tax: transportation network companies: autonomous vehicles", which adds section 5446 to the Public Utilities Code, related to transportation[279]. This rule authorizes the City and County of San Francisco, subject to the voting requirements for approval, to levy a tax on each trip originating in these places, either provided by an autonomous vehicle or by a participating driver.

From the systematic perspective of the tax system, this approach seems quite reasonable, since such these robots are not taxed for the mere fact of being so[280]. A lesser contribution to public needs from those operating digitally in the economy cannot be explained by undermining, without any proper justification, the requirements of the principle of equality in taxation. However, **the implementation of different taxes in an isolated manner could be problematic in the long-term** - especially if it tends to vary locally[281].

Obviously, a preliminary step in the transition to automation in the transport sector is to legally qualify vehicles without drivers as vehicles allowed in urban traffic. **To the extent that they are progressively assimilated to common vehicles, progress is made in subjecting them to the same legal regime**. This is usually the case in the areas of insurance and liability[282], but this also has consequences in the area of taxation.

This tax is levied on each trip with origin in the City and County of San Francisco. It is therefore irrelevant where the journey ends or where the driver resides. It can be made either by an autonomous vehicle through a transportation company or any other person, or by a participating driver. During the legislative procedure, there have been discussions about whether or not to include the use of autonomous vehicles. Basically, two reasons have been put forward against incorporating them: the brake on innovation that this may entail and, the

[278] M. A. Grau Ruiz: "La búsqueda de alternativas para la tributación de los robots: la tasa californiana aplicable a los vehículos autónomos", C. García Novoa (Dir.), Cuarta Revolución Industrial, Universidad de Santiago de Compostela, Thomson Reuters Aranzadi, Cizur Menor, 2019. M. A. Grau Ruiz: "Taxing autonomous vehicles: the Californian case", J.L. Pons (Ed.), Inclusive Robotics for a Better Society. Selected papers from INBOTS Conference 2018, 16-18 October, 2018, Pisa, Italy, Springer (forthcoming 2019).
[279] http://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180AB1184 [last access 15 Nov 2018]
[280] C. García Novoa, "Impuestos atípicos en la era post BEPS" (Capítulo 8), in A. Cubero Truyo (Dir.), *Tributos asistemáticos del ordenamiento vigente*, Tirant lo Blanch, Valencia, (2018): 223-224.
[281] SB-1184, Pan "Vehicles: City of Sacramento shared autonomous vehicle pilot Project. An act to add and repeal Section 38757 of the Vehicle Code, relating to vehicles" [available at: http://leginfo.legislature.ca.gov/faces/billTextClient.xhtml?bill_id=201720180SB1184]
[282] PALMERINI, E.: "Robótica y derecho: sugerencias, confluencias, evoluciones en el marco de una investigación europea", *Revista de Derecho Privado, Universidad Externado de Colombia*, n.º 32, 2017, pp. 53-97. DOI: https://doi.org/10.18601/01234366.n32.03

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

fact that by not excluding trips of own vehicles for private purposes their acquisition may be discouraged. They were rejected.

The amount to be recovered shall not exceed 3.25% of the applicable net rider fares for a journey, or 1.5% of the same fares for a carpool. These fares are defined as all charges for a journey, which include - but are not limited to - those based on criteria of time of use and distance in miles travelled, or both, and which exclude other additional charges such as taxes, airport or other taxes, or charges imposed by the Commission. The City and County of San Francisco may adopt a lower tax rate if the transportation service is provided with a zero-emission vehicle. In this particular point, the representatives of the automotive industry strongly criticized that full exemption should have been granted.

The sums collected shall be dedicated to finance transport operations and infrastructure within the City and County of San Francisco. A time limit is set for the duration of this tax, which may not exceed 5 November 2045.

In general, there is a current trend towards **tax hypothecation (or earmarked taxes)** in many tax systems. In particular, in relation with robotics, some proposals address this issue for the sustainability of public expenditures (e.g. Social Security, or Universal Basic Income)[283]. However, tax hypothecation has some well-known "cons": the vulnerability of resources, the fiscal populism, and the difficulty of reversing it. On the contrary, some of its "pros" are the following ones: transparency, accountability, amplification and public commitment in tax compliance. Perhaps this mechanism is **suitable for specific public interventions that just serve punctually to initiate a change in a given issue**[284].

4. Conclusions

The legal problems spill over into various disciplines or areas of knowledge, which come together to seek solutions to the effect of robots in the labour market. When robotization in the labour market means the replacement of human jobs, tasks and activities by robots, ethics should be used as a starting point to reflect on the principles, scope and methods of regulatory intervention needed to safeguard human rights, i.e. those inalienable rights which derive from human dignity.

Nevertheless, **regulation**, in its most continental sense, i.e. hard law, cannot be designed from a perspective as a defence or barrier against robotic technological innovation, but **should rather be geared towards formulating measures which provide certainty to all of those involved on the basis of the principle of socially and legally responsible robotic innovation**. To this end, based on the fundamental rights enshrined in the EC, and now in the European Pillar of Social Rights, it is necessary for everybody, public authorities and

---

[283] GARCÍA NOVOA, C.: "Impuestos atípicos…", *op.cit.*, pp. 221-222. GRAU RUIZ, M. A.; SÁNCHEZ-URÁN AZAÑA, M.Y.: "El impacto de la robótica, en especial la robótica inclusiva, en el trabajo: aspectos jurídicos-laborales y fiscales", Eprint UCM, 2018 [available at https://eprints.ucm.es/47523/] "Robotics and Work: Labor and Tax regulatory Framework", en *International Congress Technological Innovation and Future of Work*, 5-6 April 2018, Santiago de Compostela. Eprint UCM, 2018 [available at https://eprints.ucm.es/47718/].

[284] THE RESPONSIBLE TAX LAB: *Follow the money: Is the time right for (more) tax hypothecation*?, Common Vision - CoVi UK, September 2018.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

social partners, to promote legislation which balances the guarantee of entrepreneurial freedom with that of workers' rights.

From this dual perspective, **legislation should go further in guaranteeing the employability of humans and balance the incentives for launching highly technologized activities and companies against the loss of workers in sectors which are reconverting from their traditional model towards more technologically based activities**. In this respect, it will be necessary to offer, at least in the transition phase, financial incentives for those companies that truly and effectively invest in the technological training of their workers.

Furthermore, with regard to the labour relationship, labour Law needs to consider the possibility that an ('intelligent') robot can be considered a worker. It may be endowed with a certain employment status, but it cannot be legally considered as a salaried or self-employed worker. Labour Law is based on human work and a robot, even one with a high dose of AI, cannot be considered as such.

If we use this fundamental fact as a starting point, robotics has many implications for the direct or indirect regulation of working conditions, for rights and obligations, for both the entrepreneur and the worker, and from both a labour and fiscal perspective.

The starting point for the analysis is the principle of equality and non-discrimination, i.e. the real and effective equality of people, of the groups of people who are vulnerable to automation or robotization. Thus, these cannot become a direct or indirect cause of discrimination; technological neutrality cannot entail a disadvantage or barriers for certain groups of workers. In this respect, and also in the transition phase, as an exceptional, limited and conditional measure, it is necessary to reflect on the quota of humans in companies and reasonable adjustment measures for those groups which are especially vulnerable, including older workers.

On the basis of this first reflection, we need to adapt labour legislation as it applies to workers' rights and obligations. Both from the perspective of guaranteeing people's employability more than jobs themselves (which will affect the direction taken by legislation relating to the replacement effect, i.e. the replacement of humans by robots) and of guaranteeing the rights of workers in their interaction with robots in the workplace (cooperation between them).

All the expected public policies will probably have an impact on the Financial and Tax Law; but, at the same time, this discipline will condition their actual applicability in practice. **The States will have to to allocate rights and responsibilities among human beings for the actions of non-human beings, fighting inter-personal and international inequality, through strengthened cooperation.** The high costs of a universal basic income, now at an experimental phase, could be publicly funded (through tax measures or cuts in spending), and/or even privately (if there could be a particular interest, i.e. in eliminating candidates for other competitor?). In parallel, the idea of a more limited conditional basic income emerges for practical budgetary reasons. However, it will entail the need of **proportionate justification, in order not to breach constitutional or EU principles**. The financial and tax regulation should always look for a balance between technological and social progress.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

6. Assessing human acceptance and ethical concerns of interactive robotics

1. Introduction (Conceptual Analysis) and background

Artificial intelligence (AI) and robotics are technologies that seem to be of major importance for the development of humanity in the near future. They have raised fundamental questions about what we should do with these systems, what the systems themselves should do, and what risks they have in the long term. They also challenge the human view of humanity as the intelligent and dominant species on Earth.

Every robot combines mechanical, electrical and software components to develop a desired function. Therefore, in every robot there is a cogeneration of mechanical, electrical and programming (software). Such cogeneration between these three great parts has been synchronized for most of the golden age of robotics as science and in industry. This golden age ranges from 1961 when the first industrial prototype was developed in New Jersey to 1974 when Björn Weichbrodt of ABB created the first commercial microprocessor for controlling robots that laid the foundation for the modern robotics industry.

But the synchronization and/or cogeneration between the mechanical, electrical and software parts of any robot have begun to desynchronize. Right now, thanks to the great development of the computational capacity of processing chips, new algorithms and a large amount of data (Big Data), the software part has obtained an advantage by creating a heterochrony between the mechanical, electrical and software components. Machine learning - a field of Artificial Intelligence that allows extracting patterns or hidden correlations in data - which is the programmable part or software to control robots, has led to a *softwareization of robotics*. This paradigm shift or trend towards the softwarization of robotics leads to the emergence of robotics detached from mechanical and tangible embodiment and manifesting as a naked AI or software. A robot no longer has to have physical presence but can be a digital creature. The taxonomy of the machine kingdom of Hernández-Orallo (2017) highlights this **transition from the mechanistic paradigm of robotics to a digital paradigm of robotics**.

The notion of 'artificial intelligence' (AI) is understood broadly here, as any kind of artificial computational system that shows intelligent behaviour, i.e. complex behaviour that is conducive to reaching goals. In particular, we do not wish to restrict 'intelligence' to what would require intelligence if done by *humans*, as Minsky said (1985) and is often repeated[285]. This means we can incorporate machines from 'technical AI' that show only limited abilities in learning or reasoning but excel at the automation of particular tasks, as well as 'general AI' that aims at creating a generally intelligent agent. However, AI is somehow closer to our skins than other technologies since it is the project of AI to create machines that have a feature central to how we humans see ourselves, namely as feeling, thinking, intelligent beings – thus the field of

[285] Taddeo, Mariarosaria, and Luciano Floridi. "How AI Can Be a Force for Good". Science, 361 (2018): 751, http://science.sciencemag.org/content/sci/361/6404/751.full.pdf

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

'philosophy of AI' [286] . While the main purposes of an intelligent agent probably involve sensing, modelling, planning and action, current AI applications include perception and sensing, text analysis, language translation and generation, logical reasoning, game-playing, decision support systems, data analytics, predictive analytics, as well as autonomous vehicles and other forms of robotics[287]. AI may involve any number of computational techniques to achieve these aims, be that classical symbol-manipulating AI, be it inspired by natural cognition, or machine learning via neural networks – the area that currently looks most dynamic[288]. Historically, it is remarkable that the term "AI" used to be very broad in ca. 1950-1980, then it came into disrepute during the 'AI winter' ca. 1980-2000 and narrowed ('machine learning', 'natural language processing' and any 'data science' were often not labelled as 'AI'). It is now since ca. 2010 that the use broadened again, and at times almost all of computer science and cognitive science is lumped under 'AI' – now a name to be proud of, on the edge of hype again, and a booming industry with massive capital investment[289] that "… promises to drive growth of the [...] economy, enhance our economic and national security, and improve our quality of life"[290].

While AI can be entirely software and data processing, robots are physical machines with actuators that move and interact with the environment, that exert physical force on the world, such as a gripper or a turning wheel. From this point of view, autonomous cars or planes are robots (to the degree that they are autonomous), and only a minuscule portion of robots are human-shaped or 'humanoids', like in the movies. Some robots use AI, and some do not – e.g. typical industrial robots (of which there are millions) blindly follow completely defined scripts with minimal sensory input and no learning or reasoning. It is probably fair to say that while robotics systems cause the more concerns in the general public, AI systems are actually more likely to have a greater impact on humanity.

The fields of robotics and AI can thus be seen as two overlapping circles of systems: Systems that are only AI, systems that are only robotics, and systems that are both. It is interesting not just the intersection, but the union of both circles taken together.

The main division is into issues that arise with AI systems as *objects*, i.e. tools used by humans vs. AI systems as autonomous *subjects*, i.e. when ethics is for the AI systems themselves. The

---

286 Gomila, Antoni, and Vincent C. Müller. "Challenges for Artificial Cognitive Systems". Journal of Cognitive Science, 13 (2012): 453-469. http://doi.org/10.17791/jcs.2012.13.4.453; Müller, Vincent C. (ed.), forthcoming, Oxford Handbook of the Philosophy of Artificial Intelligence. New York: Oxford University Press.
287 Stone, Peter, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, . . . Astro Teller, 2016. "Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel". https://ai100.stanford.edu/2016-report
288 Goodfellow, Ian, Yoshua Bengio, and Aaron Courville, 2016, Deep Learning. Cambridge, Mass.: MIT Press; Silver, David, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, Demis Hassabis. "A General Reinforcement Learning Algorithm That Masters Chess, Shogi, and Go through Self-Play". Science, 362 (2018): 1140-1144, http://science.sciencemag.org/content/sci/362/6419/1140.full.pdf
289 Shoham, Yoav, Perrault Raymond, Brynjolfsson Erik, Jack Clark, James Manyika, Juan Carlos Niebles, . . . Zoe Bauer, 2018 (December 2018). "The AI Index 2018 Annual Report". AI Index Steering Committee, Human-Centered AI Initiative. http://cdn.aiindex.org/2018/AI Index 2018 Annual Report.pdf
290 Trump, Donald J, 2019, "Executive Order on Maintaining American Leadership in Artificial Intelligence". The White House (11.02.2019). https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

problem of a future 'singularity' or 'superintelligence' concerns ethical use of AI and the ethics for AI[291].

New technologies are often met with 'concerns' of various sorts, many of which will turn out to be rather quaint (such as that trains are too fast for souls), some predictably wrong when they suggest that humans will change fundamentally (telephones will destroy personal communication, writing will destroy memory, or video cassettes will make going out redundant - https://xkcd.com/1289/), some predictably correct but moderately relevant (digital technology will destroy industries that make photographic film, cassette tapes, or LP records), but some broadly correct and relevant (such as that cars will kill children, change the landscape, and challenge good sleep). It is important to analyse the issues, and to discard out the non-issues - always keeping in mind that technologies are situated in a social and historical context. Some technologies, like nuclear power, cars or plastics, have caused ethical and political discussion and significant policy efforts to control the trajectory that these technologies are driving – usually once the damage is done. All these concerns also exist in the case of the new technology of "artificial intelligence" or AI, plus the more fundamental fear that it may end the era of human control on planet Earth. In addition to such 'ethical concerns', new technologies challenge current norms, conceptual systems and societal structures, which is of particular interest to philosophy. **Once we have understood a technology in its context, we need to shape our societal response, including regulation and law**.

Our philosophical task here is to present an orientation and analysis of the *issues*, the *positions* and the *arguments* in the field – with an outlook on *policy*.

The ethics of AI and robotics have seen significant press coverage in recent years, which supports this kind of work, but also may end up undermining it. We often talk as if we already knew what would be ethical, and the issues are just what will happen, and what we should do about it; as if the issues were only considerations of risk and the prediction of impact (e.g. on the job market). But if we really do already know what is ethical we may have 'ethical problems' but not 'problems of ethics' of AI and robotics – a problem of ethics would require that we do not know what is the right thing to do, perhaps because we do not know which are the factors that matter or because there is a tension between different values or consequences. We do thus not follow much of the current discussion in policy and industry with its focus on image and public relations – where "ethical" is really not much more than the new "green".

A caveat is in order for our presentation: The ethics of AI and robotics is a very young field within applied ethics, with significant dynamics but few well-established issues and no authoritative overviews, though there is a promising outline[292], and there are beginnings on

---

[291] See https://www.researchgate.net/project/Ethics-of-AI-and-Robotics-for-Stanford-Encyclopedia-of-Philosophy There the sections and sub-sections are ordered by ethical issues, rather than by technologies. After a general explanation of the *issue*, *positions* and *arguments*, we look how this plays out with current *technologies* and finally what *policy* consequences may be drawn. This means we have occasion to look at ethical issues that arise from utility of consequences, including 'risk', as well as issues that arise from a conflict with rules, virtues or values. We also discuss questions of a more theoretical nature about 'artificial moral agents', in particular under which conditions an agent should be taken to have rights and responsibilities.

[292] European Group on Ethics in Science and New Technologies, 2018 (09.03.2018). "Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems". European Commission, Directorate-General for Research and Innovation, Unit RTD.01. http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

societal impact [293], sometimes with specific policy recommendations (AI HLEG 2019; IEEE 2019). So here we cannot just collect what the community has achieved thus far but must propose an ordering where very little order exists, including the identification of promising lines for future work – but without spending much time on the historical development of the field.

2. A Note on Policy

There is significant public discussion about AI ethics and frequent pronouncements from politicians that the matter requires new policy – however, this is easier said than done: Actual technology policy is difficult to plan and to enforce. **Technology policy can take many forms, from incentives and funding, infrastructure, taxation, good-will statements, regulation by various actors (self-regulation, local, national, international), to law. It will possibly come into conflict with other aims of technology policy (e.g. sustainable development) or general policy (e.g. economic growth).** One important practical aspect is which agents are involved in the development of a policy and what the power structures are. For people who work in in ethics and policy, there is probably a tendency to overestimate the impact and threats from any new technology, and to underestimate how far current regulation can reach (e.g. for product liability). Governments, parliaments, associations and industry circles in Europe and North-America have produced reports and white papers in recent years (we maintain a list on http://www.pt-ai.org/TG-ELS/policy), some have generated good-will slogans ('trusted/responsible/humane/human-centred/good/beneficial AI'), but, as of early 2019, very little actual policy has been produced – beyond funding for AI, and for AI policy.

There are beginnings: The latest EU policy document suggests trustworthy AI should be lawful, ethical and technically robust … and then spells this out as seven requirements: human oversight, technical robustness, privacy and data governance, transparency, fairness, well-being and accountability (AI HLEG 2019). Much European research now runs under the slogan 'responsible research and innovation' (RRI) and 'technology assessment' is a standard field since the advent of nuclear power. Professional ethics is now a standard field in Information Technology as well, and this includes issues relevant here (e.g. confidentiality). In AI, there is a risk that the current vision statements and self-regulation in the industry tend to delegate the decisions to experts, "a narrow circle of who can or should adjudicate ethical concerns around AI/ML" [294] , rather than incorporating it into the societal agents more deeply. A useful summary

[293] Floridi, L.et al., "Ai4people—an Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations". Minds and Machines, 28 (2018): 689; Taddeo M. et al., "How AI Can Be a Force for Good". Science, 361 (2018): 751; Taylor et al., June 2018). "Responsible AI – Key Themes, Concerns & Recommendations for European Research and Innovation: Summary of Consultation with Multidisciplinary Experts"; Walsh, Toby, 2018, Machines That Think: The Future of Artificial Intelligence. Amherst, Mass.: Prometheus Books; Bryson, Joanna J, 2019, "The Past Decade and Future of Ai's Impact on Society". In Anonymous (ed.), Towards a New Enlightenment: A Transcendent Decade. Madrid: Turner – BVVA; Whittlestone, J. et al., "Ethical and Societal Implications of Algorithms, Data, and Artificial Intelligence: A Roadmap for Research". (2019): 1-59.

[294] Greene, D. et al.,"Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning". 52nd Hawaii International Conference on Systems Science

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

of an ethical framework for AI is given in [295]. On general AI policy, see[296] as well as [297]. On wider societal implications[298]. We will discuss the theoretical approaches and the policy for each type of issue we identify, rather than for AI or robotics in general.

3. Cartography of the Values Involved in Robotics

The title of our project "Inclusive Robots for a Better Society" opens up a number of interesting ethical questions. What is meant by "inclusive" when we refer to inclusive robotics? What is meant by a "better society"? How will the demands of Responsible Research and Innovation (RRI) be met? Accordingly, we focus on conceptual analysis and clarification of these key concepts and values. For instance, we will explore which values will underpin RRI. We will examine the question of inclusive robotics in relation to society as a whole and in relation to document such as the Convention on the Rights of Persons with Disabilities (CRPD).

A particular point of ethical interest is the notion of "inclusion" and how "inclusive societies" set the benchmark for "inclusive robotics." Therefore, the title "Inclusive Robotics Robotics for a better Society" can be turned into "Better Robotics for Inclusive Societies." **The crucial question is "Do we have the robots that we need, and do we need the robots that we have?"**

a. A pluralist axiology for technoscientific practice and robotics

The idea of "value" and its ontology, that is, what it is and what its defining characteristics are, has been a question of interest for philosophers throughout the history of ideas. In philosophy, the expressions "axiology" or "value theory" are used with different meanings to refer to areas of scientific practice and knowledge that have an evaluative component. In addition to the traditional position, which concentrates on ethics and aesthetics, other philosophical currents, such as feminism and political philosophy, maintain a certain axiological perspective or contemplate a value theory at the base of their foundations. Axiology or value theory is dedicated to reflecting on what is valuable and the reasons for it being deemed so.

Traditional questions are the typology of values and how they can be analysed and decomposed into categories: intrinsic, extrinsic, instrumental values, etc., how they can be classified, if they are from a single source or if many values exist (monism or pluralism), etc.

In the field of *inclusive robotics*, we take a pluralistic position, with the goal of presenting a way for us to approach the complex system of values around which current scientific-technological practice takes place.

---

[295] European Group on Ethics in Science and New Technologies, 2018 (09.03.2018). "Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems". European Commission, Directorate-General for Research and Innovation, Unit RTD.01.
[296] Calo, Ryan. "Artificial Intelligence Policy: A Primer and Roadmap ". University of Bologna Law Review, 3 (2018): 180-218.
[297] Stahl, Bernd Carsten, Job Timmermans, and Brent Daniel Mittelstadt, 2016, "The Ethics of Computing: A Survey of the Computing-Oriented Literature". ACM Computing Surveys, 48/4 (55), 1-38.; Johnson, Deborah G, and Mario Verdicchio. "Reframing AI Discourse". Minds and Machines, 27 (2017): 575–590; Giubilini, Alberto, and Julian Savulescu, 2018, "The Artificial Moral Advisor: the "Ideal Observer" Meets Artificial Intelligence". Philosophy & Technology, 31 (2018): 169-188; Crawford, Kate, and Ryan Calo. "There Is a Blind Spot in AI Research". Nature, 538 (2016): 311-313.
[298] acobs, An, Lynn Tytgat, Michel Maus, Romain Meeusen, and Bram Vanderborght (eds.), 2019, Homo Roboticus: 30 Questions and Answers on Man, Technology, Science & Art. Brussels: ASP

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

The axiology of Javier Echeverría (1998) serves as a theoretical frame of reference to work with this complex system. He defends that scientific and technological practice are loaded with values, given that human actions themselves are guided by value systems. Minimally these values would be epistemic or cognitive (clarity, consistency, originality, verifiability, truth…) and technical (utility, innovation, trustworthiness, efficiency, accuracy…), although as we will see, other values, mainly economic, also have an influence.

Continuing with Echeverría (2002)[299], an axiological term does not have meaning on its own. On the contrary, it only acquires meaning when, inserted into a set of values, an agent applies the axiological term in question to a system or thing, creating a valuation. In this way we can go from logic based on subjects and predicates (for example, saying that an agent is innovative) to logic based on arguments and functions (certain actions of an agent have been innovative). This allows specific cases of innovative actions to be analysed as well as different degrees of innovation to be compared. As such, graduality of values is acknowledged. In other words, a value, for example, *utility*, can be satisfying to varying degrees.

Based on these considerations, a set of ideas can be useful as a bridge to move from the axiology of the technology towards the field of *inclusive robotics*. **What interests us is the pluralism of values involved in *technoscientific* practice, as well as the conflicts over values between different agents. Which elements from this axiological approach are especially relevant for interactive robotics?**

An analysis of publications from different interest groups can be useful when it comes to drawing a map of values. Each company in the robotics industry has a series of values that guide its mission and business activity. We have taken as a source various documents on values in the robotics industry, among them the Executive Summary World Robotics 2017 Industrial Robots. In this case, we have identified: innovation, security, profit maximisation, excellence, respect. The values we took from the users as stakeholders were obtained from a qualitative review of various publications from ASPAYM (a Spanish association for people with spinal cord injury). For them, the values that stand out are quality, convenience, usefulness, functionality and price. The values from society are, on the whole, generic but are present in numerous reports on Corporate Social Responsibility. In this case, we have identified community, responsibility, trust, integrity, inclusivity and diversity.

b. Ethical, axiological and sociotechnical frameworks for inclusive robotics

Robotic technologies are blurring and overlapping the borders between human subjects and technological objects. This blurring of lines has ethical implications affecting our axiological categories. In this sense, values traditionally reserved for human beings (such as autonomy, responsibility, creativity) are beginning to be projected onto intelligent technological entities, often driven by our tendency towards anthropomorphism when it comes to contemplating and evaluating robotics.

As robotic technology becomes more autonomous[300], it is necessary to identify the values and ethical principles that should regulate the interaction of robotic systems with human beings.

[299] J. Echeverría, Ciencia y valores. Destino Barcelona, 2002.
[300] Michael Funk, and Mark Coeckelbergh, "(Technical) Autonomy as Concept in Robot Ethics", J.L. Pons (Ed.), Inclusive Robotics for a Better Society. Selected papers from INBOTS Conference 2018, 16-18 October, 2018, Pisa, Italy, Springer (forthcoming 2019). They clearly explain "the empirical-descriptive application of the word autonomy in technical context is different to the normative usage of autonomy in

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

The fundamental challenge for inclusive robotics is the creation of a global strategy for interactive technologies with empirical validations and recognition of the global plurality of values. These ethical principles and values should include: a) human dignity, b) autonomy, c) privacy, d) principle of nonmaleficence, e) principle of responsibility, f) principle of beneficence, and g) justice.

Advances and developments in robotics and Artificial Intelligence (AI) are making us aware of the numerous ethical implications of these technologies on different groups of people. In the case of inclusive robotics, children will not be affected in the same ways as adults, and even gender bias must be considered for inclusive robotics having to do with care that take into account the needs and demands of individuals with disparate personal characteristics.

We believe that to identify the values present in the design of robotics, it is necessary not only to take into account all interest groups, but also that preferences and axiological hierarchies and how they come into play in different people should be tested empirically.

Nevertheless, normatively, what can the pluralist axiological model contribute to an inclusive robotics? In the first place, to defend the consideration of *inclusion* as a central value that requires its degree of satisfaction to be equal or greater than a determined level or minimum benchmark[301]. Second, to locate and analyse those evaluation mechanisms that take place throughout the entire innovation process of interactive robotics. Third, to promote *inclusion* as a core value within these mechanisms, and to encourage a greater degree of satisfaction of other values closely related to it.

We consider that this approach should be complemented with practical analysis of the interactive robots that come out on the market. Although the theoretical approach to technological innovations is an extremely important question, auditing the political properties of these robots in their interaction with human beings should be a significant objective. Winner argued convincingly that artifacts can have political characteristics in two distinct ways[302]. First, when their design, invention and implementation become a means to reach an end. Secondly, Winner points out that there are certain technologies whose own nature is very specifically politically charged. As such, adoption of a certain technological system would imply a series of

---

human life and human societies. Following this insight, and an embodied approach in philosophy of technology, six forms of technical tools are briefly introduced which could be used to describe several levels of technical autonomy. The different forms are summarized in a heuristic scheme". The more aspects/features of human bodily actions are included into a tool, the more "autonomous" it becomes.

forms of technical tools (1-6)		energy	movement / process		control / framework	
			poiesis (routine)	praxis (problem solving)	aim / advise / end	surveillance / intervention
"pre-modern" & "modern" technologies	**1 handcraft**	human body / actor (e.g. hand axes)				
	2 machine	tool	human body / actor			
	3 automat		tool	human body / actor		
"hypermodern" technologies: computer & robotics	**4 embedded technical autonomy**			tool *social robots*	human body / actor	
	5 technical semi-autonomy				tool *social robots*	human body / actor
postulate	**6 autonomy**					"tool" (?)

301 .

302 L. Winner, "Do Artifacts Have Politics?" Daedalus, 109 (1980): 121–136

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

determinants regarding human relations, favouring political and moral values such as centralisation or decentralisation, equality or inequality, repression or liberation. In other words, a technology should not be evaluated only on how it contributes to efficiency and productivity, but also for the way in which it can create certain forms of power and authority. As such, these political characteristics can work in favour or against *inclusivity*.

This perspective allows us to **explore another road towards an inclusive robotics**. Quintanilla (2018)[303] coined the term "**engaging technologies**," as opposed to other alienating technologies, proposing a series of evaluative criteria for development based on these technologies. *Openness*: a robot is open if its software is free. *Versatility*: a robot is inclusive if it allows alternative uses by its operators or users. *Docility*: a robot is docile to the extent that its functioning, control and shutdown are easily achievable for human operator. *Obsolescence*: a robot is inclusive if its repair, updating and maintenance are promoted above its substitution. *Comprehensibility*: a robot is understandable if it has some instructions and a comprehensible design for its operators or users.

c. Inclusive and participatory practices to manage value conflicts in interactive robotics

Technologies do not only transform objects, but also transform habits, customs or relationships[304]. Consequently, if we understand robotics systemically, we can say that it will transform systems, whether these are social, economic or natural. This transformative vocation creates conflicts of values in multiple phases of technological development. For example, when it comes to evaluating which technologies get financed, controversies arise over different priorities. Later, given that innovations compete in the market, this becomes a very important evaluation mechanism. When innovations reach society, controversies are often posed in terms of technophilia and technophobia. On these value conflicts underlie moral, social, religious, aesthetic and ecological values, to name a few. In this sense, **the context where robotics are applied determines, delimits or complicates the axiological pluralism** we make reference to (for example, creating controversies with other subsystems of values in the fields of care, health or the military).

Even so, given that interactive robotics includes all robots that maintain physical, cognitive and emotional interaction with a person, some controversies inherent in the technology itself can be identified. For example, *utility* or *efficiency* in these robots' performance requires deep understanding and interpretation of the movements, thoughts and emotions of the people they interact with. Economic and technical values may conflict with social and moral values such as *autonomy* and *privacy*.

Technologies are the result of a plurality of agents who take decisions to make them how they are. These decisions involve a diversity of values, among which the values that are technical, political and economic often prevail. With a view to taking into consideration the very diverse characteristics and expectations of different groups of people, the designing of products and services should be as participatory as possible. **User participation in technological development processes contributes to a better representation of the social values**

[303] M. Á. Quintanilla Fisac, "Engaging Technologies: Criteria for an Alternative Model of Technological Development," in Spanish Philosophy of Technology: Contemporary Work from the Spanish Speaking Community, B. Laspra and J. A. López Cerezo, Eds. Cham: Springer International Publishing, (2018): 103–123.

[304] J. Echeverría, La revolución tecnocientífica. Fondo de Cultura Económica de España, 2003.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

involved. Participating in these processes is not limited to interceding in technical decisions, however, for development of the technology is not conditioned on those decisions alone. All decisions involve values-related dilemmas that call for citizen participation—users, consumers, those responsible for the technological products, those affected by them, and so forth[305].

People and groups with very different and perhaps contradictory perspectives may have an influence on all facets of the development of these products as well as the situations in which they are used. The users themselves know best how technological devices can contribute to their greater quality of life and level of participation in society and can assist in evaluating their advantages as well as their drawbacks in the earliest phases of their development [306]. As the ones most familiar with their own reality, users help to create technological solutions that are more feasible in not only their technical but also their financial and social aspects. User participation also contributes to generating demand for those solutions which, in turn, stimulates their introduction into the market and inspires new lines of research. Similarly, owing to the participation of users—the presence of actors who are generally not taken into consideration in technology development processes—citizen groups are better able to generate feasible technological proposals and **steer innovation toward the real needs of society**[307].

4. Finitude and Vulnerability – Existential Societal Fundamentals for Better Robotics and Inclusive Societies

a. Introduction: The Basic Argument and Methodology

This perspective to the ethical debate is about one main argument and a wordplay with the title of the project. It will be presented in the next paragraph, while the next subsection draws on established philosophical concepts in order to methodologically justify the primary claim.

i. Argument

Societal inclusion (for instance of people with handicaps or in elderly care or rehabilitation, etc.) **relates to the general human existential situation of vulnerability and finitude[308] which is true for everyone**, not only for persons with diseases or handicaps. "Better robotics for inclusive societies" needs to take this issue into account. The benchmark is social life in which finitude and vulnerability are not seen as weakness or disability, but as universal human feature which also enables creativity, arts, and joy of life. Anything else could lead to inhumane technocratic ideologies, where "better societies" are defined by "inclusive robotics", thus by techno-economical possibilities, instead of real human and social needs. This section summarizes the conceptual analysis of vulnerability and finitude – the two key terms of the argument. Beside the human body (5.4.2.1) also lifetime (5.4.2.2), knowledge (5.4.2.3) and communication (5.4.2.4) shape the fragile conditions in which we practically apply robots.

[305] M. Toboso, "Rethinking disability in Amartya Sen's approach: ICT and equality of opportunity", Ethics and Information Technology, 13 (2010): 107-118.
[306] Clarkson, J., Coleman, R., Keates, S. and Lebbon, C. (eds.) (2003). Inclusive Design. Design for the whole population. London: Springer.
[307] Tewey, B. (1997). Building Participatory Action Research Partnerships in Disability and Rehabilitation Research. Washington, DC.: U.S. Department of Education, National Institute on Disability and Rehabilitation Research.
[308] Coeckelbergh M (2013) Human Being @ Risk. Enhancement, Technology, and the Evaluation of Vulnerability Transformations. Springer, Dordrecht

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

ii. Methodology

Following foundational approaches in the 20th- and 21st-century philosophy of technology – including phenomenology and hermeneutics[309, 310], Anthropology[311], methodical constructivism and culturalism[312, 313] – both the human body and ordinary social life can be used as methodical starting points of the theoretical conceptualisation of human practice. Vulnerability and finitude are existential characteristics of human bodies in their sociotechnical and cultural embedding[314, 315]. They serve as both the object and criterion of the critical investigation. In order to further sharpen the focus, asymmetry[316, 317] will be used as pointed hermeneutical criterion, which can be applied to the impact of robots within inclusive societies. The emphasized problem includes interactions and relations between humans and robots. In conclusion the methodically sharpened question is: **Which asymmetries in terms of vulnerability and finitude shape human-robot-relations and human-robot-interactions on the basis of human bodily, every day, sociotechnical and culturally embedded practice?** Since asymmetry is a concept which involves a wider range of aspects with a high ethical and epistemic relevance – including for instance political power-related asymmetries or asymmetric warfare[318] – the focus of this analysis remains on bodily practice in its ethical and epistemic significance. Therefore, a conceptual analysis of the meaning of finitude and vulnerability is set up, including the following four issues: human body, human lifetime, knowledge and language.

b. Conceptual Analysis of Vulnerability and Finitude

i. Human Body

The human body including organic, sensory, emotional and cognitive aspects (*Leib*) can be seen as the basic medium of human-world-relations. It includes both material and cultural aspects[319]. In these two domains, the human body is traversed by significant forms of vulnerability and finitude. Physical asymmetry shapes the material relation between human bodies and robotic

309 Husserl, E. (1970) *The Crisis of European Sciences and Transcendental Phenomenology: An Introduction to Phenomenological Philosophy,* Evanston. US: Northwestern University Press.
310 Heidegger, M. (2010). *Being and Time*. New York, US: State University of New York Press.
311 Plessner, Helmuth. „Anthropologie der Sinne (1970)." In *Anthropologie der Sinne. Gesammelte Schriften III.* by Helmuth Plessner, 317-393. Frankfurt a.M.: Suhrkamp, 2003.
312 Janich, P. (2006a) *Kultur und Methode, Philosophie in einer wissenschaftlich geprägten Welt*. Frankfurt am Main: Suhrkamp.
313 Kamlah, W. & Lorenzen, P. (1984) Logical Propaedeutic. Pre-school of Reasonable Discourse. Lanham: University of America Press.
314 Rentsch, Thomas. *Die Konstitution der Moralität. Transzendentale Anthropologie und praktische Philosophie.* Frankfurt a.M.: Suhrkamp, 1999.
315 Rentsch, Thomas. *Heidegger und Wittgenstein. Existenzial- und Sprachanalysen zu den Grundlagen philosophischer Anthropologie.* Stuttgart: Klett-Cotta, 2003.
316 Seibt, Johanna (2018) Classifying Forms and Modes of Co-Working in the Ontology of Asymmetric Social Interactions (OASIS). In: M. Coeckelbergh J. Loh, M. Funk, J. Seibt, M. Nørskov (eds.). 2018. Envisioning Robots in Society –Power, Politics, and Public Space, Proceedings of Robophilosophy 2018 / TRANSOR 2018, Series; Frontiers in Artificial Intelligence and Applications, IOS Press, Amsterdam.
317 AI HLEG (2019) Ethics Guidelines for Trustworty Artificial Intelligence. High-Level Expert Group on Artificial Intelligence. 8. April 2019. European Commission, Brüssel [https://ec.europa.eu/futurium/en/ai-alliance-consultation (12. April 2019)]
318 Münkler, Herfried 2014: Der Wandel des Krieges. Von der Symmetrie zur Asymmetrie. 3. Auflage. Weilerswist.
319 Ihde, D. (2002). Bodies in Technology. Minneapolis, MN: University of Minnesota

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

bodies, since robotic bodies are made out of cooper, iron, and other non-organic materials that are more robust. Moreover, robots don´t feel pain – even when they simulate pain. On the other hand, human bodies are specialized to interact with open environment – including social life – enable whole body perceptions and sensations, gestures and many kinds of human behaviour on the basis of tacit knowledge[320, 321]. The physical fragility of human bodies correlates to a wide range of potentials to interact in contingent situations. We are vulnerable and flexible at the same time, which means a strong benefit in evolutionary processes. Human actions based on tacit knowledge are hard to be simulated by robots[322]. The physical robustness of robot's correlates to a narrow focus of interaction only in closed situations (like chess play or other very concrete problem solving situations). In conclusion the asymmetry between human and robotic bodies includes a physical and an interactive environmental related, socio-cultural side. **On the physical side of the coin, robotic bodies are stronger, realize more power, are more robust and less vulnerable (2.1.a.). On the other, interactive side, human bodies are less vulnerable than robotic bodies, because we are much more flexible to interact and survive within unexpected (natural and socio-culturally shaped) situations (2.1.b.)**[323].

Vulnerability and flexibility of human and robotic bodies:

2.1.a. Physical asymmetry (material robustness, power): robotic body stronger than human body

2.1.b. Interactive asymmetry (potentiality to adapt to contingent natural and socio-cultural situations): human body more flexible than robotic body

ii. Human Lifetime

Based on the embodiment of humans and robots, lifetime is another crucial factor that causes a certain form of asymmetry. Whereas vulnerability was the key term in the previous section, this one is about finitude. The conceptual asymmetry in this case is situated between a limited lifetime and immorality. Currently intensive debates in the fields of trans- and posthumanism include speculations about technical enhancement of human lifetime, brain uploading and cryonic – with the aim to generate organic or mental immortality. A critical assessment from an ethical point of view justifies the argument that not the elongation of human lifetime as such, but of the good and healthy life span is an appropriate scenario. Also, in technological transhumanism the finitude of human existence cannot be denied for logical reasons[324, 325]. We don´t live forever. Immortality enabled by technical means is not a rational paradigm for the

320 Polanyi, Michael 2009: *The Tacit Dimension*. With a New Foreword by Amartya Sen, Chicago & London.
321 Merleau-Ponty, Maurice 2012: Phenomenology of Perception. Trans. by Donald A. Landes. New York: Routledge.
322 H. Dreyfus, *What Computers Can´t Do: The Limits of Artificial Intelligence.* MIT Press, New York, 1972.
323 Funk, Michael (2019) Roboter- und Drohnenethik. Eine methodische Einführung. Wiesbaden: Springer, Chapter 6.4
324 Lorenz Sorgner S (2016) Transhumanismus. "Die gefährlichste Idee der Welt !?" Herder, Freiburg / Basel / Wien
325 Lorenz Sorgner S (2017) Was wollen Transhumanisten? In Göcke BP, Meier-Hamidi F (Hrsg) Designobjekt Mensch. Die Agenda des Transhumanismus auf dem Prüfstand. Herder, Basel / Freiburg / Wien, S 153-179

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

assessment of robotics in human societies. What can be observed both in the case of the human body and mind is an ideological asymmetry between between the struggle for successful happiness and bodily wellbeing in a limited period of lifetime on the one side (2.2.a. I.). On the other side there is quasi-religious hope for immortality caused by technical means (2.2.a. II.) – not by a fountain of youth like it was believed in the past. The contention about finitude in human life is primarily normative because it is intimately linked to question of a good life within a limited timeframe.

In the case of robotic bodies, the contention remains primarily empirical descriptive. **Since robots are means for human ends, their lifetime is related to the categories of technical functionality, efficiency and sustainability.** Additionally, the software-hardware-dualism, which can be interpreted as a technical manifestation of the philosophical body-mind-problem, causes a unique perspective to the question of immortality in robotics. Here again two paradigms can be differentiated that relate to a form of techno-ontological asymmetry. Due to attrition the physical robotic body includes a limited functional time frame (2.2.b. I.). Another question relates to the software level. On the one hand software can be seen as immortal as long as it is uploaded and shared on enough physical information processing systems, IT clouds or hard discs – that can be replaced by new physical systems when their function span comes to an end (2.2.b. II.a.). On the other side, also software can be interpreted as finite insofar it depends on the physical systems, but also on permanent updating. When it comes to new data formats, versions, machine languages or algorithms also the life frame of every software appears to be finite (2.2.b. II.b.)[326].

Finitude and lifetime of human and robotic bodies:

2.2.a. Human life: Ideological asymmetry between finitude and immortality

I. Finitude: Struggle for successful happiness and bodily wellbeing in a limited period of lifetime. From a logical point of view this is the only rational option for robot ethics.

II. Immortality: Quasi-religious hope for immortality caused by technical means. This relates to some speculations in the fields of post- and transhumanism.

2.2.b. Robotic functional working span: Techno-ontological asymmetry between hardware and software

I. Physical robotic body: Due to attrition the physical robotic body includes a limited functional time frame.

II.a. Immortal software: As long as it is uploaded and shared on enough physical information processing systems, IT clouds or hard discs that can be replaced the life span of software appears to be infinite.

II.b. Finit software: Permanent updating, new data formats, versions, machine languages or algorithms limit the life span of software.

[326] Funk, Michael (2019) Roboter- und Drohnenethik. Eine methodische Einführung. Wiesbaden: Springer, Chapter 10

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

iii. Human Knowledge

Beside communication (5.4.2.4) knowledge is one example for human vulnerability and finitude. For instance, in scientific research the finitude of theoretical claims receives a methodical status in the concept of falsification. Every theoretical claim is supposed to be falsified per definition, by more economic or comprehensive theoretical explanations[327, 328]. But also new empirical insights or results of experiments can falsify an existing theoretical explanation. Generally, the Archimedean point of pure objectivity does not exist. Here insights of quantum physics correlate to results of research in philosophical and cognitive sciences. First person experience plays a crucial role in all of our embodied epistemic actions[329, 330]. Therefore, our perspective on the world is limited, depends on the methodical conditioning of our actions – methodical order – in contingent environments[331]. **Tacit bodily and culturally embedded knowledge plays a fundamental role in human epistemic actions including human-robot-interactions**[332]. Since the human body is shaped by fragility and finitude, our knowledge is generally limited – which is true for the epistemology of human-robot-interactions as well. This circumstance is prominently reflected in scepticism, one of the oldest philosophical traditions since Ancient times. We have a finite and vulnerable epistemic perspective both in theoretical and in practical knowledge. By the way, **this is another reason why we build robots: we want to understand who we are by trying to recreate human-like features with engineering methods**[333, 334].

iv. Human Communication

Many more examples for human vulnerability and finitude can be mentioned. One with a significant role in human-robot-interaction is communication including linguistic, verbal and gestural, non-verbal expression. We speak in finite and vulnerable singular situations. Every speech act, every verbal or non-verbal performance comes to an end – that´s the finite side of the coin. Vulnerability can be observed for instance in the fragility of social understanding. While talking to dialogue partners we can be misunderstood, communication can fail. But we can also understand logically incomplete sentences (for instance poetry or metaphors). **Close to the human body, vulnerability of communication enables outstanding flexibility.**

---

327 Aristoteles (2007) Analytica Priora. Buch I. Werke in deutscher Übersetzung. Band 3. Teil I. Akademie Verlag, Berlin

328 Popper KR (2005) Gesammelte Werke in deutscher Sprache Band 3: Logik der Forschung Hrsg. v. Herbert Keuth11., durchgesehene und ergänzte Auflage. Mohr Siebek, Tübingen

329 Varela, Francisco J., Evan Thompson, and Eleanor Rosch. *The Embodied Mind. Cognitive Science and Human Experience.* Cambridge & London: The MIT Press, 1997.

330 Noë, Alva (2004): *Action in Perception*, Cambridge: MIT Press.

331 Janich, P. (2006a) *Kultur und Methode, Philosophie in einer wissenschaftlich geprägten Welt*. Frankfurt am Main: Suhrkamp.

332 Funk, Michael. "Humanoid Robots and Human Knowing – Perspectivity and Hermeneutics in Terms of Material Culture". In *Robotics in Germany and Japan. Philosophical and Technical Perspectives,* edited by Michael Funk, and Bernhard Irrgang, 69-87. Frankfurt a.M.: Peter Lang, 2014.

333 Decker M (2014) Who is taking over? Technology Assessment of Autonomous (Service) Robots. In Funk M / Irrgang B (Hrsg) Robotics in Germany and Japan. Philosophical and Technical Perspectives. Peter Lang, Frankfurt a.M. u.a., S 91–110

334 Funk, Michael, Johanna Seibt & Mark, Coeckelbergh 2018: "Why Do/Should We Build Robots? – Summary of a Plenary Discussion Session" in: Coeckelbergh, Mark, Janina Loh, Michael Funk, Johanna Seibt & Marco Nørskov (eds.) 2018: Envisioning Robots in Society – Power, Politics, and Public Space. Proceedings of Robophilosophy 2018 / TRANSOR 2018. February 14–17, 2018, University of Vienna, Austria. (Frontiers in Artificial Intelligence and Applications, 311). Amsterdam a.o.: IOS Press, pp. 369-384.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

We have the potentiality to adapt to unforeseen communicative situations, to a new slang or to unfamiliar body language in foreign cultures. Communication is much more than the transmission of information[335]. It´s something like practical treatment of uncertainty[336]. The ways we use language and create meaning are related to our existential bodily limitations: we don´t life forever, so we don´t speak forever; we don´t live in a logically unified world, so we don´t speak logically unified. In times of voice operated systems and linguistically controlled social robots the human-voice-robot-interface receives a new ethical topicality, such as the human-gesture-robot-interface[337, 338].

5. Discursive Frameworks for the Development of Inclusive Robotics

Within the relevant discursive frameworks for development of an inclusive robotics, important ethical-political ideas must be analysed: autonomy, dependency, vulnerability, functioning, care and disability. This development must take into account the regulatory framework of the UN *Convention on the Rights of Persons with Disabilities* (CRPD, 2006), and prospectively orient itself towards facilitating autonomy in the achievement of human functionings in inclusive environments. To this end, it is important to also pay attention to theoretical frameworks such as the capability approach (Nussbaum & Sen, 1993) and ethical conceptions of care (Kittay, 1999; Tronto, 2013). Additionally, another social group that should be considered regarding the consequences coming from the introduction of robotics is the group of children. As a mediating factor in the socialisation process, a responsible integration of inclusive robotics is necessary to maintain careful protections in regard to the inherent vulnerability of this group. The general regulatory obligations regarding protection, full development and wellbeing underlying the discourse in the *Convention on the Rights of the Child* (CRC, 1989) are a reference in order to safeguard the greater interests of the minor.

The CRPD and the CRC are international regulatory frameworks having legal repercussions in the majority of countries, without forgetting that as modulations in universal human rights, they also imply ethical requirements (Habermas, 2012) that are indispensable in different professions (Cortina, 2005). Hence, **to advance responsibly in a model for interactive inclusive robotics, in terms of wellbeing and justice, the professionals involved in their development should receive training in these judicial-moral documents. The implementation of a truly inclusive robotics should take into account the regulatory aspects of the CRPD and the CRC**, given the importance socio-technical mediations have for full enjoyment of human rights (Winner, 2007). Additionally, it should guarantee the conditions for good care, facilitating autonomy in the usual environments for activity and social participation.

---

335 Janich, Peter 2006b: Was ist Information? Frankfurt a.M.: Suhrkamp.

336 Wittgenstein, Ludwig. 1984b. "Über Gewißheit. Herausgegeben von G. E. M. Anscombe und G. H. von Wright." In *Werkausgabe Band 8. Bemerkungen über die Farben. Über Gewißheit. Zettel. Vermischte Bemerkungen*, 113–257. Frankfurt a.M.: Suhrkamp.

337 Coeckelbergh, Mark & Michael Funk 2018: "Wittgenstein as a Philosopher of Technology: Tool Use, Forms of Life, Technique, and a Transcendental Argument" in: Human Studies. June 2018, Volume 41, Issue 2, pp. 165–191. (Online first 12 January 2018) [(DOI) 10.1007/s10746-017-9452-6]

338 Funk, Michael 2018d: "Repeatability and Methodical Actions in Uncertain Situations: Wittgenstein's Philosophy of Technology and Language" in: Coeckelbergh, Mark, Michael Funk & Stefan Koller (eds.): Wittgenstein and Philosophy of Technology. Techné: Research in Philosophy and Technology. Special Issue, Volume 22, Issue 3 (2018), pp. 351-376. [(DOI) 10.5840/techne201812388]

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Today, numerous human functionings are mediated by robotic devices. Think about, for example, the environment in which people who are dependent, children, the elderly or those with disabilities are cared for. In this particular environment, for varying reasons, the introduction of such devices is encouraged as a substitute for traditional human care (Ortega, 2016). Nevertheless, resorting to these devices should not be a source of new discrimination regarding access to basic care and attention, and should indeed, however, **contribute to a "democratization of care"**, in the words of Tronto (Tronto, 2013), that also overcomes the enormous gender bias that has befallen this fundamental activity for social reproduction (Kittay, 1999).

a. Inclusive robotics under the framework of the Convention on the *Rights of Persons with Disabilities*

In addition to its regulatory dimension, the underlying theoretical discourse in the CRPD is important which, based on the *social model of disability*, interprets the disability as a social construction produced by the interaction between people with diverse bodily or mental functionings and exclusionary social structures (Palacios, 2008). The social model substitutes for the traditional *medical-rehabilitation model*, which restricts the disability to the individual sphere, by conversely trying to eliminate all barriers: physical, regulatory, political, economic, social, cultural or attitudinal that affect people with disabilities. Neither current developments in assistential robotics nor the majority of reflections on robo-ethics sufficiently take these questions into account, nor do they expressly refer to the CRPD (Unesco, 2017; EGE, 2018).

With respect to the values in the CRPD to be taken into account in developing robotics, the following are noted: 1) Respect for every type of human functional diversity. **Robotic care must deal with the wide diversity of human functionings**. 2) Inclusion requires universal accessibility and universal design. **Robots and robotic environments must be accessible and promote, in addition to health, social inclusion**. 3) Social participation of people receiving assistance demands having a **voice in all the phases** of development and implementation of robotics.

The change in discourse the CRPD has brought about allows the idea of care and other related ideas (vulnerability, autonomy, dependency) to be redefined. Care is now conceived as a right that is reinforced by the set of rights in this Convention and by technological mediation. Thus, what is stressed is a "public" vision of care, as opposed to another that is benevolent and with an enormous gender bias. There is also more awareness now that physical or mental vulnerability interacts with the social vulnerability resulting from stigmatization and from the discriminatory configuration of the socio-technological environment (Nussbaum, 2007). Limitations on personal autonomy in dependent people are no longer conceived as an inherent quality, but rather as a contextual condition that can be modified and is subject to attention in the framework of human rights. Nor is dependency, as a situation that requires care, conceived apart from the possibility of an independent life.

Altogether, some **implications of this conceptual redefinition for assistential robotics** are: 1) It needs to contribute to improving the social task of care-giving; 2) It needs to facilitate the autonomy of people receiving assistance in the social environments in which they develop; 3) It needs to form part of the socio-technological foundation that supports their rights as a whole.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Awareness on the rights of the disabled (applicable also to rights for the elderly) can favour reorientation of interactive robotics towards inclusive robotics, with objectives that are not only focused on rehabilitation, promoting greater human development in more inclusive social environments (Nussbaum, 2007). The possibility of inclusive environments for care, some of which involving technological and, especially, robotic mediation, requires that the diversity of functionings be valued socially. In this way, the discourse on "functional diversity" (Romañach & Lobato, 2005) places the need for care implied in natural human vulnerability within an inclusive social perspective.

b. Inclusive robotics under the framework of the Convention on the *Rights of the Child*

Structural changes that have taken place in Western families since the middle of the twentieth century have led to a redefinition of care to make it compatible with the understanding of children and adolescents as human beings who are in a growth process in regard to exercising their own autonomy (Honneth, 2014; Rojas Marcos, 2008). This conception of childhood and adolescence as relevant stages in personal maturation is clearly reflected in the policy implications in the CRC: the right to the necessary protection and assistance for their wellbeing, right to harmonious development of their personality in a family environment, etc.

Developments in robotics also have to be consistent with the policy guidelines arising from the rights established in the CRC. Supervision of the prospective processes to design, build and evaluate robotic devices aimed at children, generally more inclined towards an uncritical "technological fetishism", must be done from an ethical-legal principal focused on the best interests of the minors. These interests will be respected if, i) with such devices, any harm to the minors' abilities and personalities are avoided, subject to a progressive training process (principle of nonmaleficence); ii) if it is ensured that benefits will include comprehensive training of the personality and wellbeing (principle of beneficence); iii) if the minors' autonomy is progressively encouraged to make them capable of having self-control over their lives in a robotised socio-technical context (principle of autonomy); and iv) if equality in access to robotic devices responding to such principles is fostered (principle of justice).

Accordingly, **inclusive robotics must conform to an axiological framework characterised by**: a) **security**, which implies the construction of robots that are safe and collaborate in protecting life, physical and mental integrity, and the privacy of children; b) **graduality**: the design and construction of robots must adapt to the evolution of the physical and psychological characteristics in each stage of life, and c) **respect for development of the personality**: implementation of robotics must be respectful of the complete and harmonic development of the cognitive, emotional and moral dimensions that make up the developing personality of children.

In regard to the social group made up of children and adolescents, the CRC establishes in Article 23 that their education must be directed towards maximum development of physical and mental capacities, development of the personality and the assumption of a responsible and unprejudiced life. Accordingly, **robotic artefacts must be programmed in such a way that they can reinforce this designated direction in education**. The objective of incorporating these technological devices in educational processes should be to contribute to stimulating the different capabilities; **they must also be monitored**, as their continued use may actually lead

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

to undermining these same capabilities. It is also necessary to take into account that the use of robots does not negatively interfere in the shaping of the personality, with the consequent loss of self-esteem, and that this use does not facilitate the acquisition of prejudices due to the existence of biases in their programming.

Article 24 of the CRPD explicitly states that the right to an education is the right to an inclusive, quality education, recognising the right to an education without discrimination and on the basis of equal opportunities within an inclusive educational system for people with disabilities. With the aim of promoting their full participation and on equal terms in education and as members of the community, pertinent and personalised measures with reasonable accommodations and support should be adopted for full social inclusion (facilitating the learning of Braille, alternative writing, and orientation and mobility skills; learning of sign language and encouragement of linguistic identity for deaf people; appropriate augmentative and alternative means and formats of communication; and educational techniques and materials to support disabled students, etc.). The realisation of many of these measures is based on technological devices and, foreseeably in the short term, robotics.

c. Inclusive robotics and social innovation

The introduction of robotic devices in domestic environments or institutional ones (schools, hospitals, etc.) must be done in such a way that the overall dimensions of the care offered to create, consolidate and support personal autonomy **contribute to good development of human beings in all their different life stages**. This introduction should guarantee security, mental and physical integrity, emotional stability and respect for the dignity of the person (CRPD, 2006; CRC, 1989),

More than mere instruments, robotic artefacts are mediations arising from a certain socio-cultural context and which open up new possibilities in human functioning. **Interactive robots incorporate values coming from the different social agents involved in their design, manufacture and use. Each agent may have their own discourse in regard to their possible benefits or harm. Confrontation between different discourses by means of inclusive public debates constitutes a necessity** (Unesco, 2017; EGE, 2018) so that the introduction of interactive robots in society favours wellbeing and justice.

To avoid a "robotic divide", it will be necessary to pay close attention to the requirements for access to devices in relation not only to persons with disabilities. It should be kept in mind that universal accessibility (Article 9 of the CRPD) constitutes a fundamental right of persons with disabilities as it is the axis upon which rests fulfilment of many other rights. But it is also timely to point out that **although universal accessibility and reasonable accommodations are demands coming from the minority group of people with disabilities, these benefit a majority of users in several environments**. Actions to make demands and implement them transform the environments, and this results in benefits for many other users, so these actions could be considered examples of social innovations.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

6. Socially assistive robots for people with intellectual and developmental disabilities

a. Introduction

The global prevalence of intellectual disability is thought to be between 1 and 3% of the population[339]. "Intellectual disability means a significantly reduced ability to understand new or complex information and to learn and apply new skills (impaired intelligence). This results in a reduced ability to cope independently (impaired social functioning), and begins before adulthood, with a lasting effect on development" (World Health Organisation, 2019). Increased longevity in the general population has also resulted in more people developing disabilities, including cognitive disabilities, e.g. people with dementia (PwD). Intellectual disability can range from mild to severe or radical. Whilst mild cognitive disability would include learning difficulties or attentional disorders, severe cognitive disability might "limit or preclude the development of … the consciousness of oneself as a temporally-extended being; practical rationality—the capacity to govern one's actions by reasoning about how to act; and the capacity to make and respond to moral demands" [340].

These cohorts are likely to be one of the main beneficiaries of the development of *social robotics*. A variety of robots will be used in care contexts for people with intellectual and developmental disabilities (PIDD) and PwD. Robots are likely to be used to perform menial tasks (changing sheets, washing, picking things up, lifting patients). Also, robotic prostheses might be used to help patients perform physical actions. Some of the most successful robots perform social functions such as PARO, developed by the Japanese National Institute of Advanced Industrial Science and Technology, which operates as a stimulant for improved social interaction amongst patients and between the patients and their caregivers. These robots are commonly called ***socially assistive robots*** (SAR), a subcategory of social robots. Whilst social robotics "involves robots that engage in some form of social interaction with humans, through speech, gestures, or other means of communication," [341], SAR are "robots that **are designed to help through advanced interaction driven by user needs** (e.g. tutoring, physical therapy, daily life assistance, emotional expression) **via multimodal interfaces** (speech, gestures, and input devices)" [475]. SAR are designed to assist users though social interaction as opposed to physical interaction. SAR are used to **coach, motivate, and influence changes in behaviour**[342] . SAR may or may not incorporate artificial intelligence (AI). It is expected that AI might help SAR achieve a greater degree of autonomy, in order to lighten the burden on

---

[339] Maulik, P. K., Mascarenhas, M. N., Mathers, C. D., Dua, T., & Saxena, S. (2011). Prevalence of intellectual disability: a meta-analysis of population-based studies. Research in Developmental Disabilities, 32 (2011): 419

[340] asserman, D., Asch, A., Blustein, J., & Putnam, D. (2013). Cognitive Disability and Moral Status. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Fall 2013). Retrieved from http://plato.stanford.edu/archives/fall2013/entries/cognitive-disability/;

[341] Conti, D. et al., Robots in Education and Care of Children with Developmental Disabilities: A Study on Acceptance by Experienced and Future Professionals. International Journal of Social Robotics, 9 (2017): 51

[342] Scassellati, B., et al., "Robots for use in autism research". *Annual Review of Biomedical Engineering*, 14 (2012): 275

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

therapists and to provide a better therapeutic experience for patients[343].

SAR will likely transform the care sector. There is increasing evidence that SAR can produce benefits for users, particularly for users with **autism** spectrum disorder[344]. SAR are also being considered for people with **down syndrome**[345] and for people with cognitive disabilities such as **dementia**[346] There is evidence "that the presence of a physically embodied robot is associated with improved task compliance and more positive perceptions of the interactions" amongst people generally[347].

Whilst there is evidence of significant benefits to be realised from the use of SAR for PIDD or for people with cognitive disabilities (i.e. people with dementia), **certain ethical issues need to be addressed**. In order to maximise the benefits of SAR, it would be wise to address such ethical concerns as early as possible. We will outline a number of these concerns, focusing on privacy, autonomy, and responsibility. This section will primarily concentrate on the impact SAR are likely to have on people with intellectual and developmental disabilities or cognitive impairments such as dementia, but in some cases will note the ethical implications of widespread adoption of SAR for the wider society.

b. Privacy

Unlike robots designed to do menial tasks (e.g. lifting patients), and robotic prosthetics, SAR will likely gather large amounts of data on PIDD. After all SAR are more beneficial if they are

---

343 Esteban, P. G. et al.,"How to Build a Supervised Autonomous System for Robot-Enhanced Therapy for Children with Autism Spectrum Disorder". *Paladyn, Journal of Behavioral Robotics*, 8 (2017): 18–38.

344 Liu, C. et al., "Online affect detection and robot behavior adaptation for intervention of children with autism". IEEE Transactions on Robotics, 24 (2008): 883; Feil-Seifer, D., & Matarić, M. J. "Toward Socially Assistive Robotics for Augmenting Interventions for Children with Autism Spectrum Disorders". In O. Khatib, V. Kumar, & G. J. Pappas (Eds.), Experimental Robotics , 2009, 201–210, Springer Berlin Heidelberg; Scassellati, B., Admoni, H., & Matarić, M. Robots for use in autism research. Annual Review of Biomedical Engineering, 14 (2012): 275; Peca, A. et al.,. "How do typically developing children and children with autism perceive different social robots?" Computers in Human Behavior, 41 (2014): 268; Shamsuddin, S. et al.,"Design and Ethical Concerns in Robotic Adjunct Therapy Protocols for Children with Autism". Procedia Computer Science, 42 (2014): 9; Yun, S. et al., "A robot-assisted behavioral intervention system for children with autism spectrum disorders". *Robotics and Autonomous Systems*, 76 (2018): 58; Coeckelbergh, M. et al., "A Survey of Expectations About the Role of Robots in Robot-Assisted Therapy for Children with ASD: Ethical Acceptability, Trust, Sociability, Appearance, and Attachment". Science and Engineering Ethics, 22 (2016), 47; Sartorato, F. et al., "Improving therapeutic outcomes in autism spectrum disorders: Enhancing social communication and sensory processing through the use of interactive robots". Journal of Psychiatric Research, 90 (2017): 1; Esteban, P. et al., "How to Build a Supervised Autonomous System for Robot-Enhanced Therapy for Children with Autism Spectrum Disorder". Paladyn, Journal of Behavioral Robotics, 8 (2017): 18

345 Mehmann, H. et al., "Robot companions for children with down syndrome: A case study". Interaction Studies, 15 (2014): 99

346 Moyle, W. et al.," Exploring the effect of companion robots on emotional expression in older adults with dementia: a pilot randomized controlled trial". Journal of Gerontological Nursing, 39 (2013): 46; O'Brolcháin, F. (2017). Robots and people with dementia: Unintended consequences and moral hazard. Nursing Ethic; Tapus, A., Fasola, J., & Mataric, M. J. (2008). Socially assistive robots for individuals suffering from dementia. In ACM/IEEE 3rd Human-Robot Interaction International Conference, Workshop on Robotic Helpers: User Interaction, Interfaces and Companions in Assistive and Therapy Robotics; Valentí Soler, Met al., (2015). Social robots in advanced dementia. Frontiers in Aging Neuroscience, 7; Vitelli, R. (2013). Can Robots Help Care for the Elderly? Retrieved 22 March 2017, from Psychology Today.

347 Rabbitt, S. M. et al., "Integrating socially assistive robotics into mental healthcare interventions: Applications and recommendations for expanded use". Clinical Psychology Review, 35, (2015): 35

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

more autonomous, i.e. are able to interpret signals from users [348]. It is plausible that social robots will employ technologies that utilise facial recognition software, eye-tracking, emotion-recognition software, as well as image and sound recording.

This obviously poses a threat to users' privacy, which can be defined as the "right to appropriate flow of information"[349]. Allen has a broader understanding and distinguishes 5 different kinds of privacy: "informational privacy (e.g., confidentiality, anonymity, secrecy and data security); physical privacy (e.g., modesty and bodily integrity); associational privacy (e.g., intimate sharing of death, illness and recovery); proprietary privacy (e.g., self-ownership and control over personal identifiers, genetic data, and biospecimens); and decisional privacy (e.g., autonomy and choice in medical decision-making)" [350]. Beyond its moral importance privacy is enshrined as a right in numerous international documents (United Nations, 1948, art. 12) (UNESCO, 2005: art. 9) (United Nations, 2006: art. 22), the European Union (European Union, 2012: art. 7) (Council of Europe, 1997).

SAR might violate the privacy of both PIDD as well as those who interact with them. SAR pose threats to each of Allen's five types of privacy. A SAR capable of recording and transmitting data might gain confidential information or secrets. If the SAR is hacked this information could be accessed. This risk is real not just for the patient but for anyone in close proximity to them. In terms of physical privacy – a SAR might be considered to breach physical privacy particularly if PIDD or their careers are concerned about what the robot might transmit or record. An anthropomorphic SAR might upset some PIDD in terms of protecting modesty or upset their carers depending on how the SAR is perceived (e.g. as a person watching). Risks to autonomy (discussed below) pose some threats to the idea of self-ownership, thus constituting a threat to proprietary privacy. These are unlikely to manifest for carers or visitors who only have slight interactions with the SAR. A SAR might also be able to record and store personal identifiers, genetic data and biospecimens as the SAR will be in close proximity to the user and be able to access these. This risk is also present for carers and others but is of course less of a threat due to their limited time in the company of the SAR. Finally, if the SAR is used to nudge the user, then it will pose a potential threat to autonomy and choice in medical decision making. PIDD are most at risk of having this type of privacy breached, but the SAR might also be able to influence carers. PIDD often struggle to understand privacy settings and what is being asked of them in relation to privacy[351]. Thus, whilst consent may be requested and received, it will not always be *informed* (more on this below). Those who interact with PIDD risk having their privacy breached as well. Anyone visiting the user would risk being recorded as soon as they are in the vicinity of the SAR. At the very least, SAR must include settings designed to protect the privacy of users and others.

---

[348] Yun, S.-S. et al., "A robot-assisted behavioral intervention system for children with autism spectrum disorders". Robotics and Autonomous Systems, 76 (2018): 58

[349] Nissenbaum, H. (2009). Privacy in Context: Technology, Policy, and the Integrity of Social Life. Stanford, CA, USA: Stanford University Press.

[350] Allen, A. (2016). Privacy and Medicine. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Winter 2016).

[351] Chalghoumi, H. et al., (2017). Information Privacy for Technology Users with Intellectual and Developmental Disabilities: Why Does It Matter? Ethics & Behavior

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

The loss of privacy, i.e. the loss of control over information about oneself, has significant implications in other areas. To put it succinctly, the more data gathered about a person, the easier they are to "nudge" or to manipulate as will be discussed below.

c. Autonomy

Promoting the autonomy of PIDD is often cited as a goal of carers working the area. Increasing or improving the autonomy of PIDD is a goal of the United Nations Convention on the Rights of Persons with Disabilities (United Nations, 2006: Preamble (n)). Autonomy, however, is not always clearly defined. Often it is taken to mean living independently or semi-independently. The philosophical understanding of the concept emphasises "self-rule that is free from both controlling interference by others and limitations that prevent meaningful choice, such as inadequate understanding"[352].

It can be expected that SAR will have a prominent position in the lives of PIDD. As mentioned in the section on privacy, it is likely that SAR will possess a lot of information about users – about their preferences, needs, and deficits. Such data are likely to improve the performance of SAR as assistive devices. As such, SAR will have the capabilities to "nudge" or manipulate PIDD to behave in a certain way. While manipulation carries negative connotations, nudging is more ambiguous. According to Thaler and Sunstein, "A nudge, as we will use the term, is any aspect of the choice architecture that alters people's behavior in a predictable way without forbidding any options or significantly changing their economic incentives. To count as a mere nudge, the intervention must be easy and cheap to avoid. Nudges are not mandates. Putting the fruit at eye level counts as a nudge. Banning junk food does not" [353]. Ménard's definition of a nudge is particularly pertinent to SAR. "Essentially, a 'nudge' consists in an intervention, which aims to suggest one choice over another by gently steering individual choices in welfare enhancing directions yet without imposing any significant limit on available choices" [354] . Clearly, not every instance of being influenced by someone or something else constitutes a violation of autonomy or something more generally morally problematic. Thus, it is easy to see how SAR (or robots more generally) could help PIDD or PwD live more autonomously in the sense of living without as much need for human help. As well as doing menial chores and helping with physical activities, robots might be able to remind people to do certain tasks, take medicine and so on.

However, there are also concerns in relation to autonomy and SAR. There are ongoing debates about the "nudging" capabilities of new technologies, including assistive devices for PIDD[355]. Robots, as distinct from chatbots or robotic prosthetics in being both physically embodied and separate, have distinct advantages over other technological devices in terms of potential to mould behaviour. They will be present in the user's life in a way that chatbots are not and be capable of interacting with the user in myriad ways (Borenstein & Arkin, 2016: 35). SAR will be

---

352 Beauchamp, T. L., & Childress, J. F. (2009). Principles of Biomedical Ethics (6th ed.). Oxford: Oxford University Press.
353 Thaler, R. H., & Sunstein, C. R. (2008). Nudge: Improving Decisions about Health, Wealth, and Happiness. Yale University Press.
354 Ménard, J.-F. A 'Nudge' for Public Health Ethics: Libertarian Paternalism as a Framework for Ethical Analysis of Public Health Interventions? Public Health Ethics, 3 (2010): 229–238. https://doi.org/10.1093/phe/phq024
355 O'Brolcháin, F. (2017). Robots and people with dementia: Unintended consequences and moral hazard. Nursing Ethics,

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

designed to mould the user's behaviour, perhaps helping persons with Autism Spectrum Disorders develop their social skills.

This means that we need to be cautious with SAR and the sorts of influence they might have over users but does not mean that we need to ban their use. Not all instances of nudging or influencing other people constitute morally repugnant violations of an individual's autonomy. Parents and teachers influence children, friends make suggestions to each other, doctors provide advice to patients and so on. We are a social species that learn from each other and SAR can be a useful tool in this regard. Attempts to influence someone so that they flourish – have a more fulfilled or healthier life – is at the core many ethical systems. Yet, insofar as we value autonomy, there are limits on how far this influence should go – we like to ensure that the person still has a choice regarding the advice, that they can decide for themselves how to respond to attempts to influence them. It appears then, that there is a strong case that nudging in certain scenarios – in order to promote the user's well-being – is acceptable. Nonetheless, there remains the question of the degree to which "nudging" is permissible. **Beyond a certain point, such nudging will veer from being helpful to being a violation of autonomy**, i.e. the SAR might nudge a person into doing something they would not have desired to do (in normal circumstances) or something that is harmful, self-destructive, illegal, or immoral.

We must also question whether the use of SAR to promote moral behaviour in PIDD is acceptable [356] . There seems to be a case for using SAR to ensure that users take medication at the correct times and other health related matters as it will benefit the user directly, but the case for nudging PIDD to behave in *moral* ways is less clear. By *moral*, we can consider suggesting that users be kind, honest, generous, brave, etc. The first problem would be to decide which moral principles should be chosen. Even if this objection was answered, some might feel that nudging people into ethical behaviour is a violation of the person's autonomy and constitutes too much interference in a person's personal identity. However, this assumes that personal autonomy is sacrosanct and should never be interfered with. Many traditions (religious traditions, virtue ethical traditions) will have no difficulty in training people to behave in certain ways and will argue that this is in their interest. Even within a liberal individualist tradition, there is an argument that such nudging is acceptable when dealing with children or those who have diminished competency, as they do not possess sufficient rationality to make decisions for themselves [490]. In these scenarios, the assumption is that due to youth or diminished competence, certain people are not capable of making rational autonomous decisions. It follows then that nudges (or more) made in their interest will benefit them. On the other hand, if one of the goals of SAR (and assistive technologies in general) is to promote the autonomy of users, determining their values (i.e. nudging them to behave in certain *moral* ways) risks undermining their autonomy via the very tools designed to promote it. All options may well be left on the table, but the values imparted to the user will mean that they are extremely unlikely to choose does that conflict with the chosen values.

If SAR are used in conjunction with carers or therapists, these latter will play a role in supervising how the SAR interacts with the user, including the prompts and nudges that take place. **Protecting the autonomy of the user will be the responsibility of the carer**.

[356] Borenstein, J., & Arkin, R.. "Robotic Nudges: The Ethics of Engineering a More Socially Just Human Being". Science and Engineering Ethics, 22 (2016): 31

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

However, if SAR are used to help users live independently, **oversight will be needed to ensure that the SAR are not nudging people to such an extent that they end up controlling the user. Government agencies** that oversee or regulate the care sector would be the most likely sector capable of providing such oversight.

d. Informed Consent

The solution most frequently advocated in relation to questions of privacy and autonomy is that of *informed consent*. Informed consent has a dominant role in bioethics, having usurped older paternalistic models of doctor-patient relationships. If SAR are to be used as *assistive* devices for PIDD, informed consent will be required. So if SAR are going to collect data, nudge users, and generally interact with users, users would normally be required to consent to this, i.e. accept that data will be collected, or that the SAR will nudge them, or that interacting with the SAR constitutes an element of treatment (i.e. that their interactions with the SAR are teleological – the goal being the well-being of the patient). Furthermore, whether the SAR are being used in a domestic or a research context will alter the requirements of informed consent. **The requirements of informed consent are more stringent in a research context than in a therapeutic context.** This is because research contexts are generally less focused on the welfare of the users (they are often focused on the needs of the researchers) than therapeutic contexts thus requiring additional protections of the wellbeing and safeguards for the autonomy of research participants. In a therapeutic context, the interests of the physician (or the SAR) can be expected to align with the interests of the patient. Normally, informed consent is not required in a domestic context, except when the user is signing up to terms and conditions. Given the privacy implications of SAR, the terms and conditions governing the use of SAR will no doubt require some element of informed consent. Yet, the conditions of informed consent in relation to both PIDD and PwD are complex as they frequently struggle to understand what is being asked of them and what is at stake in terms of privacy and autonomy[357].

As intellectual and developmental disabilities form a spectrum, each user will have different capacities. Some will be able to provide informed consent in relation to SAR, others will not be able provide informed consent depending on their *decisional capacity*. This is "defined as the *ability* of health care subjects to make their own health care decisions"[358] and is the "cornerstone of the legal determination of competence (whether or not one is legally able to make self-determining decisions" [359]. Decisional capacity of PID and PwD will vary according to the degree and severity of the cognitive impairment of the subject. Intellectual disability is a spectrum, and PwD may be subject to 'fluctuating capacity', i.e. changes in their in decision-making capacity as a result of cognitive fluctuations[360]. A person in the early stages of dementia might be able to provide informed consent for most of the issues at hand most of the time, but

---

357 Chalghoumi, H. et al.,(2017). Information Privacy for Technology Users with Intellectual and Developmental Disabilities: Why Does It Matter? Ethics & Behavior.
358 Charland, L. C. (2015). Decision-Making Capacity. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Fall 2015).
359 Fields, L. M., & Calvert, J. D." Informed consent procedures with cognitively impaired patients: A review of ethics and best practices." Psychiatry and Clinical Neurosciences, 69 (2015): 462
360 Bradshaw, J. et al., "Fluctuating cognition in dementia with Lewy bodies and Alzheimer's disease is qualitatively distinct". Journal of Neurology, Neurosurgery, and Psychiatry, 75 (2004), 382; Trachsel, M. et al., "Cognitive fluctuations as a challenge for the assessment of decision-making capacity in patients with dementia". *American Journal of Alzheimer's Disease and Other Dementias,* 30 (2015): 360

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

as the disease progresses this capacity will diminish. Similarly, a person with a mild intellectual disability might be competent enough to provide informed consent to some aspects of their interaction with the SAR but not to others. In cases, where the user is not competent to provide consent, a surrogate will be required, e.g. a parent, guardian, or someone responsible for the user. Surrogates generally adopt one of two standards – the best interests standard or the substituted judgement standard [361]. The first aims at doing what is in the best interest of the incompetent person, the second aims a reconstruction of the incompetent person's preferences, e.g. through accounts of proxies or advance directives. If both approaches are feasible, **respect for autonomy demands prioritising the substituted judgement standard over the best interest's standard. However, in cases of PIDD where the impairment is of such severity that they never have had the capacity to make judgements, the proxy will have to rely on the best interest's standard.**

While consent must be given voluntarily, there are various obstacles such as coercion, undue inducement, and the absence of a choice[362]. **It might be possible for the SAR to regularly prompt the user to consider whether they still consent to the interactions. The SAR would then be able to facilitate *rolling consent***, whereby the user is constantly able update their preferences in relation to data-gathering, etc. However, if SAR are designed to prompt users to behave in certain ways, it is plausible that they would equally be able to prompt users to provide consent. Similarly, the very presence of the SAR, particularly if it is life-like and friendly, might encourage the user (or their guardians) to trust them and to not properly consider whether they ought to consent. As such, SAR might make the *voluntariness* condition of informed consent difficult to determine.

Whilst informed consent will be needed if SAR are to be used with PIDD, it will not be sufficient to allay concerns regarding loss of privacy and threats to autonomy. Users and their proxies are likely going to see the benefits SAR offer and want to avail of them, whilst producers of SAR will be keen to emphasise the benefits. This emphasis on the positive aspects of SAR is not undeserved, but government oversight will be required to make sure that privacy and autonomy norms are upheld. This is particularly the case in relation to such vulnerable populations.

e. Deception

Concerns have been raised about whether the use of "emotional" SAR will be deceptive[363]. The concern here is that SAR are deceptive in that they a) intend to deceive, b) their emotions are unreal, and c) the robots pretend to be something they are not[364] . Coeckelbergh suggests that a proper understanding of *appropriate* emotional responses in the social context should replace

---

[361] Jaworska, A. (2017). Advance Directives and Substitute Decision-Making. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Summer 2017).
[362] Eyal, N. (2012). Informed Consent. In E. N. Zalta (Ed.), The Stanford Encyclopedia of Philosophy (Fall 2012).
[363] Wagner, A. R., & Arkin, R. C. " Acting Deceptively: Providing Robots with the Capacity for Deception". International Journal of Social Robotics, 3 (2011): 5; Coeckelbergh, M.et al., "A Survey of Expectations About the Role of Robots in Robot-Assisted Therapy for Children with ASD: Ethical Acceptability, Trust, Sociability, Appearance, and Attachment". Science and Engineering Ethics, 22 (2016), 47; Borenstein, J., & Arkin, R.. "Robotic Nudges: The Ethics of Engineering a More Socially Just Human Being". Science and Engineering Ethics, 22 (2016): 31;
[364] Coeckelbergh, M. Are Emotional Robots Deceptive? IEEE Transactions on Affective Computing, 3 (2012): 388

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

ideas about *authentic* emotional responses. However, Coeckelbergh is focusing on affective robots designed for the general population. The issues he raises are complicated where PIDD are concerned as many PIDD may only partially understand the context within which interaction takes place and thus be at greater risk of being deceived. **PIDD are therefore more likely to attribute emotional states to robots or to believe they are interacting with a person**.

f. Designer Bias

The values that a SAR promotes have to be encoded into the robot, thus placing a considerable degree of responsibility on the designer. Designers will have to ensure that SAR do not harm the user, or others; what data they will collect and how they will use it; how they will respond to what the user, and others, do; what behaviours are promoted; and, if ethical nudges are permitted, which ethical principles the robot promotes. The choices relating to how SAR behave, and which behaviours they promote in the user, are of course value laden. They will implicitly endorse certain ethical values and condemn others. As such, the issue of designer bias must be addressed, both in terms of explicit and implicit bias.

Explicit bias is easier to address – as regulations can be put in place to ensure that certain behaviours are not encoded into SAR, e.g. that they will not be homophobic or racist. Implicit bias will be more difficult to address. Given that SAR are intended to help PIDD, the presence of implicit but obscured societal biases against people with disabilities in the algorithms underpinning the SAR would be pernicious. For instance, insofar as there are biases regarding the sexual desires of PIDD (that they either do not or should not have sexual desires), these could be built in to SAR – with SAR's discouraging or condemning sexual activity. Every effort should be made to avoid this. Should SAR make extensive use of AI, the issue of bias will be further complicated. If AI is involved in building algorithms that will determine the behaviour of the SAR, there is the potential that biases inherent within any system will be exacerbated – the AI will use existing data sets that might contain biases. **New algorithms built by the AI risk being so opaque that it will be impossible for anyone to determine the source of any biases**. For more on this, see (Mueller, Forthcoming)

g. Caring

Much of the literature on SAR stresses that they will be best used in conjunction with human carers/therapists. The benefits offered by SAR suggest that they will provide a valuable tool. Consideration of SAR use in care contexts (including for PIDD) must also take into account the wider societal implications of SAR and automation. One of the chief worries in this regard is job losses as a result of automation. **The many benefits of SAR for PIDD (and in care in general) should not be used as a justification for a more general adoption of SAR, as a trojan horse. However, many of the risks associated with SAR for PIDD will also be present in SAR designed for the general population**.

There is also a risk that fears about job losses arising from the adoption of SAR and automation in general will create resistance to SAR in the care professions. This appears to be misplaced - it has been suggested that the caring professions are unlikely to be badly affected by automation because the caring professions involve "soft" skills that are hard for machines to replicate. This might mean that the caring professions become highly desirable and more competitive. The standards amongst human employees might be raised even higher, though wages might stagnate.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

If we widen our perspective again, we should consider how the potential loss of tax revenue from automation in the rest of the economy might impact on the care sector. There is a risk that a lower tax intake might see the care sector for PIDD squeezed insofar as it is reliant on state funds. Lower tax revenue might be allocated to areas with better representation for instance. This might create a temptation to save money by replacing human carers with robotic carers where possible.

In a more philosophical sense, the use of SAR in care, will mean that the rest of society has less reason to interact with PIDD. Philosophers such as MacIntyre argue that **an awareness of human vulnerability is essential for human flourishing and that this awareness rests on practices and interactions with those who are vulnerable**[365]. Potentially, the use of SAR in care settings might result in the diminishing of the sense of human vulnerability necessary for a flourishing society[366] or the exacerbation of the mechanistic side of contemporary healthcare[367].

7. Robotics as an instrument for social mediation

a. Activity and social participation in functioning environments

Different environments for activities and social participation, such as the urban environment, domestic environment, educational environment, work environment, environments for public, political, social and economic participation, environments for information, communication, science, culture, leisure or health, can be considered "functionings environments", in the sense of the capability approach from Amartya Sen and Martha Nussbaum. This idea deals with spaces (real or virtual) where we carry out actions and practices, we consider to be important and valuable for our wellbeing and quality of life [368]. In this approach, wellbeing consists of evaluating life conditions defined by functionings. These represent what a person achieves or becomes in the development of their life, considered a set of interrelated functionings[369].

It is characteristic of functionings environments to contain a diversity of possible functionings in them, an aspect that relates to one of the dimensions in the "functional diversity" concept (Toboso, 2010), which can be integrated quite naturally into the capability approach.

Taking into consideration human diversity, of individual characteristics as well as contextual circumstances, constitutes a very important aspect in Amartya Sen's capability approach (Sen, 1998). The relationship has already been noted between the concept of functional diversity and the importance of considering human diversity in Sen's approach, making the case that this consideration should also be broadened to include the particular characteristics of realising the different functionings. **In an evaluation of wellbeing and quality of life, the "capability**

---

[365] MacIntyre, A. (1999). Dependent Rational Animals: Why Human Beings Need the Virtues. London: Duckworth.
[366] O'Brolcháin, F. (2017). Robots and people with dementia: Unintended consequences and moral hazard; Robson, A. (2018). Intelligent machines, care work and the nature of practical reasoning. Nursing Ethics,
[367] coeckelbergh, M., B. "A Survey of Expectations About the Role of Robots in Robot-Assisted Therapy for Children with ASD: Ethical Acceptability, Trust, Sociability, Appearance, and Attachment". Science and Engineering Ethics, 22 (2016), 47.
[368] M. C. Nussbaum & A. Sen (Eds.), The Quality of Life, New York: Oxford University Press, 1993.
[369] A. Sen, "The Standard of Living", in G. Hawthorn (Ed.), The Standard of Living. Cambridge: Cambridge University Press, 1987.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

set" (the set of functionings that are of value to people, carried out in their relevant environments) should include all the variety of possibilities in the performance of valued functionings, in line with the fact that different people have different ways of achieving identical functionings. Thus, for each particular functioning it should be taken into account the different possible ways of carrying it out and not be limited to the majority, standard and most common ways.

The assumed margins of human functional normality are very narrow, and to be outside of them implies accepting the diversity of functionings that have been altered by circumstances. An inclusive society that aspires to equal opportunities should promote and keep these other possible functionings active so that a larger number of people can have access to them if they need them. This is one of the reasons why we consider it important and necessary to value functional diversity socially, given that this expands the space of possibilities for functioning to other less common ways that, nevertheless, could be taken advantage of by everyone [370].

The wider the set of functionings available in a society the more inclusive it will be, and it will offer greater life possibilities to people. A set like this will also offer greater freedom of choice and will favour equal opportunities.

b. Technological mediation in functionings environments: transformation of practices, representations and values

Functionings environments are discursive spaces (of practices and of representations) where the values that characterise them and define them, "live". The relationship between sets of practices and representations in a given environment are mediated by the environment's own characteristic values.

The introduction of technological elements and, in particular, robotic elements, into functionings environments transforms the way in which the actions are carried out and, consequently, can modify evaluation of the functionings that are mediated by such elements. We can verify this statement by looking at the case of *smartphones*, a mediation technology present today in practically every environment: as almost everything we do is now done with this device, it appears that its use even defines what is valuable to do; it is as if the actions we do by means of using this device have greater value than those that do not need it. The practice of use of the device seems to define the value of what we do with it.

Let's look at the case of a community that begins a relationship with a certain technology by means of sets of values, representations and practices of the community in regard to it. This way of understanding the relationship allows us to talk about the community's discourse about this technology. But there is an analogous discourse, linked to the technology in question, which also brings with it values, practices and representations coming from the communities responsible for its idea, design, development, manufacturing, marketing and commercialisation. This artefactual discourse, materialised in the technology, is introduced into the functionings environments where it is used and can come to modify the community's discourse.

[370] M. Toboso, "Rethinking disability in Amartya Sen's approach: ICT and equality of opportunity", Ethics and Information Technology, 13 (2010): 107-118.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

As devices that mediate actions, robots not only transform the practices carried out in an environment, but also its characteristic values.

The aspects that a community of users can evaluate in their relationship with any type of device are very numerous. Among them we highlight the following values, distributed in the instrumental dimension in connection with access and use, and in the functional dimension linked to security and privacy: availability, affordability, necessity, appearance, simplicity, ergonomics, accessibility, usability, versatility, efficiency, quality, reliability, security, intimacy and privacy.

But it is also necessary to evaluate devices in an ethical dimension. To do this, it is essential to define which values should be taken into account in the evaluation. Accordingly, Romero (2017)[371] considers the following four principles in order to create a more complete evaluation process for technological devices:

1. Sustainability: requires verifying and analysing the impact of technologies on the contamination of land, the atmosphere, or on the system for recycling materials.
2. Precaution: parallel to the principle of non-maleficence (one of the traditional principles in Bioethics), these criteria supports adopting cybersecurity and protective measures to confront the suspicions related to future risks associated with the implementation and use of certain technologies. This principle should involve carrying out evaluations on impacts prior to their implementation.
3. Privacy: the user should know or be informed about privacy procedures online, for their security and anonymity, as well as about the privacy systems in the hardware and software. At this point it should be remembered that, with regard to privacy, in addition to a personal dimension, this is also present socially. One of the key functions of privacy is to control the perverse and abusive consequences deriving from the asymmetrical power relationships between individuals and organizations.
2. Democracy: along with the principle of autonomy, the defence of digital rights should be promoted, in institutional organisms, as Human Rights, just the same as cybersecurity in domestic, professional, state and crucial infrastructures (such as hospitals, nuclear power plants, airports, water supplies, etc.). The definition of this principle poses some problems and could be substituted by the possibility of citizens being able to intervene in the procedures used for handling data in these institutions and infrastructures.

c. Social appropriation of technologies in functionings environments

Let's now consider how the practice of using a technology is connected to the values that motivate said practice and, at the same time, how it can come to define these values. In the first case, we assume the motivational priority of the users' discourse over the artefactual discourse. That is, certain values forming part of the users' discourse motivate them towards the practice of using the technology in question. In the second case, we have the priority of the practice of use over the value, which reflects the priority of the artefactual discourse over the community's discourse, as if through this practice of use the values that motivate the practice are constructed and encouraged, in a type of circular feedback.

[371] J. Romero, "CiberÉtica como ética aplicada: una introducción", Dilemata. Revista Internacional de Éticas Aplicadas, 24 (2017): 45-63.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

We could ask ourselves if the "value" appearing in both cases is the same. That is, **does the motivational value from the community (which guides the practice of use of the technology) coincide with the artefactual value (that is encouraged by said practice)?** Usually these values are different, and the difference between both is precisely what accounts for the difference between the community's discourse and the technology's artefactual discourse.

In the case in which both values coincide it could be said, in respect to this value, that both discourses are concurrent and that the technology in question "satisfies," optimally, the motivational value of the community. Easily understandable and desirable examples of this satisfaction are those that refer to values, for example, such as availability, price, accessibility, quality or security in the use of the technology in question. The satisfaction of values, such as those previously referred to (instrumental, functional, ethical and others), which can also be interpreted as the concurrence in them of the community's discourse and the artefactual discourse, is the basic condition for "social appropriation" of the devices by the community, which we should differentiate from the simple concept of "adopting" them. **Adoption happens through the practice of use, as if it was simply guided by artefactual values, while appropriation happens by means of the co-constructive combination of practices and values**, requiring as such satisfaction of these values in practices, as values from the community, **and also involves social representations** that are favourable towards the device in question.

One of the most important dimensions of social appropriation is the axiological dimension, which relates to the values that guide the practices and representations that different communities have in the functionings environments coming from the implicated technological innovations.

If in a given environment the situation arises in which mediation from a device is obligatorily needed to perform a certain functioning, the limitations of this device will define the limitations of this action. In the most extreme case, but at the same time the most evident, if the device is not available, performance of the functioning will be nullified, which will negatively affect the scope of wellbeing associated with its achievement.

Artefactual dependency is one of the undesirable effects of its adoption. In the case of robotic devices, the barriers will produce a "**robotics divide**" analogous to that already known as the "digital divide." These **"divides" reflect inequalities of access and use, whose combination produce an inequality of greater importance: inequality of opportunities for the affected people or communities.** Demanding genuine equal opportunities brings up important questions having to do with ethics, politics and social justice, and related to which value frameworks should be considered the most relevant when considering what is needed for equality.

d. Robotic mediation devices "for a better society"? Addressing the complexity of this issue

Within a research on inclusive robotics for a better society, we should ask ourselves which aspects of current society are the ones seeked to improve with the introduction of these robotics. We could also ask which human functionings environments aspire to be better, and to what extent, by means of said introduction.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

More than one hundred years ago, Schumpeter spoke about processes for innovation as being "creative destruction" [372]. Recently, Echeverría has been reflecting on the good and bad of innovation, assuming that all innovation has its beneficiaries but also those who are harmed [373]. The question that concerns us is how to distribute the benefits and the damages. It would be desirable for this to be done in a just and equitable way toward the conditions of functionings environments, on which the wellbeing and quality of life of people may, critically, depend.

The introduction of devices (technological, robotics, etc.) in functionings environments transforms not only the landscape of practices in the environment, but also the panorama of its values. **The introduction of robotic mediation elements can notably affect the structure of the functionings environments, just like when, in the same way, the biological balance of an ecosystem is affected** if an invasive or predatory species is introduced. Similarly, the balance in the space from the functionings that certain actors do within this environment can be critically affected.

In the case of the work environment, the functionings affected refer to the daily chores and circumstances that constitute the work itself of the workers in their posts[374]. In the same way, profoundly transformative effects can be predicted in the educational environment, with the promotion of education mediated by assistants, not technological but rather robotic, and with the consequent introduction of the figure of the robotic professor. The same influence is currently predicted in care-giving environments, where the supposed advantages of affective robots, robotic assistants and care-giving robots are already being advertised. This is being done without taking into account that care-giving relationships imply elements that are essentially human, intrinsically intersubjective, and to care for a person with Alzheimer's, to name one case, the best "technology" is another human being[375].

In relation to this, we propose to study the effect of interactive robotic devices on the wellbeing and quality of life of people and communities by means of the capabilities approach and functionings of Amartya Sen and Martha Nussbaum. The main advantage of using this approach for analysing the effect of robotic innovations on functionings environments is that it introduces important ethical and social considerations on questions that, at first glance, could appear to only be technical or instrumental.

---

[372] . A. Schumpeter, Theorie der wissenschaftlichen Entwicklung, Berlin, Duncker & Humblot, 1912: english translation, The Theory of Economic Development, Boston, Harvard University Press, 1934.
[373] J. Echeverría, Innovation and Values: A European Perspective, Reno, NV, University of Nevada Reno, Center of Basque Studies, 2014.
[374] M. Goos, "The impact of technological progress on labour markets: policy challenges", Oxford Review of Economic Policy, 34 (2018):362–375.
[375] R. de Asís Roig, "Ethics and Robotics. A First Approach", The Age of Human Rights Journal, Nº 2, (2014):1-24.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

8. Ethical issues relating to the environmental impact of robotics

a. Introduction and background

The advent of a society in which robots are commonplace brings with it many ethical questions. Whilst many of these issues are receiving increased attention[376] , the ethical issues relating to the environmental impact of widespread use of robots has garnered little attention. Nonetheless, **if one of the aims of using robots is to create a *better* society, the environmental impact of robotics cannot be ignored.** New industries, institutions, practices and markets will develop around robotics. This can be more or less environmentally sustainable. This section touches on some of the basic issues that should be considered, specifically environmental justice, consumerism and well-being. Building on that some topics are recommended for further research, such as energy, resource use, disposal, moral hazard, social unrest, and nudging.

Given the stark warnings about the human impact on the natural environment (IPCC, 2014) (World Wildlife Fund, 2018), ensuring that the development of interactive robots is minimally disruptive – or better yet is beneficial – to the natural environment ought to be a priority. The merits of an anthropocentric (human centred) or ecocentric (ecologically centred) ethical approach, or whether non-human nature is merely instrumentally valuable or inherently valuable will not be a focus of this section. Suffice to say, irrespective of ethical or political stance – a functioning environment is required to pursue any individual or social good. This nascent stage of the development of interactive robotics is the best opportunity to this – if framed correctly, it will be possible to ensure that the development of robotics will be environmentally beneficial or minimally harmful. If not, there is a risk that society will become locked into systems and industries that continue to degrade the natural environment and thus imperil the continued existence of non-human animals, human civilisation, and potentially humanity's continued existence.

In 2011, John Sullins (2011) suggested that **we need a green agenda for robotics**. More recently van Wynsberghe and Donhauser focused on critical applications of **robots "for environmental research, engineering, and remediation** have received next to no attention in the roboethics literature to date" [377] . They suggest "that environmental robotics will (and arguably should) play an increasingly prominent role in environmental protection and resource management in years to come" (1780). This is because it is becoming easier "becoming easier and easier to use robots to monitor environmental conditions and hazards and to apply robots to aid in addressing certain environmental issues[511] (van Wynsberghe & Donhauser, 2018: 1780).

The focus of that paper was on various types of robot that would be used in environmental work, rather than on the environmental impact of robots *as such*. They distinguish between

[376] Müller, V. (Forthcoming). Ethics of Artificial Intelligence and Robotics. More info needed
[377] Van Wynsberghe, A., & Donhauser, J. "The Dawning of the Ethics of Environmental Robots". Science and Engineering Ethics, 24 (2018): 1777

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

"robots-in-ecology"[378], "robots-for-ecology"[379] and "ecologically-functional-robots" or ecobots[380] (van Wynsberghe & Donhauser, 2018: 1778). However, many of these robots will have little to no interaction with humans (depending on their level of autonomy). Nonetheless some of the issues briefly alluded to in the paper are extremely relevant to us; specifically, "the **environmental impact of the type of materials used to make the robots; the environmental impact of the degradation process when the robot is no longer in use; the process for testing the robot**"[381].

b. Environmental Justice

These latter issues – issues of environmental justice – merit considerable talk and debate. For instance, the provenance of the raw materials required for robots and the supply chains involved might cause environmental problems either locally (e.g. mines in areas with minimal environmental protections) or globally (e.g. the carbon footprint of the supply chain). While rich markets might benefit from access to interactive robotics, the environmental costs will be imposed on poorer nations and peoples. We will also need to consider the disposal of obsolete robots (avoiding a market for robot's dependent on obsolescence would be preferable). We see that obsolete computers and mobile phones have had a massive impact on environments and livelihoods, e.g. in Ghana e-waste is polluting the environment around Accra [382]. Such concerns are not always raised when novel technologies initially emerge.

There is also the issue of power required for robots as for many new technologies. For instance, the cryptocurrency Bitcoin is now known to require huge amounts of energy [383]. The generation of the power required by interactive robots could further entrench our collective dependency on fossil fuels. Again, **the costs of this are largely borne by people in poorer nations and by non-human animals**.

c. Consumerism and Well-Being

Given the facts known about the environmental crisis (in terms of climate and biodiversity loss) and the role contemporary consumerism plays in generating that, more research will be required on the development of markets for interactive robotics as status symbols or *consumer*

---

[378] "Robots-in-ecology are robot technologies used for environmental research applications; including uses of general robotics technologies for such research" (van Wynsberghe & Donhauser: 2018: 1783). These would include drones and rovers used to monitor pollution, poaching, and deforestation and so on.
[379] "Robots-for-ecology are here conceived as those service robots used in environmental research that are invented and designed for the express purpose of carrying out more highly specialized research tasks with maximum efficiency" (van Wynsberghe & Donhauser, 2018: 1786). These might fill ecosystemic niches or perform things like maintenance or pest control.
[380] Ecobots are 'ecologically functional robots,' (van Wynsberghe & Donhauser, 2018: 1788) that perform ecological roles, such as hunting down and destroying invasive species.

[381] van Wynsberghe, A., & Donhauser, J. "The Dawning of the Ethics of Environmental Robots". Science and Engineering Ethics, 24 (2018): 1777
[382] Beaumont, P. (2019, April 24). Rotten eggs: e-waste from Europe poisons Ghana's food chain. The Guardian; Petrlik, J., Adu-Kumi, B., Hogarh, J., Akortia, E., Kuepouo, G., Behnisch, P., … DiGangi, J. (2019). Persistent Organic Pollutants (POPs) in Eggs: Report for Africa.
[383] Hern, A. (2018, January 17). Bitcoin's energy usage is huge – we can't afford to ignore it. The Guardian; Lee, T. B. (2018, May 17). New study quantifies bitcoin's ludicrous energy consumption. Retrieved 23 April 2019, from Ars Technica website: https://arstechnica.com/tech-policy/2018/05/new-study-quantifies-bitcoins-ludicrous-energy-consumption/

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

goods (goods to be consumed, or devoured, rather than durable, fabricated goods,[384]. This raises questions regarding the flourishing life and the role goods play in achieving it. While interactive robotics can play an important role in this, **the creation of new needs (i.e. the need to purchase an interactive robot) on the part of the consumer might have negative environmental consequences**. If interactive robots are developed within a paradigm that assumes that well-being is achieved through the satisfaction of individual preferences, it will be very difficult to place limits on the wants of individuals. However, there is no reason to assume that well-being can only be achieved through the satisfaction of individual preferences – both Aristotelian and Epicurean accounts of well-being suggest that there are limits to the good required for human flourishing.

Furthermore, when discussing such issues, the distinction between consumer preferences and citizen judgements needs to be remembered. While many potential users of interactive robotics might *prefer* cheaper robots that impose greater environmental costs, the judgements of **citizens** (supposedly based on reasoned argument) **might prefer a system that does not further damage the natural environment.** If the aim of inclusive robotics is indeed a *better* society, it is vital that the environmental cost is minimised.

d. Topics of Further Research

Against the backdrop of the above, there are a number of areas that need to be investigated in greater detail.

- *Energy.* How to minimise the energy requirements of interactive robotics? At this early stage in the development of a society making extensive use of interactive robotics, the default energy requirements for those robots is yet to be established. This will need to be considered both in terms of individual robots and on a societal level. Stringent regulations governing the energy requirements of interactive robotics at the design stage and at the use stage provide one potential solution.
- *Resource Use.* How to ensure that interactive robots do not require enormous resources (other than energy) to develop and use (particularly at a societal level)? An interactive robotics market requiring built-in obsolescence (i.e. getting users to upgrade their robots repeatedly) is likely to be extremely resource intensive. Again, stringent regulations designed to minimise resource use, or better ensure that robots do not become obsolete quickly (or at all) might be required.
- *Disposal.* How to ensure that obsolete robots can be safely disposed of without causing further environmental harms? To this end, utilising biodegradable materials as much as possible should be considered. This can be seen as an issue of environmental justice – many of the costs of the resource-heavy lifestyles in the richer parts of the world are exported to the least well-off parts of the world.
- *Moral Hazard.* If robots begin to fill ecological niches, there exists a possibility that people, states or corporations will fail take other actions to mitigate the impact of climate change or biodiversity loss, as robots will fulfil the instrumental role. From an environmental ethics perspective, particularly from the perspective of theories that highlight the inherent value of natural systems and nonhuman animals, this scenario should be avoided. **This is not to say that ecobots should not be used to help restore or rebalance nature, only**

[384] Arendt, H. (1958). The Human Condition: Second Edition. University of Chicago Press.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

that they should not be designed to replace natural processes, nor should they be relied upon to do so.

- *Unrest.* Policymakers should also consider the risk that job-losses arising from automation, including robots, is likely to cause social unrest. Such conditions may make it harder to pursue the sorts of environmental policies that are now required. This is not to claim there exists a direct causal relationship between automation and resistance to environmental policies, but examples such as France's *gilets jaunes* highlight the risks.
- *Nudging.* Interactive robots might also bring environmental benefits. Such robots could be equipped with sensors so extremely localised environmental data can be gathered (e.g. air quality, species diversity, etc). **Interactive robots could nudge people to behave in environmentally better ways, or indeed report if their behaviour is egregiously environmentally harmful (subject to privacy norms, etc.).**

e. Conclusion

Policymakers have the opportunity to shape the development of interactive robotics. Interactive robotics are often framed in terms of opportunities and risks to individuals and societies, without reference to the natural environment. The danger now is that the development of interactive robotics – the economies, markets, institutions, and practices – takes place without consideration of environmental costs. This would risk putting us on a technological treadmill – where robots are required – despite enormous environmental harms. Alternatively, **if environmental values are built into the economies, markets, institutions and practices relating to interactive robotics, the environmental harms can be minimised, and interactive robotics can serve as a model for other novel technologies.**

7. Assessment of corporate social responsibility (CSR) and inequalities

1. Corporate social responsibility (CSR) in the European Union for Global Sustainability - also affecting robotics

In the international scenario, there are, at this time, various interlinked fronts: the United Nations Sustainable Development Goals[385], the United Nations Guiding Principles on Business and Human Rights, the OECD Responsible Business Conduct initiative or the EU Corporate Social Responsibility Strategy. All of these objectives and means push for a better world (one with values on which we can build a more cohesive society) and its implementation deserves careful attention.

The EU adopted in June 2017 the new European Consensus on Development based on the "5 Ps" of the 2030 Agenda: People, Planet, Prosperity, Peace, and Partnership, and has been systematically integrating social, economic and environmental dimensions in its initiatives since

[385] UNITED NATIONS: "Transforming our world: The 2030 Agenda for Sustainable Development", Resolution 70/1 adopted by the General Assembly on 25 September 2015.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

then. More recently, in January 2019, the EU Commission has issued a Reflection Paper "Towards sustainable Europe 2030" to show its significant commitment[386].

As part of the EU Commission's work on supporting the implementation of the United Nations 2030 Agenda, a Working document has been published during March 2019 where stocktaking is made of all the progress made in the European Union since 2011 on Corporate Social Responsibility[387].

Corporate social responsibility (CSR) is a dynamic concept that mostly depends on circumstances of time and place. It can be expressed in different manners. A pure concept of CSR implies going beyond the binding regulation and its application. There are different standards available to voluntarily adhere to, in order to prove good entrepreneurial behaviour; depending on the level of commitment wanted. Often the standards evolve progressively, requiring the subsequent adoption of a higher one. Once a certain level is commonly accepted and generally spread in the public perception, it is easier to turn the minimum standards into mandatory legislation.

It is said that "companies can become socially responsible: - by integrating social, environmental, ethical, consumer, and human rights concerns into their business strategy and operations; - following the law"[388]. However, in our view, the latter is clearly binding if an imperative rule is adopted, because only rules with a promotional nature would allow voluntary compliance (as it is the case with tax incentives for research and innovation with the potential of emerging technologies, or for smart green investment and the circular economy). So, **public authorities can play a supporting role through a smart mix of voluntary policy measures and, where necessary, complementary regulation, as businesses should integrate Sustainable Development Goals in their operations. Responsible Business Conduct, a.o. is an important horizontal enabler for sustainability change if "geared towards an innovative green, inclusive and socially just economic transition"** [389].

Irrespective of whether it is labelled 'Corporate Social Responsibility', 'Responsible Business Conduct', 'Business and Human Rights', 'Sustainable Development Goals' or all four together, action is needed. **This is also applicable to robotics**.

It is quite obvious that companies do not only offer products or services, and create jobs and opportunities, but also have impacts on society in terms of working conditions, human rights, health, environment, innovation, education, and training. The robotization affects all these areas. The European Union Strategy defined broadly the concept of Corporate Social Responsibility: "the responsibility of enterprises for their impacts on society". In a more detailed

[386]See https://ec.europa.eu/europeaid/sites/devco/files/european-consensus-on-development-final-20170626_en.pdf and https://ec.europa.eu/commission/files/reflection-paper-towards-sustainable-europe_en

[387] In 2011, the Commission adopted its renewed strategy for CSR, which combines horizontal approaches to promote CSR with more specific approaches for individual sectors and policy areas. It aimed to align European and global approaches to CSR. EUROPEAN COMMISSION**:** "Corporate Social Responsibility, Responsible Business Conduct, and Business & Human Rights: Overview of Progress", *Commission Staff Working Document*, Brussels, 20.3.2019 [SWD (2019) 143 final], p. 4. Commission services issued in 2014 a Compendium on national public policies in the area of CSR. See https://ec.europa.eu/digital-single-market/en/news/corporate-social-responsibility-national-public-policies- european-union-compendium-2014

[388] See http://ec.europa.eu/growth/industry/corporate-social-responsibility/

[389] EUROPEAN COMMISSION: *Reflection paper…*, cit., p.14.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

perspective, *ad intra*, this means that **companies "should have in place a process to integrate social, environmental, ethical, human rights and consumer concerns into their business operations and core strategy in close collaboration with their stakeholders**, with the aim of maximising the creation of shared value for their owners/shareholders and civil society at large and identifying, preventing and mitigating possible adverse impacts". How to consider all these factors and agents' views, where to add them, and which purposes should be followed, is key.

In addition, *ad extra*, the Non-Financial Reporting Directive[390] requires certain largest EU companies, as well as other public-interest entities, to disclose their business model, policies, outcomes, principal risks and risk management, and key performance indicators relevant to the particular business. **They must report on the due diligence process that they implement with regard to environmental, social and employee issues, respect for human rights, and bribery and anti-corruption**[391]. **The influence of robotics on these topics could be also voluntarily explained, as a voluntary action showing CSR.**

Apart from that, the use of Green Public Procurement or Socially Responsible **Public Procurement** (to be updated in 2019 to reflect the 2014 Directives) in public purchasing can create additional market opportunities for sustainable products, promote supply chain due diligence and encourage the market to shift towards more environmentally friendly and socially responsible solutions. Here, **mention could be made to responsible robotics** (both in the creation and implementation phase).

2. Inclusive Robotics in relation with Responsible Research and Innovation

It is wise to align *inclusive robotics* within the paradigm of *Responsible Research and Innovation* (RRI). Although RRI was coined a decade ago, it has taken on a leading role due to its inclusion in the programme *Science with and for Society* (SWAFS), promoted by the European Commission in the research strategy framework of Horizon 2020.

The main goal of RRI strategy is to reduce the divide existing between the scientific community and society, encouraging different stakeholders (entities from civil society, the educational and scientific communities, people in charge of policy and the business and industrial sectors) to work together during the entire research and innovation process. Part of society looks with concern at

[390] Directive 2014/95/EU of the European Parliament and of the Council of 22 October 2014 amending Directive 2013/34/EU as regards disclosure of non-financial and diversity information by certain large undertakings and groups; OJ L 330, 15.11.2014, p. 1–9. Small and medium-sized enterprises often have a naturally responsible approach to business (due to their close relations with employees, local community, and partners) through informal and intuitive processes. The EU has developed CSR handbooks and manuals for them. EUROPEAN COMMISSION: "Corporate Social Responsibility…", cit., p. 59.

[391] Companies were reporting for the first time in 2018 covering financial year 2017. In mid-2019 a first review of the Directive will be published, as part of a broader Fitness Check exercise on the overall EU framework for company reporting. The Commission's Non-Binding Guidelines for reporting non-financial information, published in July 2017 will be updated by mid- 2019, to integrate the recommendations of the Task Force on Climate-related Financial Disclosures established by the G20's Financial Stability Board. The European Financial Reporting Advisory Group hosts the European Corporate Reporting Laboratory to identify and document innovations in reporting practices, where work is envisaged in the area of environmental accounting. EUROPEAN COMMISSION: "Corporate Social Responsibility…", cit., pp. 28-29.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

the dangers of a technified society. This feeling has to do with the alienating nature of many technical systems. They influence and condition our life, but we practically cannot influence them. As such, starting from mechanisms for cooperation between different actors, **it is possible to better align the research process and its results with the values, needs and expectations of current society**.

RRI can be understood thus as an **effort to justify innovation not on grounds of uncritical, or taken for granted macro-economic assumptions, but on the basis of societally-beneficial objectives, or challenges, as openly defined and debated by a plurality of societal actors**. As such, RRI-based EU policy aims to introduce "broader foresight and impact assessments for new technologies, beyond their anticipated market-benefits and risks" [392]. RRI's radical rhetoric on openness and socialization regarding techno-industrial innovation processes has been claimed to ultimately reflect four **fundamental principles of scientific governance: anticipation, reflexivity, deliberation and responsiveness**[393].

The theoretical framework for RRI is which is called "post-normal science" [394], a science characterized by uncertainty of facts, disputed values, enormous challenges and urgent decisions. The challenges that techno-science tackles produce important disagreements between the experts and involve the individual and collective decision-making, and the assumption of risks in contexts of uncertainty ("society of risk"). They require, therefore, ethical debate, public deliberation, and policies (social control). Hence, scientific research, as in robotics, should favour social participation in public debate, promoting information, transparency and the intervention of non-experts in deliberation and decision-making.

RRI acknowledges the transformative power of innovation to create futures, that innovations are often socially and politically constituted [395] and that they embed values[396] .

The RRI approach aims to make society a participant in science and innovation from its most initial phases, involving all actors and aligning the processes and results of the research with the values, needs and expectations of society. In practice, the RRI concept implies designing and implementing research and innovation policies that pay special attention to 6 interrelated agendas: ethics, gender equality, governmental agreements, open access to scientific information, public participation and scientific education.

A pluralistic approach with these characteristics could **lessen the risks and make the most of the opportunities from innovations in robotics, including the diversity of agents and values that will be involved in its development in the decision-making process**. The set of ideas and initiatives brought together in the paradigm of RRI take us to the proposal of *inclusion* as a central value within these processes as well as giving more positive weight to other values related to it.

---

[392] von Schomberg, R. (2013). "A Vision of Responsible Research and Innovation?, in R. Owen, J. Bessant and M. Heintz (eds.), Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society, Chichester, UK: Wiley, 51-74.
[393] Stilgoe, J., Owen, R., Macnaghten, P. "Developing a framework for responsible innovation", Research Policy 42 (2013), 1568-1580.
[394] Funtowicz, S. O., and Ravetz, J. R. "Post-normal science: A new science for new times", Scientific European, October 1990, 20-22
[395] Winner, L. (1980). "Do Artifacts Have Politics?" Daedalus, Vol. 109, No. 1: 121-136. Reprinted in Donald A. MacKenzie and Judy Wajcman (eds.), The Social Shaping of Technology, London: Open University Press, 1985; second edition 1999.
[396] van den Hoven, M. J., Lokhorst, G. J. C. and van de Poel, I. (2012) "Engineering and the problem of moral overload", Science and Engineering Ethics, (2012): 18, 1–13

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

As RRI points out, inclusive and participatory methodologies should be used in all the stages during the research and innovation processes. To locate and analyse the evaluation mechanisms that take place throughout the innovation process of interactive robotics, aimed at making it easier to satisfy values and criteria such as *inclusion,* is an important element to advance in the road towards an inclusive robotics. This movement implies "bringing social and moral values" from the citizenry to processes that traditionally have acted independently.

Public engagement is essential for researchers to reach better ethical solutions on difficult issues and to ensure public acceptance and trust of such research to identify ethical solutions acceptable to the public. However, there is a need to go beyond participation towards *deliberation*. Deliberation is a condition for RRI, mainly focused on the question of inclusiveness, fair cooperation and rational decision-making, and could take advantage of the precautionary principle, one possible understanding of responsibility, to be focused on the controversies relative to the future. The precautionary principle has the potential to structure the meta-deliberation of minds at the intersection of ethics, politics, sciences and technologies[397]. Thus, public deliberation, with good and plural information and transparency, provides knowledge (epistemic virtues), changes attitudes, and increases the perceived legitimacy of institutions involved in the deliberation process. As we have said, participatory methodologies are crucial for inclusiveness in robotics research and development.

However, it remains open some questions:

- What is "the public"?

  In general, it would be the society as a whole but the level of public interest in the issue under discussion although even a low response rate but with people representing a broad demographic range may nevertheless provide insights about the public's informed views that would otherwise be left to guesswork. Moreover, democracies cannot expect everyone to be interested in every issue and the proper audience for a deliberation may be the portion of the public most capable of being interested in the issue ("attentive public"). This may be the audience that could become activated in a controversy over the issue: civic organizations, trade unions, civil servants, policy makers, opinion leaders and stakeholders concerned in the introduction and development of robotics.

- How to reduce public distrust or outrage?

One of the main problems of developing and introducing new technologies as robotics is the distrust and even rejection in society. Directly communicating information about projects and their rationale as well as public engagement may reduce that risk.

- What channels should be used to that "robotics literacy"?

As a widespread communication is mandatory and unavoidable, the use of websites with clear explanations, YouTube channels, Twitter and social network feeds, movies, TED conferences, TV dialogues and other outreach efforts are some ways to do it. Open science and science education is a key issue of RRI.

- How to concrete public deliberation?

[397] Reber, B. "RRI as the inheritor of deliberative democracy and the precautionary principle", Journal of Responsible Innovation, 5 (2018): 38-64.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

There exist numerous methods of public deliberation that can be utilised under different circumstances and constraints or for different purposes (see http://www.participedia.net/). Some examples are the participatory budgets, the community development plans, the citizen juries, the consensus conferences, the parliament consultation or the deliberative councils (as the Comision Nationale du Débat Public in France or the Bureau d'audiences publiques sur l'environnement in Quebec). At this point, Ethical Review Boards (national, institutional or sectoral) could play an important role in order to promote such a public deliberation.

- What about minorities and future generations viewpoints?

A huge limitation of deliberative processes is about inequalities, asymmetries and disadvantages. Thus the design of participatory tools should be especially careful to cope with future generations and minority viewpoints[398], as those of vulnerable groups as children or disabled people.

- The problem of biases and, concretely, gender biases.

The under-representation of women in research and management must be addressed. RRI clearly stands that the gender dimension must be integrated in research and innovation content.

- Does ethics prevent research development and innovation?

To respect fundamental rights (as the Convention on the Rights of Persons with Disabilities or the Convention on the Rights of the Child) and the highest ethical standards both ensure the acceptability of research and innovation outcomes. **Ethics should not be perceived as a constraint to research and innovation, but rather as a way of ensuring high quality results (excellence).**

- Which good practices should be promoted in robotics research?

Good practices establish procedures and guidelines for desirable conduct pointing out and censuring mispractices and corruption. They build trust and public engagement to overcome barriers to robotics research. Some examples of good practices are: Supervision of research personnel in training; preparation of written research protocols where methodology is explained; recording, documentation, storage and custody of data and materials resulting from the research; clear terms for research projects financed by the industry and public organisations (property rights and compensations); management criteria for conflicts of interest; transparent practices for publication, protection and dissemination.

- How to articulate the benefits and added value of RRI in robotics from a commercial perspective?

For companies, especially small and medium, RRI incurs a potential financial cost for which the added value would need to be demonstrated, particularly to investors. Some companies had very little capital and for them the priority was that they were commercially viable. They wanted to know what the return on an RRI investment would be, given the hard-commercial realities they faced.

3. Women and Robotics

The regulation of robotics should be based on the following principles: equity, reliability,

[398] Sunstein, C.R. (2005). Why Societies Need Dissent, Harvard University Press.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

privacy, inclusiveness, transparency and responsibility[399].

From a CSR perspective, there are two main aspects to consider in the relation between robotics and diversity: the treatment of women in robotics, and how far robotics advances or hinders women's position in the workplace and the business world[400].

a. The treatment of women in robotics

One area that should be particularly highlighted in this first aspect is the data "processing"; that is, the data that are used as input for the robots, and in the area of artificial intelligence in general.

The means and the criteria used to input data in a robot or in any element involving AI – although they may be valid– are essential for ensuring that the results are fair and offer equal opportunities and treatment to everybody. If we input data in a search for job candidates with certain characteristics in which men tended to have more expertise or historically better results, women would be affected by negative discrimination.

The use of algorithms, even those created with objective data poses a potential danger of maintaining or even widening the gender gap as even objective data will be shaped by the history of human thought, with its inherent biases and prejudices – again, certainly unconscious or unintentional. The risk does not derive from innovation, technology or the new era, but from the past.

i. The gender gap

In spite of the progress in the robotics sector in both the business and academic spheres, there continues to be a significant gender divide. The source of this gap is the fact that there are far fewer women working in robotics and AI, both in research and in the use of these resources. On this point it is worth noting that women still tend to self-exclude themselves from the fields of engineering and mathematics[401].

The experts warn that **the lack of women in the robotics industry in the threefold areas of research, programming and enterprise is a problem** for the actual development and evolution of artificial intelligence[402].

[399] González Espejo, M.J., "Desayuno INNOTECH ¿Deben ponerse límites legales y/o éticos a la Inteligencia Artificial?". CEU, https://institutodeinnovacionlegal.com/events/limites-inteligencia-artificial/.
[400] AMESTI MENDIZABAL, M.C. "Women and Robotics", Inclusive Robotics for a Better Society. Selected papers from INBOTS Conference 2018, 16-18 October 2018, Pisa, Italy, Springer (forthcoming 2019).
[401]EDUCAWEB, "La formación, clave para afrontar la era robótica del mundo laboral, https://www.educaweb.com/noticia/2018/09/26/formacion-clave-afrontar-era-robotica-18567/; UNIVERSIA, La diversidad en la industria robótica, un reto en común UNLEASH, http://noticias.universia.es/practicas-empleo/noticia/2018/09/26/1161933/diversidad-industria-robotica-reto-comun.html.
[402] ORELLANA, M.C. . "Barbie Ingeniero de robótica", en http://abcblogs.abc.es/tareas-pendientes/2018/09/26/barbie-ingeniero-de-robotica/; UNIVERSIA, "La diversidad en la industria robótica, un reto en común", UNLEASH, http://noticias.universia.es/practicas-empleo/noticia/2018/09/26/1161933/diversidad-industria-robotica-reto-comun.html.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

ii. Robots are genderless

Robots may have a male (android) or female name (gynoid), but it should be stressed that they have no gender. **Robots are not classified into female or male according to the functions they perform, which is very appropriate for improving equality.** The trend is for robots to be neutral.

b. How far does robotics advance or hinder women's position in the business world?

With regard to the second question, namely the consequences of robotics on women's jobs, or how robotics will influence women's employment, most experts believe that robotics will improve the position of women in two regards.

First – and perhaps most importantly – physical strength will cease to be a differentiating element in favour of men for certain types of jobs. Robots are already doing the heavy work, and women now have access to operating the robotic machinery to carry out these functions, a situation that will become even more commonplace in the future.

Second, the most highly demanded jobs in the near future will be data analysts, software developers, app creators and developers, researchers and experts in e-commerce and social networks. **It will therefore be important in this new era – the age of robotics and AI – to train women to ensure they retain their place in the labour market.**

According to the experts, the jobs that are currently done by women have a far lower risk of disappearing than those done by men (except in Japan), from which they deduce that women will be less affected than men by the potential threat to jobs from robotization.

Robotics is not sexist and is not in itself discriminatory.

c. Conclusions

In conclusion, robotics does not marginalise women, and it can initially be assumed that it will not have a negative effect on them, but rather the reverse. However, it is essential that women should not be left behind, and this can be avoided by placing special emphasis on training and education to ensure that their jobs will be safeguarded in this new era[403].

Corporate directors and company managers will play an important role in all this. They will need to work on advancing and reinforcing corporate social responsibility in the governance and administration of their organisations. The support for gender diversity and the observance of women's rights in the workplace are essential aspects that inform the concept of this corporate social responsibility that must be present in the approach to the new challenges posed by robotics and artificial intelligence today and in the future.

4. Minors and Robotics

Minors can interact with robots in various fields (such as, entertainment, care or training). The Toys Directive and rules on video games are useful in the entertainment and leisure ambit. For care, the rules applicable to elderly and disabled can inspire future regulation,

[403] EDUCAWEB, "La formación, clave para afrontar la era robótica del mundo laboral", https://www.educaweb.com/noticia/2018/09/26/formacion-clave-afrontar-era-robotica-18567/

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

apart from considering the specific ones related to healthcare (e.g. surgical interventions, rehabilitation exoskeletons or data management).

Additionally, a differentiation can be made among the **types of robots**, depending on the share of human control: fully controlled by human being, partially controlled (at certain point) or not controllable.

From a relational perspective, it is important to observe **when** the robot is interacting with minors, if it keeps relation with other robots, or establishes relations between users, under certain ages or robots.

Minors are subjects in need of special protection that should be granted through hard law. At EU level, the regulatory framework affected would include:

- the Charter of Fundamental Rights of the Chile, where a reference to the digital environment and robotics could be added;
- Directive 2009/136 on the processing of personal data and the protection of privacy in electronic communications, where references could be made to the Directive 2000/31 on the Information Society Services;
- -Regulation 2120/2015 on open Internet. In its article 3 the exceptions to restrict free access to minors could be developed.

Under article 29 of the Directive 95746 (WG 29) some measures to protect minors in the digital environment have been proposed: information and awareness raising, separate groups by age, technologies for verification of access and use, among others. These can be applicable to interactive robotics as well.

Directive 2009/48/EC on the safety of toys affects the use of toys by children under 14 years. It includes video games. The period from 14 until 18 years for virtual games –out of the scope of this Directive, should be taken into account and regulated.

Regarding privacy, as a right regulated in the RGPD 679/2016 that affects the collection of data, among others, consent is required from the age of 16. This should be assessed and controlled in the field of robotics. As problems of espionage can appear (affecting the manufacturer, distributor or seller), a restrictive interpretation of the provision of consent can be defended. **Special protection is deserved, because of the reduced awareness of possible risks, during several moments**:

- Before using the toy, it is necessary to inform users about the fact that personal data are collected (what and how) and will be used or shared (by whom, for what). Information can be provided through labelling and description of contents, showing licenses, certifications or registration. Parents' or guardians' consent is required for children under age of 16. In the moment of access, the user age must be verified, and eventually apply restrictions.

  Perhaps there should be regulations to ensure that the toy can be used (with somewhat limited functionality) without collecting personal data. Informed consent in this context could be inadequate to protect society against data mining by large corporations etc. People tend, in reality, not to read the terms and conditions and consent to whatever the company requires. Stronger regulation might be required.

- During the life of the toy, it is important to maintain access control and safety mechanisms (mandatory signalling and content filters, qualification of sensitive data and geolocation).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

- In the case of transfer, destruction or abandonment, the data in the robot toy and the ones transferred to the company should be deleted, similarly as with the right to forget.

When promoting the use of robotics, quantity of users cannot take precedence over the quality, to provide services to minors. **Regulation should take into account the psychophysical development, in accordance with parental control, emotional and social skills**. Access control must be checked, e.g. with codes or messages to parents or persons in charge of the minor. Every activity (its contents, language or timing) and its development (predictable or not) must be controlled. Implementing protection mechanisms for minors must be considered an **added value to the services rendered**. There must be a **fast procedure for reporting illegal or harmful content** (e.g. violence or other that may affect the child's own health), for the **detection and removal**. Sound and visual signals may be an option.

Regarding training, a **correct use of robotics by children** should focus on simple and basic tasks (tutorials, facilitators…), not on non-cognitive ones such as creativity, teamwork, social skills and confidence.

If the robot's activity infringes the right of a minor, its temporary suspension (even through remote control), its deposit or seizure, and eventually the opening of a sanction procedure should be regulated.

The **degree of linkage of the child with the robot** should be even more controlled in the (health) care and education area (as educational robots help children with autism[404] or learning disabilities). Again, control is decisive according to the age and real possibilities of interaction of the child with the robot.

The possible access of a minor to autonomous vehicles also raises concerns and there should be regulations on this.

Currently, the mentions to minors in standards are quite vague and do not favour their proper protection. Issues on prevention or limits still can be thoroughly developed. The main actors in the robotics and AI environment should move forward, together with parents and guardians. With respect to the age of access, we doubt that the current disparity can be maintained *sine die*. **Minimum ages** should be set in the interest of minors.

5. Ways to protect the right to equality in case of discriminatory impacts caused by Robotics and Artificial Intelligence

The future of our society depends on finding the right balance between technological development and human rights protection. The international framework for protecting human rights could be applied against discriminatory impacts caused by robotics and AI. States should take targeted measures to ensure that discrimination in the exercise of rights is eliminated, and regularly assess whether the measures chosen are effective in practice.

---

[404] See DREAM EU Project No. 611391, D7.2 Ethics white book for child-robot interaction for children with ASD. Available at https://www.dream2020.eu/deliverables/

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

The assumption that the design, deploying machine learning systems, development and use of responsible technology is a shared responsibility by the States and private actors has led to the setting of Codes of Ethics and standards under the auspices of several organizations, such as the Toronto Declaration.

However, **the more effective way for promoting a solid protection for future robotics and AI concerns, relies at a firm level, and specially, through the design of a compliance system** in line with other existing compliance systems in corporate governance and according to the principles and values of Corporate Social Responsibility.

In practice States will often rely on private contractors to design and implement these technologies in a public context. In such cases, **States must not relinquish their own obligations around preventing and ensuring accountability and redress for discrimination and other human rights harms in delivery of services.** The compliance program proposed relies on two main pillars: the shared responsibility by States and private sector and on principles suited for each kind of actor[405].

The main task of future regulation in the development of robotics and artificial intelligence is that it be put at the service of the human being, promoting an inclusive and equitable social progress, and also avoiding supremacist positions of some individuals against others and situations of abuse and manipulation.

Notwithstanding the need to regulate the future impacts of robotics and artificial intelligence, the asymmetries in knowledge and the awareness of its impacts among society, experts, government authorities and industry; the objective of not reducing the innovation and competitiveness of a sector that is still incipient in many States; the lack of flexibility and agility of regulatory mechanisms and processes; and the different capacity of action of the affected interest groups, have led to a wide body of statements by experts and industry sectors, which represent a first step towards the establishment of agreed ethical principles and minimum guidelines for the future regulation of the sector.

Some voluntary initiatives have arisen so far, and we can find some basic principles of a future due diligence system that could be articulated by the private sector as complementary to the regulation.

a. Some problems posed by Big data, Robotics and AI

Several problems may arise: the digital divide; the use of covert artificial intelligence systems that increase the risk of manipulation and control of the human being; the limitations of the generation, collection, analysis and use of data; the imbalance between supply and demand among robotics and AI professionals working for the commercial sector and the social sector (as the former has a greater capacity to attract talent, which has consequences not only for research but also for the resulting applications); and a given economic model can condition the future use of the data and the potential of new 'species' models (and the data they use) to penetrate the economic ecosystem.

[405] Helena Ancos, "AI And Discrimination. A Proposal For A Compliance System For Protecting Privacy And Equality", Inclusive Robotics for a Better Society. Selected papers from INBOTS Conference 2018, 16-18 October, 2018, Pisa, Italy, Inclusive Robotics for a Better Society. Selected papers from INBOTS Conference 2018, 16-18 October, 2018, Pisa, Italy, Springer (forthcoming 2019).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

b. Towards a governance of Artificial Intelligence and Robotics

The challenges posed by artificial intelligence and robotics are so broad in their thematic coverage, so complex in their implications and so unpredictable in the depth of their impacts that many of the existing formal and informal institutions are not adequate to address these challenges.

Regarding the impact on privacy, the application of automation to decision-making in the field of justice or health, the impact on employment or the interaction of robots and androids with human beings, **a certain degree of institutional innovation is necessary** to guarantee the governance of these technologies in society and provide adequate accountability.

While **international human rights law** has been invoked as a universally accepted framework for considering, evaluating, controlling and, ultimately, correcting the impacts of artificial intelligence and robotics on individuals and society, its limited effectiveness and flexibility to adapt to a rapidly changing system, **requires the cooperation of other protection mechanisms**.

In this sense, regulatory loopholes not only derive from a new playing field where the law has not yet entered, but also from the complexity of **an area that moves much faster than the capacity to generate laws that cover them**. This dynamism, its complexity and its lack of coverage, have attracted the necessary concurrence of the private sector for the construction of the future governance of the digital society.

The contribution of the private sector to the debate has borne fruit in a wide panoply of statements of principles that cover both the research phase, the design of algorithms and marketing, and its exploitation. From the Asilomar Declaration for an ethical investigation in Artificial Intelligence, through the FAT / ML Principles for Responsible Algorithms, the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems, the Montreal Declaration for responsible development of artificial intelligence, up to the recent Guide to ethical principles in AI prepared by the group of high experts of the European Commission, a corpus of soft law has been established with a minimum consensus on the principles that should govern the life cycle of artificial intelligence.

This **proliferation of codes of conduct is a clear evidence that each application affects a multitude of rights, and sometimes in a contradictory way**. The necessary participation of the private sector (researchers, engineers, systems developers) and representatives of civil society (experts from different areas, users, non-governmental organizations) will contribute to the development of inclusive proposals and throughout the life cycle of AI.

The codes of conduct offer a consensus framework adapted to the particularities of the developments of each system or service. **However, they present limitations derived from their voluntariness and the absence of sanctions** and in some cases, an overly technical approach that avoids ruling on legal consequences.

As common characteristics to these statements, we can highlight the following:

1. The concern about the possible negative impacts of the development of an artificial supra-intelligence that may exceed the limits of human control, on the one hand, and the possible

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

harmful uses of particular developments, such as lethal weapons or the indiscriminate use of data beyond the knowledge and human consent.

2. On the other hand, the principles contained in them are not presented in a hierarchical way. Its reading must be done in a complementary and inclusive manner, in a consistent interpretation where the limits of the application of a principle are traced by the scope of another principle.

3. Basically, the interest of voluntary initiatives rests on:

i. Strengthening the legitimacy of proposals for a responsible AI, by incorporating the initiatives of private actors;

ii. They collectively arbitrate ethical and social controversies about AI by conceiving themselves as open, reviewable and adaptable processes according to the evolution of knowledge and techniques, and feedback on the use of artificial intelligence in society.

iii. Improving the quality of thinking about responsible AI by presenting the vision of experts and specialized groups.

4. Some of them represent a more advanced stage.

The problems raised by AI and robotics are of such magnitude that self-regulation will not be enough. The important issues raised could, without proper regulation, question fundamental values on which our society is based. Even in a constantly evolving technological context, regulation already needs to be adapted in many aspects. One of them is the protection of personal data and, more specifically, that of the notion of anonymous information. The question of the anonymization of data becomes increasingly complex; some even claim that it is a myth, especially in the context of the exponential dissemination of data of all kinds and increasingly effective cross-checking techniques.

In the same way, the evolution experienced in the management of personal data by citizens, where there has been a shift from a posteriori consent and the Robinson lists to prior consent and the conduct of privacy impact assessments with doubtful efficacy, highlight that there are powerful reasons to maintain the control of citizens over their personal information and consent management, in information management processes and the need to adopt a preventive approach.

The preventive approach or due diligence can be developed in a range of tools throughout the life cycle of the product or service to help developers and users in order to preserve due guarantees on the fundamental rights involved. In addition, the establishment of a due diligence process in **both public and private entities** that work with artificial intelligence and robotics in its different stages of the product, will allow for the **implementation of a co-responsibility system in the management of risks and the establishment of correction and remedy measures.**

As usual, before initiating a due diligence process in any organization, an unambiguous definition of a corporate Human Rights policy must be made. This corporate policy may well be based on the principles set out and agreed upon in the Asilomar statements, responsible algorithms, Montreal or the Working Document of the High-Level Expert Group on Artificial Intelligence.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

The advantages of incorporating these statements into corporate policies are that they are already based on specialized advice, gathering the opinion of various stakeholders, have a mechanism for continuous updating and include operational procedures to facilitate a corporate risk management system. Precisely, the wide range of voluntary declarations allows the evaluation of the impacts of corporate practices on human rights throughout the entire value chain and allows identifying and documenting the risks posed by the design, development and implementation. The principle of prevention on which all these statements are based makes it possible to identify the most important problems in a preventive way and to find operational solutions and opportunities for improvement from the design stage of a project and to take into account the human rights problems in an upward direction throughout the life cycle of products and services.

Based on the aforementioned impact assessment, the classification of possible measures to be implemented by the organization will lead to the design of a road map to integrate different change projects aimed at reducing future impacts on human rights and mitigating risks by incorporating formulas for the protection of rights in management systems and decision-making processes.

The measures of transparency established in these statements should contribute to giving **visibility not only to the risks but also to the sequencing of responses to them, also facilitating the establishment of monitoring and follow-up systems**, making changes in the way of proposing some corporate activities; increasing the awareness of workers and users about particular risks and increasing the supervision of the media and other sources of information on human rights, among others.

Transparency in turn allows for accountability as a condition for social acceptance. In the rendering of accounts, the organization must not only inform the public of the positive aspects but also of the negative impacts including the establishment of systems of claim and mediation destined to repair the damage already caused if any.

Transparency practices allow for **increasing citizens' awareness and education, empowering them and encouraging their critical thinking as well as increasing informed automation decisions**. On the other hand, transparency practices would provide confidence in a context where the use of algorithms for personalization and assistance in decision-making have already given rise to contrasted cases of inequality and discrimination and to the fear that existing social differences will aggravate. The opacity of machine learning systems is largely due to the speed and dimension in which data is evolving, which is particularly crucial in deep learning cases. At this stage, the accountability of systems based on machine learning is therefore a real scientific challenge, which is creating tension between the need for explanations, intellectual property and trade secrets, security and public order, and the interests in innovation.

For its part, the application of the principles of **prevention, mitigation, and protection** is relevant at a time when the principles and process contained in the Guiding Principles of Business and Human Rights are beginning to be widely understood by companies, also helped for the development of national business and human rights frameworks. But together with the establishment of due diligence processes by public and private actors, the involvement of fundamental rights invokes the **role of the State in the establishment of economic and institutional incentives** (for example, in public procurement mechanisms that encourage the

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

use of responsible technologies) **for systems that respect human rights in the deployment of artificial intelligence and robotics. Along with this, the establishment of certification programs or labels on responsible artificial intelligence and robotics**, taking as background the programs provided in the European Commission Regulation on the protection of personal data. Companies or organizations that deserve this recognition based on objective criteria could benefit from a very significant and reassuring competitive advantage for people (consumers, customers, employees, service users, etc.).

8. Conclusions and recommendations

In order to determine whether robots might be deemed agents, thence subjects rather than objects of the law, eventually bearing responsibility for the harmful consequences deriving from their operation, it is necessary to clarify the fundamental concepts involved both with respect to the philosophical and legal paradigms. By doing so, it may be demonstrated how agency, and moral agency, presuppose the capability of identifying one own end, and actively pursuing them, what today corresponds to a notion of strong autonomy. Existing applications do not display such degree of autonomy and therefore may not, on ontological grounds, be deemed agents, or subject of the law. The fact that they are objects, thence products, causes those who designed them to be primarily responsible for having conceived them in a given way, and programmed them according to chosen criteria. A functional analysis, however, that takes into account applicable law, and analyses the incentives it provides, might give rise to considerations that support the need to reform some specific aspects of the legislation in order to achieve preferable outcomes. Alternative approaches might also suggest treating the machine as a juridical person, yet solely grounded on functional considerations, not as the recognition of some intrinsic subjectivity.

With respect to IP law, it is recommended that a number of new legal reforms be enacted at the European level to bring a measure of clarity to this rapidly developing area. In the areas of copyright and design law, where the EU has substantial competence, new legislative guidance on: (i) whether interactive robot-created works of e.g. music, literature, art can be protected by copyright; and (ii) if so, who should own these works? Similar reforms could be: brought forward by the European Patent Office to clarify the circumstances regarding (i) whether an interactive robot-created invention is patentable or whether a human inventor is required; and (ii) who should be the owner of such a patent (e.g. the consumer/user of the robot, the developer of the robot)?

Regarding the future of work, we have identified four key trends that impact on it: 1) technological progress and automation; 2) international trade and urbanization; 3) a rising diversity of work forms; and 4) population aging.

They are all interrelated. Rules-based and prediction-based technologies are replacing workers in some tasks, while complementing them in others. Prediction-based technologies such as machine learning and other forms of Artificial Intelligence are seeing increasing applications across labour markets, including in human resource management. High-skilled workers have increasingly moved to cities. New work forms are increasingly important in our economies, including rises in part-time and temporary work (such as contractor work and freelancing), and working through online platforms. Ageing workforces are a mediating factor for other key

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

trends: countries with older workforces adopt more robotic technologies, and the workforce of rural regions is ageing more rapidly than in cities.

Technological progress, international trade, and diverse work arrangements have increased productivity, economic wealth and opportunity. However, these forces are also accompanied by several challenges related to the future of work: rising inequality, job reallocation, and skill gaps. While technological advances have not decreased total employment, they have led to increased wage inequality, displaced individual workers from their jobs and hollowed out the skill distribution, leading to job polarization. Economic opportunity has been increasingly concentrated in cities, and in favor of skilled workers. Science, Math, Engineering and Technology (STEM) skills as well as social skills are in high demand in our labour markets.

To try to solve many of these problems, whether we follow the path of regulation, which we can call hard law, or whether we opt for the path of soft law, in both cases we need to ensure transparency and accountability regarding the social and economic costs and benefits. It is a question of moving forward on the basis of 'the principle of caution' as applied to freedom of scientific investigation and, beyond rules of 'technological neutrality' to implement the principle of socially and legally responsible technological innovation

Robotics has many implications for the direct or indirect regulation of working conditions, for rights and obligations, for both the entrepreneur and the worker, and from both a labour and fiscal perspective.

A minimum, albeit necessary, mandatory legal intervention is needed to ensure a balance between entrepreneurial freedom and the function and purpose of labour Law, especially with regard to the protection and guarantee of human work. The starting point for the analysis is the principle of equality and non-discrimination, i.e. the real and effective equality of people, of the groups of people who are vulnerable to automation or robotization. Thus, these cannot become a direct or indirect cause of discrimination; technological neutrality cannot entail a disadvantage or barriers for certain groups of workers. In this respect, and also in the transition phase, as an exceptional, limited and conditional measure, it is necessary to reflect on the quota of humans in companies and reasonable adjustment measures for those groups which are especially vulnerable, including older workers.

We need to adapt labour legislation as it applies to workers' rights and obligations. Both from the perspective of guaranteeing people's employability more than jobs themselves (which will affect the direction taken by legislation relating to the replacement effect, i.e. the replacement of humans by robots) and of guaranteeing the rights of workers in their interaction with robots in the workplace (cooperation between them).

All the expected public policies will probably have an impact on the Financial and Tax Law; but, at the same time, this discipline will condition their actual applicability in practice. The States will have to to allocate rights and responsibilities among human beings for the actions of non-human beings, fighting inter-personal and international inequality, through strengthened cooperation. For example, a system of temporary permits, with a fee to be paid by hauliers to obtain authorisation, would make it possible to manage the speed of automation and obtain funds for the retraining of workers and the provision of assistance to redundant workers. In addition, the insurance companies could cover the technological obsolescence implications for workers, on top on the maintenance of robots.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Robots should not be taxed for the mere fact of being so. Those who defend the creation of a tax on robots do not do so by affirming the existence of (a special electronic) ability to pay, but as an extrafiscal measure, as a way of maintaining the collection despite the loss of jobs, which would even allow for the creation of a universal minimum income. However, Public Finances could promote socially responsible behaviours through the tax expenditure budget, while establishing clear controllable conditions.

As robotic technology becomes more autonomous, it is necessary to identify the values and ethical principles that should regulate the interaction of robotic systems with human beings. Technologies are the result of a plurality of agents who take decisions to make them how they are. These decisions involve a diversity of values, among which the values that are technical, political and economic often prevail. But all decisions involve values-related dilemmas that call for citizen participation—users, consumers, those responsible for the technological products, those affected by them, and so forth.

Societal inclusion relates to the general human existential situation of vulnerability and finitude which is true for everyone, not only for persons with diseases or handicaps. "Better robotics for inclusive societies" needs to take this issue into account. The benchmark is social life in which finitude and vulnerability are not seen as weakness or disability, but as universal human feature which also enables creativity, arts, and joy of life. Anything else could lead to inhumane technocratic ideologies, where "better societies" are defined by "inclusive robotics", thus by techno-economical possibilities, instead of real human and social needs.

Within the relevant discursive frameworks for development of an inclusive robotics, important ethical-political ideas must be analysed: autonomy, dependency, vulnerability, functioning, care and disability. The implementation of a truly inclusive robotics should take into account the UN Convention on the Rights of Persons with Disabilities and the Convention on the Rights of the Child, given the importance socio-technical mediations have for full enjoyment of human rights.

Whilst there is evidence of significant benefits to be realised from the use of social assistive robotics for people with intellectual and developmental disabilities or cognitive impairments such as dementia, certain ethical issues need to be addressed, focusing on privacy, autonomy, and responsibility.

The introduction of technological elements and, in particular, robotic elements, into functionings environments (such as the urban environment, domestic environment, educational environment, work environment, environments for public, political, social and economic participation, environments for culture, leisure or health) transforms the way in which the actions are carried out and, consequently, can modify evaluation of the functionings that are mediated by such elements. As devices that mediate actions, robots not only transform the practices carried out in an environment, but also its characteristic values. If in a given environment the situation arises in which mediation from a device is obligatorily needed to perform a certain functioning, the limitations of this device will define the barriers of this action. In the case of robotic devices, these barriers will produce a "robotics divide". The main advantage of using the capabilities approach for analysing the effect of robotic innovations on functionings environments is that it introduces important ethical and social considerations on questions that, at first glance, could appear to only be technical or instrumental.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

If one of the aims of using robots is to create a *better* society, the environmental impact of robotics cannot be ignored. We need a green agenda for robotics, paying attention to robots for environmental research and remediation, the environmental impact of the type of materials used to make the robots; or of their degradation process, etc. The creation of new needs on the part of the consumer might have negative environmental consequences. Citizens might prefer a system that does not further damage the natural environment.

Policymakers have the opportunity to shape the development of interactive robotics. Interactive robotics are often framed in terms of opportunities and risks to individuals and societies, without reference to the natural environment. The danger now is that the development of interactive robotics – the economies, markets, institutions, and practices – takes place without consideration of environmental costs. This would risk putting us on a technological treadmill – where robots are required – despite enormous environmental harms. Alternatively, if environmental values are built into the economies, markets, institutions and practices relating to interactive robotics, the environmental harms can be minimised, and interactive robotics can serve as a model for other novel technologies.

Public authorities can play a supporting role through a smart mix of voluntary policy measures and, where necessary, complementary regulation, as businesses should integrate Sustainable Development Goals in their operations. Responsible Business Conduct is an important horizontal enabler for sustainability change if geared towards an innovative green, inclusive and socially just economic transition. Robotics companies must report on the due diligence process that they implement with regard to environmental, social and employee issues, respect for human rights, and bribery and anti-corruption. The influence of robotics on these topics could be also voluntarily explained, as a voluntary action showing their Corporate Social Responsibility.

In a similar vein, the Responsible Research and Innovation approach aims to make society a participant in science and innovation from its most initial phases, involving all actors and aligning the processes and results of the research with the values, needs and expectations of society. In practice, this concept implies designing and implementing research and innovation policies that pay special attention to six interrelated agendas: ethics, gender equality, governmental agreements, open access to scientific information, public participation and scientific education.

There are two main aspects to consider in the relation between robotics and diversity: the treatment of women in robotics, and how far robotics advances or hinders women's position in the workplace and the business world. The use of algorithms, even those created with objective data poses a potential danger of maintaining or even widening the gender gap as even objective data will be shaped by the history of human thought. The lack of women in the robotics industry in the threefold areas of research, programming and enterprise is a problem for the actual development and evolution of artificial intelligence and robotics. Robotics does not marginalise women, and it can initially be assumed that it will not have a negative effect on them, but rather the reverse. However, it is essential that women should not be left behind, and this can be avoided by placing special emphasis on training and education to ensure that their jobs will be safeguarded in this new era. Corporate directors and company managers will have to play an important role in this ambit.

Currently, the mentions to minors in standards are quite vague and do not favour their proper protection. Issues on prevention or limits still can be thoroughly developed. The main actors in

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

the robotics and artificial intelligence environment should move forward, together with parents and guardians. With respect to access, minimum ages should be set in the interest of minors.

Of course, the application of the principles of prevention, mitigation, and protection is relevant at a time when the principles and process contained in the Guiding Principles of Business and Human Rights are beginning to be widely understood by companies, also helped for the development of national business and human rights frameworks. But together with the establishment of due diligence processes by public and private actors, the involvement of fundamental rights invokes the role of the State in the establishment of economic and institutional incentives (for example, in public procurement mechanisms that encourage the use of responsible technologies) for systems that respect human rights in the deployment of artificial intelligence and robotics. Along with this, the establishment of certification programs or labels on responsible artificial intelligence and robotics might be useful.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

9. Annex 1

1. Workshop on Ethics and Corporate Social Responsibility for Inclusive Robotics[406]

Robots in society can be seen as a threat to human dignity, privacy, freedom, equal access or desirable social effects – perhaps furthering marginalization through a form of the digital divide (the robotic gap). It has thus been said that robots should be "inclusive", which would contribute to their acceptance in society. The main idea of this workshop is to bring to the fore the diversity, potential dissonances and points of conflict regarding existing definitions and expectations of inclusiveness of interactive robotics. This Workshop allows an exploration of ethical and corporate social responsibility issues regarding the design and implementation of inclusive robotics, with particular focus on the role of robot designers, industry and societal stakeholders.

We examine the question of inclusive robotics in relation to society as a whole and in relation to documents such as the Convention on the Rights of Persons with Disabilities (CRPD). Different models from which interactive robotics can be approached, with special emphasis on the dichotomy between the social approach vs medical rehabilitation model of disability, are presented in relation to vulnerability, functional diversity and care. In order to overcome these barriers and to promote the implementation of truly inclusive interactive robotics, we raise principles and good practices that should constitute the ethical framework for responsible robotics committed to fairness, justice and the well-being of people.

The Project partners report on the work done on these issues, and the Experts from the current Groups set by the European Commission share the state of the art and explain their views and the next actions planned to overcome difficulties.

Agenda:

Ethics of AI and Robotics.Vincent C. Müller (UAF, University of Leeds)

Inclusiveness of wearable robotics.Heike Felzman (NUI Galway, Cost Action 16116)

Corporate Social Responsibility for inclusive robotics.Aníbal Monasterio (CSIC, INBOTS)

A responsible entrepreneurial view.Freygarður Thorsteinsson (ÖSSUR, INBOTS)

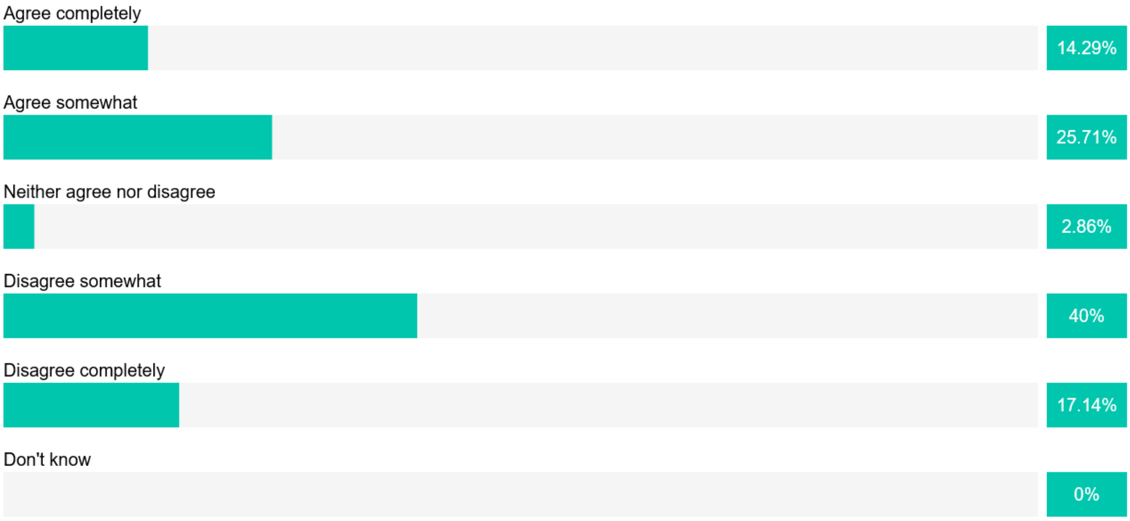
Corporate Social Responsibility for Inclusive Robotics. Amparo Grau (Complutense University of Madrid, INBOTS)

More information at Workshop: Workshop: Ethics and Corporate Social Responsibility for Inclusive Robotics

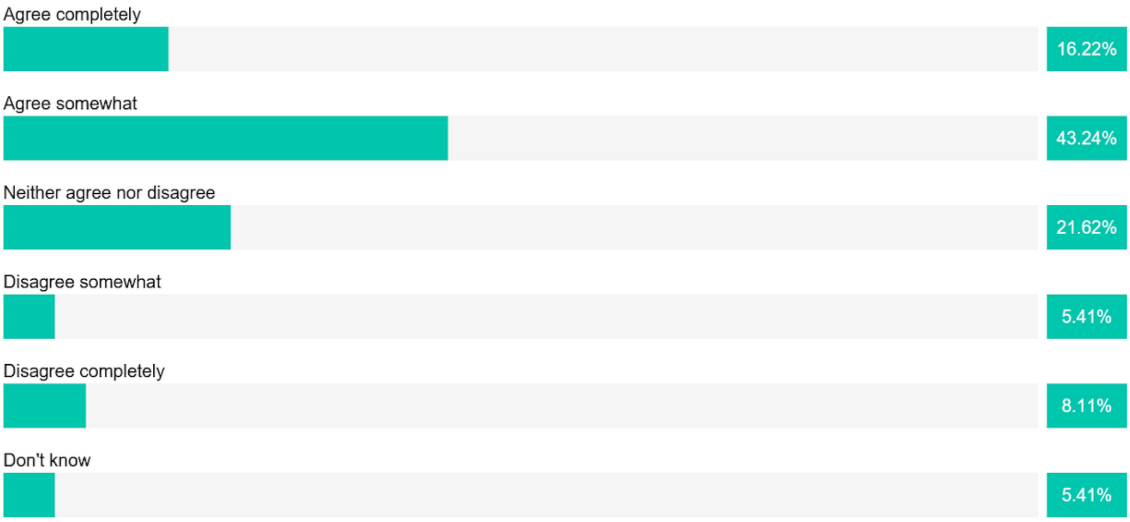
[406] ERF – INBOTS Conference 2019, 20th of March from 14:00 to 15:30. Room Galati
Moderator: Vincent Müller (University of Leeds, INBOTS)
Co-Moderator: Heike Felzman (NUI Galway, Cost Action 16116)

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

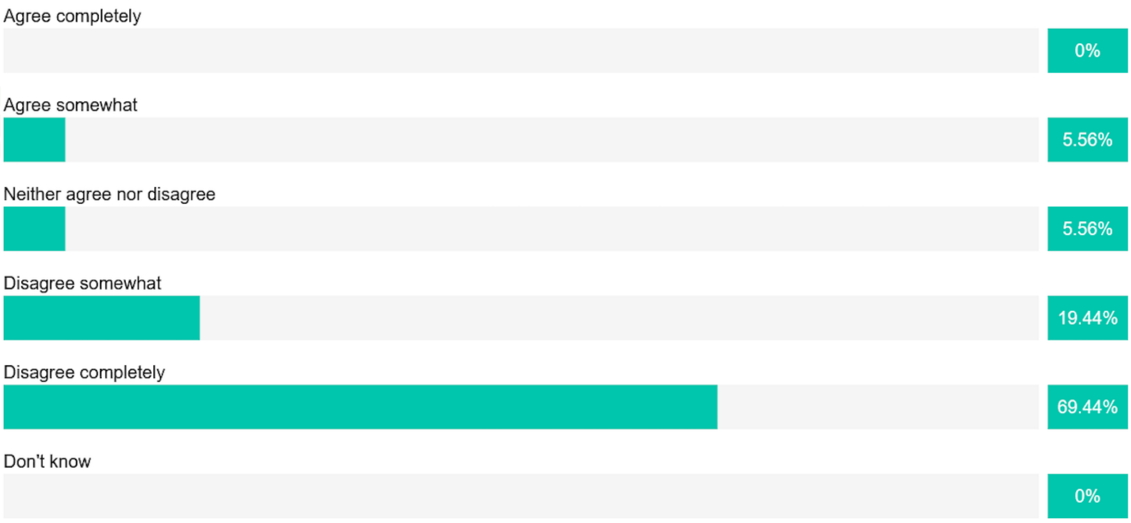
I am very concerned about the presence of robotics in society and in people's daily lives.



I am sure that, despite the challenges, society will ultimately become better with the increase in automation, robotics, and AI.

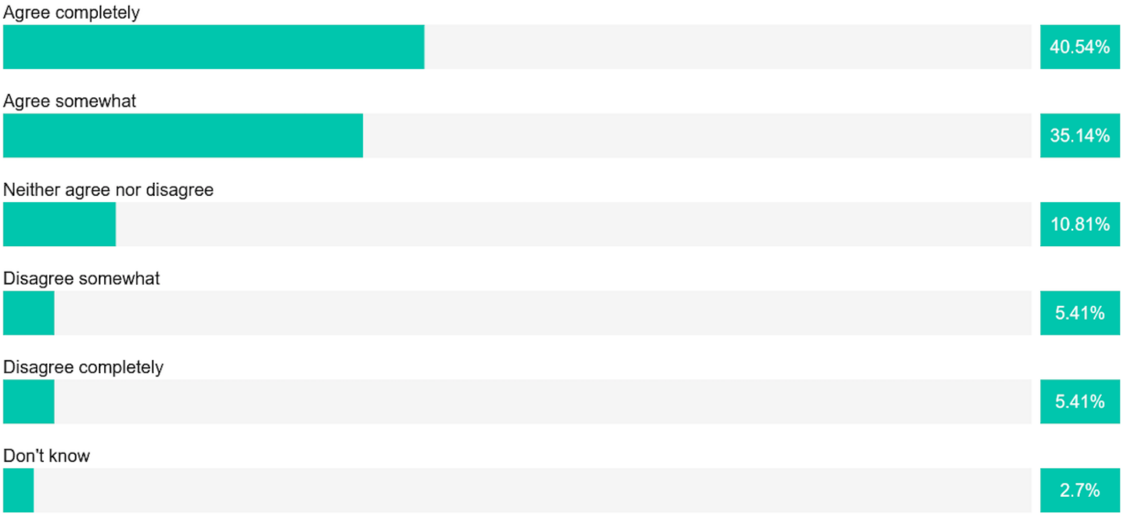


I think that economic and efficiency aspects need to come first in the development of robots; ethical, legal or social values only come in once the robot is being deployed.



This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

I fear that future developments in robotics may negatively affect the protection of personal data and privacy.



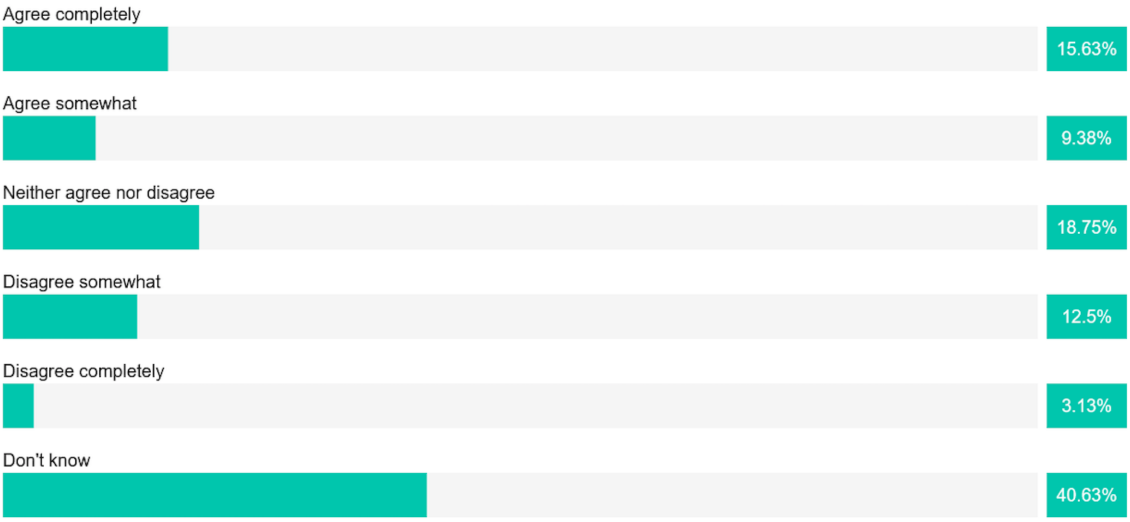
I believe too little attention is being paid to the needs and perspectives of vulnerable groups in the design of robots.

I believe too little attention is being paid to the needs and perspectives of workers in the design of robots.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

In my opinion, robot companies need to pay more attention to the frameworks and practices of Responsible Research and Innovation (RRI) and CSR.

In my opinion, meaningful Responsible Research and Innovation (RRI) and CSR are very challenging to realize for robot companies in practice.



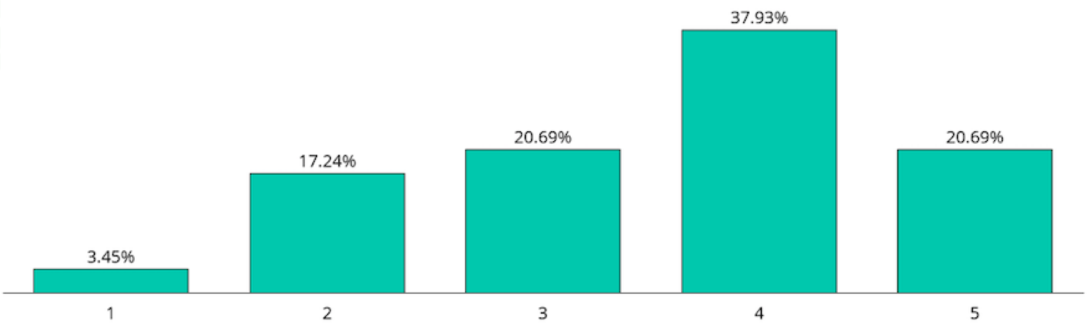
Which values should be promoted through RRI and CSR?

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

What do you think are the most important barriers for the realization of meaningful RRI and CSR regarding robotics?

Do we need to have public regulation to make robotics inclusive?

Average response: 3.55



This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Which incentives do we need to promote CSR in robotics?

2. Workshop on Sustainable Public Policies for Innovation and the Future of Work[407]

In this workshop we explore how different considerations coming from robot designers, industry and societal stakeholders in relationship to robots and the future of work are translated into public policy, and how this process could be improved. Introducing interactive robots in our society will have economic and legal consequences. The High-Level Expert Group on the impact of the digital transformation on EU labour markets alerts the society to the potential positive or negative impacts of robotics in the labour market. A robotized administration, for instance, may trigger the redefinition of the legal and financial systems. The international institutions have reluctantly admitted the use of tax benefits as a way to foster innovation through Public Finance but stressing their proportionality. This approach could be useful to solve the current needs recently observed for training in the digital economy transition phase, not to leave anyone behind. However, there is still an apparent contradiction between the innovation policy and the uncertain future of workers that needs to be unveiled. This situation calls for sound legislation making compatible their fair protection and promoting EU companies' competitiveness, productivity and sustainability. Under these circumstances, and at the risk of losing control over decision-making processes in the hands of autonomous processes, these decision processes will have to be clearly and explicitly defined. In this sense, mechanisms allowing stakeholder involvement, including workers, are deemed to be necessary. This workshop serves as a platform to define decision-making processes for the insertion of robotics in society and in the workplace, and to devise fair and transparent stakeholder involvement instruments. Experts from economics,

[407] ERF – INBOTS Conference 2019, 21st of March from 10:45 to 12:15. Room Galati
Moderator: Amparo Grau (Complutense University of Madrid, INBOTS)
Co-Moderator: Eduard Fosch Villaronga (Leiden University, Cost Action 16116)

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

business, law, labor market and policymakers share their perspectives with the robotics community in depth in order to reach some consensus.

Agenda:

Tax incentives for human reskilling in the transition to a robotized world?Amparo Grau (Complutense University of Madrid, INBOTS)

Robotics and Healthcare: Convergence Frameworks.Robin L. Pierce (Tilburg University)

ILO and the future of work: a human centred agenda.Mari Luz Vega (International Labour Organization)

The impact of the digital transformation on EU labour markets.Maarten Goos, Ronja Roettger (Utrecht University, INBOTS)

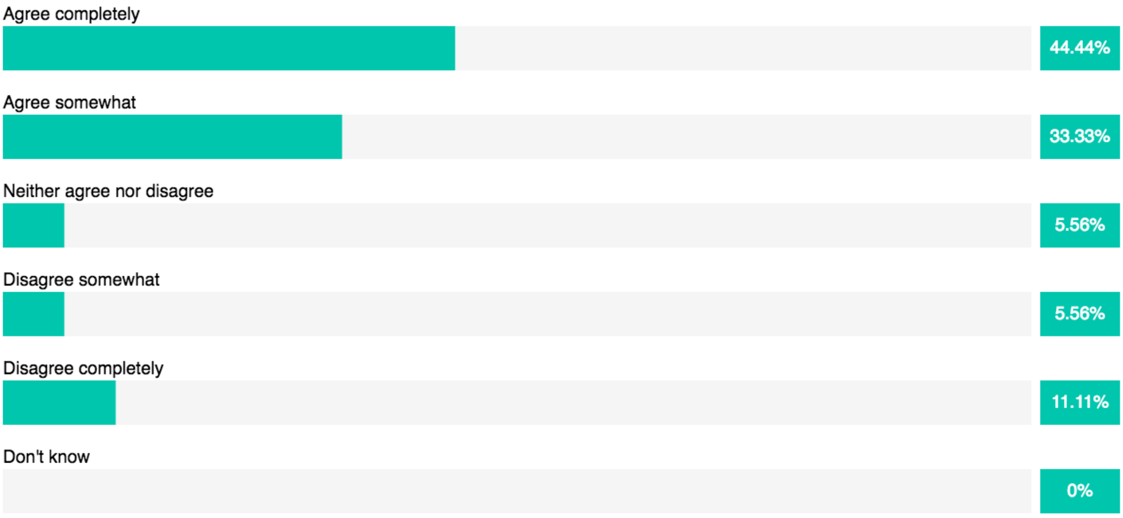
Innovation and IP policies. Luke Mc Donagh (CITY University of London, INBOTS)

Cooperation to drive robotics innovations. Francesco Ferro (PAL Robotics, INBOTS)

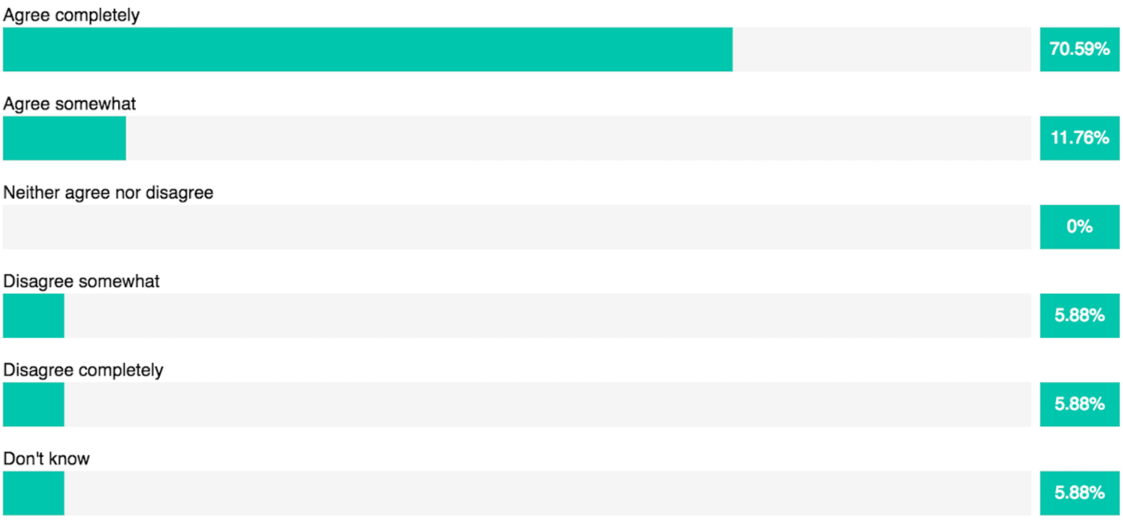
More information at Workshop: Sustainable Public Policies for Innovation and the Future of Work

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

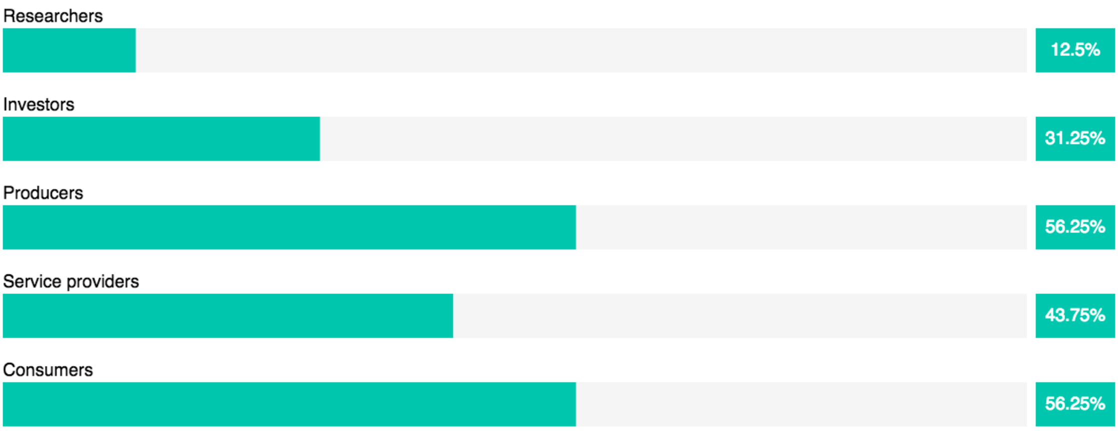
The introduction of interactive robots into our society should be regulated to avoid inequality in society.



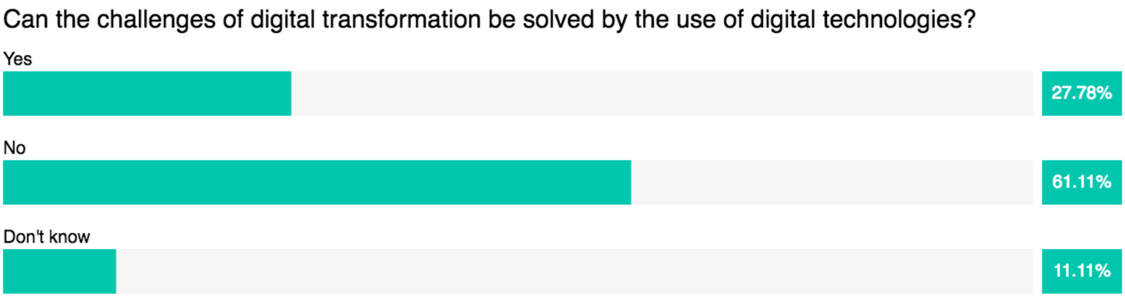
The introduction of interactive robots into our society should be regulated to ensure access of potential users to the technology.

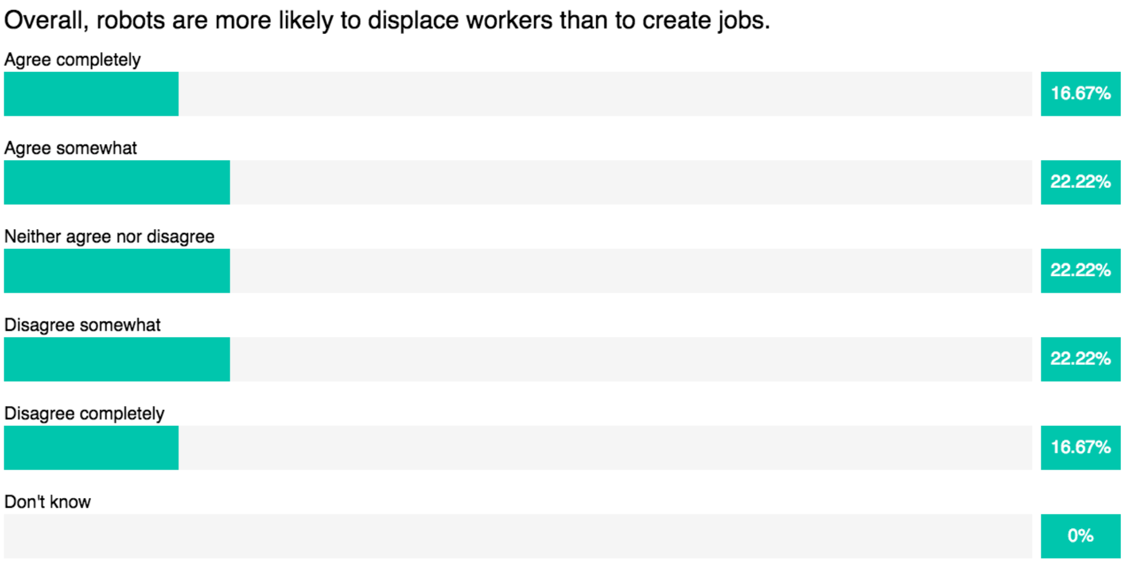


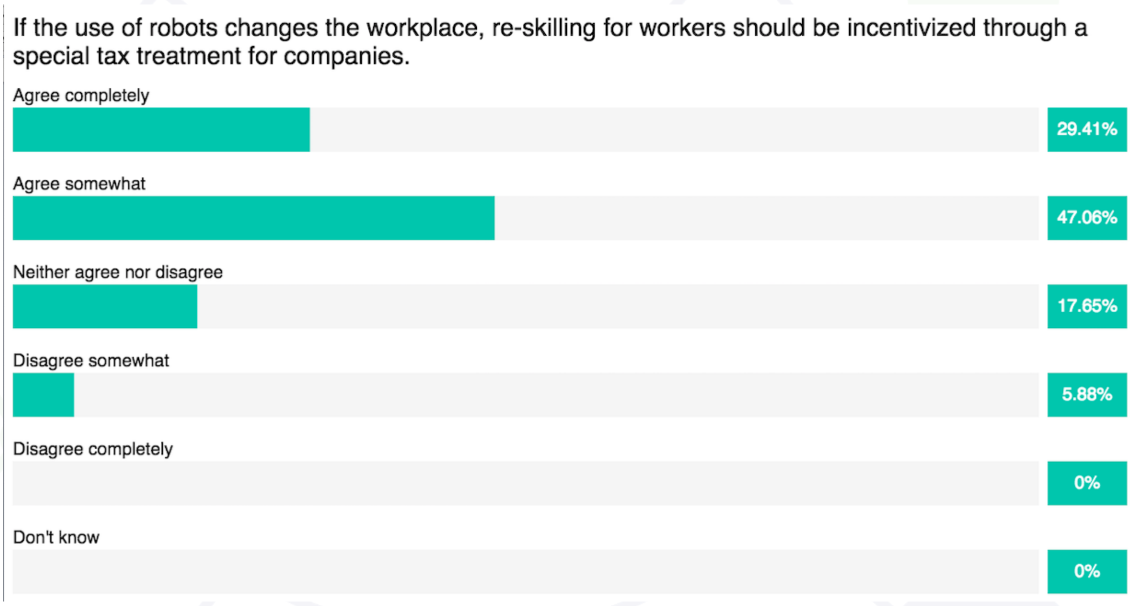
The regulation of robots entails risks for… (pick a maximum of 3)



This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

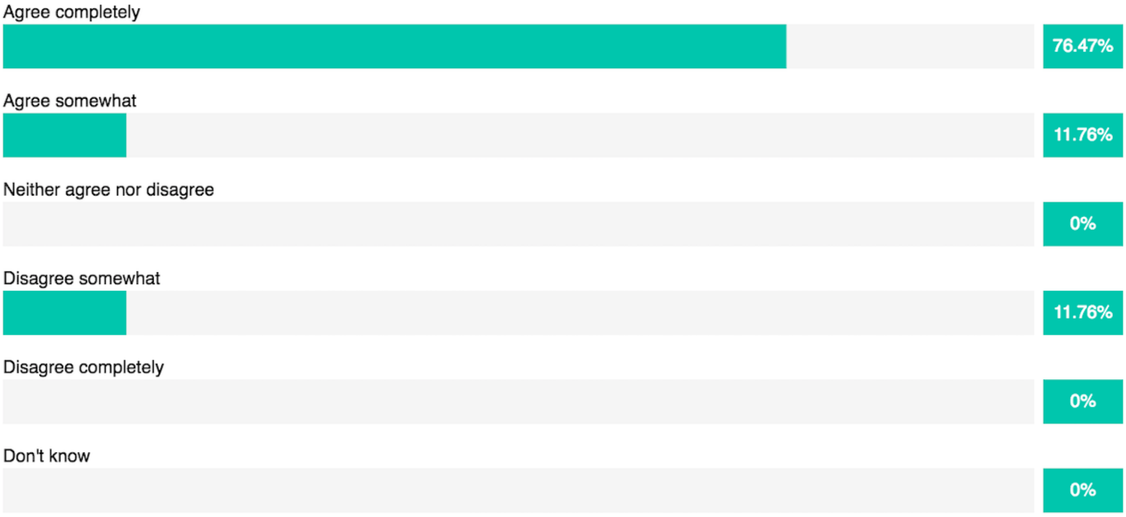






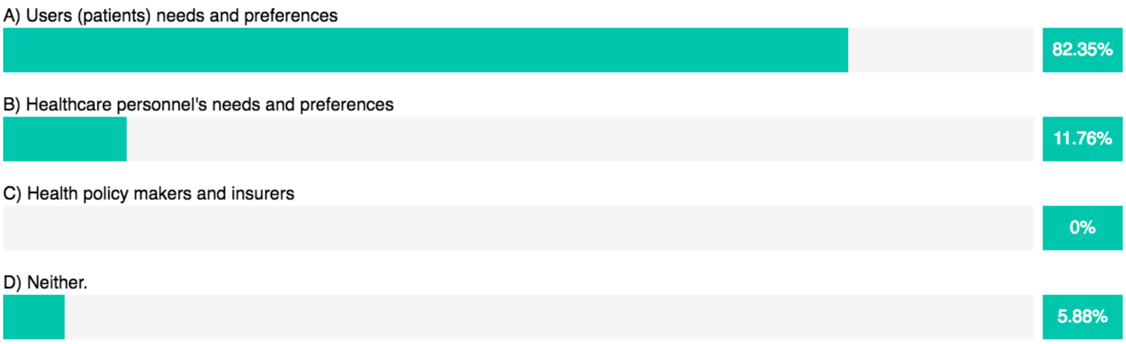
This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Roboticists should take into account Environmental, Social and Governance criteria for sustainable development in the work.

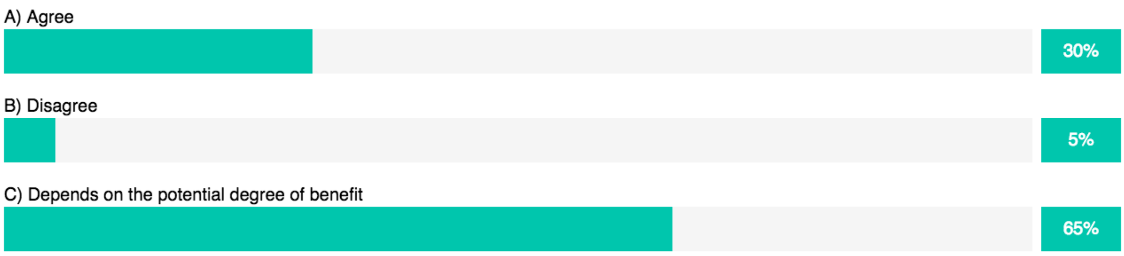


Would it be useful to create an Observatory, Forum or Platform that follows the regulatory and legislative processes with regards to robotics in society?

How robots are integrated into the health domain should depend primarily on:



The longstanding norms and practices of healthcare (e.g. Doctor-Patient relationship) should not be altered by the introduction of robots.



This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Healthcare workers should be able to refuse to work with robots, e.g. wearing of exoskeletons for elderly care, without penalty or impact on job security.

A more detailed explanation of these materials has been disseminated in the following article:

Eduard Fosch-Villaronga y María Amparo Grau Ruiz, "Expert considerations for the regulation of assistive robotics. A European Robotics Forum Echo", DILEMATA International Journal of Applied Ethics [ISSN 1989-7022] 30 (May 2019 - Ethics, Robotics, and Assistive Technologies). Available at the following link: https://www.dilemata.net/revista/index.php/dilemata/issue/view/31

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

10. Annex 2

1. Written dissemination of ideas to promote debate

Bertolini A. *Artificial Intelligence and Civil Law: Liability Rules for Drones*. Policy Department far Citizens' Rights and Constitutional Affairs Directorate Generai far Internal Policies of the Union: Brussels, 2018.

Bertolini A. Human-Robot Interaction and Deception. Osservatorio del diritto civile e commerciale 2018.

Bertolini A, Aiello G. "Robot companions: A legal and ethical analysis". *The Information Society* , 34 (2018): 130-140.

Bonadio, E., McDonagh, L. and Arvidsson, C. "Intellectual property aspects of robotics". *European Journal of Risk Regulation*, 9 (2018): 655–676. doi:10.1017/err.2018.58.

Grau Ruiz, M.A. "Tributación y Robótica Responsables para una Economía Digital Sostenible", in *Braccia, M.: Tributación de la Economía Digital*, Thomson Reuters Aranzadi, 2019.

Grau Ruiz, M.A.; Sánchez-Urán Azaña, M.Y. López Sánchez, J.I.; Grau Ruiz, M.A.; Sánchez-Urán Azaña, M.Y. "The impact of robotics and computerization on the labour market: Inclusive insight from a Law and Economics perspective". *Derecho Digital e Innovación, Wolters Kluwer,* 3 (2019).

Sánchez-Urán Azaña, M.Y.; Grau Ruiz, M.A. "El impacto de la robótica, en especial la robótica inclusiva, en el trabajo: aspecto jurídicos-laborales y fiscales". *Aranzadi de Derecho y Nuevas Tecnologías*, (2019). Also as Eprint UCM https://eprints.ucm.es/47523/, 2018.

Sánchez-Urán Azaña, M.Y.; Grau Ruiz, M.A. "Nuevas Tecnologías y Derecho: Retos y Oportunidades planteados por la Robótica y la Inteligencia Artificial". (2019), Juruá (in press).

Grau Ruiz, M.A. "The Importance Of Ethics In The World Of Ai And Robotics". *Otrosí (Review of the Madrid Bar),* 2 (2019)

https://www.otrosi.net/hemeroteca/numero-2-2019-7a-epoca

Eduard Fosch-Villaronga and María Amparo Grau Ruiz. "Expert considerations for the regulation of assistive robotics. A European Robotics Forum Echo", *DILEMATA International Journal of Applied Ethics [ISSN 1989-7022]* 30 (May 2019 - Ethics, Robotics, and Assistive Technologies). Available at the following link: https://www.dilemata.net/revista/index.php/dilemata/issue/view/31

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Aníbal Monasterio Astobiza: "Ética para máquinas: Similitudes y diferencias entre la moral artificial y la moral humana", *Dilemata, International Journal of Applied Ethics* , 30, 129-147.

Aparicio-Payá M, Toboso-Martín M, Ausín-Díez T, Monasterio-Astobiza A, Morte-Ferrer R, López-Castro D. (2019). "Un marco ético-político para la robótica asistencial". **Artefactos. Revista de estudios sobre la ciencia y la tecnología** 8(1): 97-117. Disponible en: http://revistas.usal.es/index.php/artefactos/article/view/19837

DOI: http://dx.doi.org/10.14201/art20198197117

Mario Toboso and Manuel Aparicio. "Entornos de funcionamientos robotizados. ¿Es posible una robótica inclusiva?" *Dilemata, International Journal of Applied Ethics* , 30 (2019): 171-185.

Mario Toboso, Manuel Aparicio, Daniel López*. "Ethics, robotics and assistive technologies". Dilemata. International Journal of Applied Ethics*, 30.

Aníbal Monasterio Astobiza, Txetxu Ausín, Mario Toboso, Ricardo Morte, Manuel Aparicio, Daniel López. "Traducir el pensamiento en acción: Interfaces cerebro-máquina y el problema ético de la agencia". *Revista de Bioética y Derecho*; 46(2019) : 29-46.

Aparicio-Payá M, Toboso-Martín M, Ausín-Díez T, Monasterio-Astobiza A, Morte-Ferrer R, López-Castro D. "Un marco ético-político para la robótica asistencial". *Artefactos. Revista de estudios sobre la ciencia y la tecnología* 8 (2019): 97-117, http://revistas.usal.es/index.php/artefactos/article/view/19837

DOI: http://dx.doi.org/10.14201/art20198197117

Aníbal Monasterio, Mario Toboso, Manuel Aparicio, Txextu Ausín, Daniel López, Ricardo Morte, José L. Pons. "Bringing inclusivity to robotics with INBOTS". *Nature Machine Intelligence*, 1 (2019), 164. ISSN: 2522-5839

DOI: https://doi.org/10.1038/s42256-019-0040-5

Mario Toboso, Ricardo Morte, Aníbal Monasterio, Txextu Ausín, Manuel Aparicio, Daniel López. "Robotics as an instrument for social mediation". *Proceedings of the 1st Anual INBOTS Conference*, "Inclusive Robotics for a Better Society" (INBOTS – 780073 – H2020), October 16–20, 2018, Pisa, Italy. Springer (in press).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Manuel Aparicio, Mario Toboso, Txetxu Ausín, Daniel López, Ricardo Morte and Aníbal Monasterio. Discursive Frameworks for the Development of Inclusive Robotics. *Proceedings of the 1st Anual INBOTS Conference*, "Inclusive Robotics for a Better Society" (INBOTS – 780073 – H2020), October 16–20, 2018, Pisa, Italy. Springer, (in press).

Daniel López, Aníbal Monasterio, Mario Toboso, Manuel Aparicio, Txetxu Ausín, Ricardo Morte Cartography of the Values Involved in Robotics. *Proceedings of the 1st Anual INBOTS Conference*, "Inclusive Robotics for a Better Society" (INBOTS – 780073 – H2020), October 16–20, 2018, Pisa, Italy. Springer (in press).

Aníbal Monasterio, Daniel López, Manuel Aparicio, Ricardo Morte, Txetxu Ausín, Mario Toboso (2019). Conceptual Analysis: technology, machine and robot. *Proceedings of the 1st Anual INBOTS Conference*, "Inclusive Robotics for a Better Society" (INBOTS – 780073 – H2020), October 16–20, 2018, Pisa, Italy. Springer (in press).

Sánchez-Urán Azaña, María Yolanda y Grau Ruiz, María Amparo (2019) *Nuevas tecnologías y Derecho. Retos y oportunidades planteados por la Robótica y la Inteligencia Artificial.* In Nuevas tecnologías y Derecho. Retos y oportunidades planteados por la Robótica y la Inteligencia Artificial, 25 a 28 de febrero de 2019, Facultad de Derecho, UCM. https://eprints.ucm.es/56042/

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

11. **References**

1. Clarifying liability, risk management and insurance issues

Alpa, Guido, and Vincenzo Zeno-Zencovich. *Italian Private Law.* Routeledge-Cavendish, 2007.

Anscombe, G.E.M. *Intention.* Oxford: Basil Blackwell, 1957.

Aristotle. *The Nicomachean Ethics.* Translated by Terence Irwin. Indianapolis: Hackett Publishing, 1985.

Barandiaran, X.E., E Di Paolo, and M. Rohde. "Defining Agency: Individuality, Normativity, Asymmetry, and Spatio-Temporality in Action." *Adaptive Behavior* 17, 5 (2009): 367–86.

Bertolini, Andrea. "Insurance and Risk Management for Robotic Devices: Identifying the Problems." *Global Jurist*, 2 (2016): 1-24.

———. "Robots and Liability - Justifying a Change in Perspective." In *Rethinking Responsibility in Science and Technology*, edited by Fiorella Battaglia, Julian Nida-Rümelin and Nikil Mukerji, 143-66. Pisa: Pisa University Press, 2014.

———. "Robots as Products: The Case for a Realistic Analysis of Robotic Applications and Liability Rules." *Law Innovation and Technology* 5, no. 2 (2013): 214.

Bobbio, Norberto. "Sanzione." In *Novissimo Digesto*, 350-540. Torino: UTET, 1969.

Bryson, Joanna J , and Philip P. Kime. "Just an Artifact : Why Machines Are Perceived as Moral Agents." In *Proceedings of the Twenty-Second International Joint Conference on Artificial Intelligence: Barcelona, Catalonia, Spain, 16–22 July 2011*, edited by Toby Walsh, 1641-46. Menlo Park, CA, USA AAAI Press, 2011.

Calabresi, Guido, and Douglas A. Melamed. "Property Rules, Liability Rules, and Inalienability: One View of the Cathedral." *Harvard Law Review* 85, no. 6 (1972): 1089.

Calo, Ryan. "Robotics and the Lessons of Cyberlaw." *California Law Review* 103, no. 3 (2015): 513-63.

Coleman, Jules, Scott Hershovitz, and Gabriel Mendlow. "Theories of the Common Law of Torts." In *The Stanford Encyclopedia of Philosophy*, edited by Edward Zalta. \url{https://plato.stanford.edu/archives/win2015/entries/tort-theories/, Winter 2015.

"Communication from the Commission to the European Parliament, the European Council, the Council, the European Economic and Social Committee and the Committee of the Regions. Artificial Intelligence for Europe." Brussels: European Commission, 2018.

Davidson, D. "Actions, Reasons, and Causes." (1963).

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Endicott, Thomas. "Law and Language." In *The Stanford Encyclopedia of Philosophy*, edited by Edward Zalta, https://plato.stanford.edu/entries/law-language/, 2016.

———. *Vagueness in Law.* Oxford: Oxford University Press, 2000.

Eshleman, Andrew. "Moral Responsibility." In *The Stanford Encyclopedia of Philosophy (Winter 2016 Edition),*: T Edward N. Zalta, 2016.

Floridi, Luciano, and J.W. Sanders. "On the Morality of Artificial Agents." *Minds and Machine* 14 (2004): 349–79.

Fossa, Fabio. "Artificial Moral Agents: Moral Mentors or Sensible Tools?". *Ethics and Information Technology* 20, no. 2 (June 01 2018): 115-26.

Frankfurt, H., 1971. "Freedom of the Will and the Concept of a Person." *Journal of Philosophy* 68, no. 1 (1971): 5-20.

Ginet, C. *On Action.* Cambridge: Cambridge University Press, 1990.

Gutman, Mathias, Benjamin Rathgeber, and Tareq Syed. "Action and Autonomy: A Hidden Dilemma in Artificial Autonomous Systems." In *Robo- and Informationethics. Some Fundamentals*, edited by Michael Decker and Mathias Gutman, 231-56: Lit Verlag, 2012.

Himma, Kenneth Einar. "Artificial Agency, Consciousness, and the Criteria for Moral Agency: What Properties Must an Artificial Agent Have to Be a Moral Agent?". *Ethics and Information Technology* 11, no. 1 (March 01 2009): 19-29.

The Cambridge Dictionary of Philosophy. Cambridge: Cambridge University Press, 1999.

Lanzarone, Gaetano Aurelio, and Federico Gobbo. "Is Computer Ethics Computable?". In *Conference Proceedings of Ethicomp 2008: Living, Working and Learning Beyond Technology*, edited by Terrell Ward et Al. Bynum, 530. Mantova: Tipografia Commerciale, 2008.

Leroux, Christophe, Roberto Labruto, Chiara Boscarato, Franco Caroleo, Jan-Philipp Günther, Severin Löffler, Florian Münch*, et al.* "Suggestion for a Green Paper on Legal Issues in Robotics. Contribution to Deliverable D.3.2.1 on Els Issues in Robotics." 2012.

Matthias, Andreas. "The Responsibility Gap: Ascribing Responsibility for the Actions of Learning Automata." *Ethics and Information Technology* 6 (2004).

O'Connor, Timothy, and Christopher Franklin. "Free Will." In *The Stanford Encyclopedia of Philosophy (Winter 2016 Edition),*: T Edward N. Zalta, 2018.

Palmerini, Erica, and Andrea Bertolini. "Liability and Risk Management in Robotics." In *Digital Revolution: Challenges for Contract Law in Practice*, edited by Reiner Schulze and Dirk Staudenmayer, 225-59. Baden-Baden: Nomos, 2016.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Polinsky, Mitchell A, and Steven Shavell. *Handbook of Law and Economics.* Vol. I: North-Holland, 2007.

Rest, J. R. *Moral Development: Advances in Research and Theory.* New York: Praeger Publishers, 1986.

Reynolds, Scott J. "Moral Awareness and Ethical Predispositions: Investigating the Role of Individual Differences in the Recognition of Moral Issues." *Journal of Applied Psychology* 91, no. 1 (2006): 233-43.

Schlosser, Markus. "Agency." In *The Stanford Encyclopedia of Philosophy (Fall 2015 Edition)*: Edward N. Zalta.

Shavell, Steven. "Liability for Accidents." In *Handbook of Law and Economics*, edited by A. Mitchell Polinsky and Steven Shavell, 142. Amsterdam: Elsevier, 2007.

Strawson, Peter F. "Freedom and Resentment." In *Proceedings of the British Academy, Volume 48: 1962*, edited by Gary Watson, 1-25: Oup Oxford, 1962.

Taylor, C. "What Is Human Agency?". In *He Self: Psychological and Philosophical Issues*, edited by T Michel, 103-35. Oxford: Blackwell, 1977.

Toulmin, Stephen. *The Uses of Argument.* Cambridge: Cambridge University Press, 1958.

Toulmin, Stephen E. *The Uses of Argument.* Vol. 34: Cambridge University Press, 1958.

Van Gerven, Walter , Jeremy Lever, and Pierre Larouche. *Tort Law* Oxford: Hart Publishing, 2000.

Waldron, Jeremy. "Vagueness in Law and Language: Some Philosophical Issues." *California Law Review* 82, no. 1 (1994): 509.

Walen, Alec. "Retributive Justice." In *The Stanford Encyclopedia of Philosophy*, edited by Edward Zalta. URL = <https://plato.stanford.edu/archives/win2016/entries/justice-retributive/>. Winter 2016 Edition.

2. Clarifying the role of IP law

C. Andrew Keisner, Consultant, Julio Raffo and Sacha Wunsch-Vincent, Breakthrough Technologies – Robotics and IP, December 2016, Economics and Statistics Division, WIPO 2016, available at http://www.wipo.int/wipo_magazine/en/2016/06/article_0002.html.

"Making Your Robotics Company a More Attractive Investment", in Robotic Business Review of 21 October 2012, at https://www.roboticsbusinessreview.com/unmanned/making_your_robotics_company_a_more_attractive_investment**...**

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

3. Determining the impact of interactive robotics in the labour market

Abraham, K.G. 1988. "Flexible Staffing Arrangements and Employers' Short-Term Adjustment Strategies."

Acemoglu, D. and Restrepo, P. 2019. "Demographics and Automation." MIT working paper.

Alabdulkareem, A., Frank, M. R., Sun, L., Hidalgo, C. and Rahwan, I. 2018. "Unpacking the polarization of workplace skills", Science Advances, 4(7).

Agrawal, A., Gans, J., and Goldfarb, A. 2018. Prediction Machines: The Simple Economics of Artificial Intelligence. Harvard Business Review Press.

Arntz, M., Gregory, T. and Zierahn, U. 2017. "Revisiting the risk of automation," *Economics Letters*, 159(C): 157-160.

Autor, D.H. 2019. "Work of the Past, Work of the Future," Richard T. Ely Lecture, American Economic Association: Papers and Proceedings, forthcoming, May 2019.

Autor, D.H., Dorn, D., and Hanson, G. 2016. "The China Shock: Learning from labour Market Adjustment to Large Changes in Trade." *Annual Review of Economics*

Autor, D.H., Dorn, D., and Hanson, G. 2013. "The China Syndrome: Local labour Market Effects of Import Competition in the United States." *American Economic Review*

Autor, D.H. and Fournier, J. 2019. "No Country for Old Men: The Inversion of the Rural-Urban Age Gradient." Working paper, February.

Autor, D.H., Katz, L. and Krueger, A. 2006. "The Polarization of the U.S. labour Market," *American Economic Review Papers and Proceedings* 96 (2), 189–194.

Autor, D.H., Katz, L.F. and Krueger, A. 1998. "Computing Inequality: Have Computers Changed the labour Market?" *The Quarterly Journal of Economics*, 113(4): 1169–1213.

Autor, D.H., and Salomons, A. 2019. "New Frontiers: The Evolving Content and Geography of New Work in the 20th Century." Working Paper.

Autor, D.H., and Scarborough, D. 2008. "Does Job Testing Harm Minority Workers? Evidence from Retail Establishments," *Quarterly Journal of Economics*, 123: 219–277.

Appelbaum, Kalleberg, A., and Hye, J.R. 2019. "Nonstandard Work Arrangements and Older Americans, 2005–2017." https://www.epi.org/publication/nonstandard-work-arrangements-and-older-americans-2005-2017/

Bagues, M. F. & Labini, M. S. 2009. Do Online labour Market Intermediaries Matter? In: D. H. Autor (ed.), Studies of labour Market Intermediation. Chicago: University of Chicago Press, pp. 127-154.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Baker, G. and Hubbard, T. "Make Versus Buy in Trucking: Asset Ownership, Job Design, and Information", *The American Economic Review*, 93 (2003): 551-572.

Baldwin, Richard. 2019. 'The Globotics Upheaval: Globalization, Robotics and the Future of Work', *Oxford University Press*, February 2019.

Bessen, J., Goos, M., Salomons, A. and Vandenberge, W. 2019. "Automatic Reaction: What Happens to Workers at Firms that Automate?" Boston University School of Law, Law and Economics Research Paper. https://ssrn.com/abstract=3328877

Bolukbasi, T., Chang, K.-W., Zou, J. Y., Saligrama, V., and Kalai, A. T. "Man Is to Computer Programmer as Woman Is to Homemaker? Debiasing Word Embeddings." *Adv. Neural Inf. Process. Syst.*: (2016): 4349–4357.

Bughin, J., Hazan, E., Lund, S,. ; Dahlström, P., Wiesinger, A., and A. Subramaniam. 2018. "Skill Shift Automation and The Future Of The Workforce", McKinsey Global Institute, May 2018. https://www.mckinsey.com/featured-insights/future-of-work/skill-shift-automation-and-the-future-of-the-workforce

Card, D., Kluve, J., and Weber, A. "Active Labour Market Policy Evaluations: A Meta-Analysis." *The Economic Journal* 120.548. Chen, M.K., Chevalier, J.A., Rossi, P.E. and Oehlsen, E. 2019. "The Value of Flexible Work: Evidence from Uber Drivers." *Journal of Political Economy*, forthcoming.

Ciarli, T., Marzucchi, A., Salgado, E., Savona, M. 2018. "The Effect of R&D Growth on Employment and Self-Employment in Local Labour Markets." SPRU Working Paper Series 2018-08, SPRU - Science Policy Research Unit, University of Sussex Business School.

Clark, A.E. 2001. "What Really Matters in a Job? Hedonic Measurement Using Quit Data," *Labour Economics*, 8(2): 223-242.

CNBC. 2019. "IBM Artificial Intelligence Can Predict With 95% Accuracy Which Workers Are About To Quit Their Jobs," https://www.cnbc.com/2019/04/03/ibm-ai-can-predict-with-95-percent-accuracy-which-employees-will-quit.html.

Cook, C., Diamond, R., and Oyer, P. 2019. "Older Workers and the Gig Economy." American Economic Review Papers and Proceedings, May, forthcoming.

Cowgill, B. 2018. "Bias and Productivity in Humans and Algorithms: Theory and Evidence from Resumé Screening." Working paper.

Cortés, M., Jaimovich, . and Siu, H. "The "End of Men" and Rise of Women in the High-Skilled labour Market", Working paper.

Council of Economic Advisors. 2010. Work-Life Balance and the Economics of Workplace Flexibility. https://obamawhitehouse.archives.gov/files/documents/ 100331-cea-economics-workplace-flexibility.pdf

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Datta, A., Tschantz, M.C. and Datta, A. 2015. "Automated Experiments on Ad Privacy Settings," *Proceedings on Privacy Enhancing Technologies*, 1: 92–112.

De Backer, K., Menon, C., Desnoyers-James, I., and Moussiegt, L. 2016. "Reshoring: Myth or Reality?", OECD Science, Technology and Industry Policy Papers, No. 27, OECD Publishing, Paris. http://dx.doi.org/10.1787/5jm56frbm38s-en.

Delaney, J. and Devereux, P.J. 2019. It's Not Just for Boys! Understanding Gender Differences in STEM. CEPR Discussion Paper 13558.

Deming, D.J. 2017. The Growing Importance of Social Skills in the labour Market. *Quarterly Journal of Economics*, 132(4) :1593-1640.

Deming, D.J. and Kahn, L. B. 2018. "Skill Requirements across Firms and labour Markets: Evidence from Job Postings for Professionals," *Journal of labour Economics* 36(S1): S337-S369.

Deming, D.J. and Noray, K.L. 2018. STEM Careers and Technological Change. NBER Working paper 25065.

EFIP and Malt, Freelancing in Europe, 2019, https://news.malt.com/wp-content/uploads/2019/02/ FREELANCING-IN-EUROPE-2-1.pdf

Faberman, J. and Kudlyak, M. 2016. "What Does Online Job Search Tell Us About the labour Market?" Economic Perspectives, Federal Reserve Bank of Chicago, No. 1, January.

Freeman, B. 2002."The Labour Market in the New Information Economy." *Oxford Review of Economic Policy*, 18(3): 288–305.

Goldin, C. and Katz, L.F. 2009. The Race Between Education and Technology. Harvard University Press.

Goos, M., Manning, A. and Salomons, A. 2009. "Job Polarization in Europe." *American Economic Review Papers & Proceedings*: 99(2), 58-63.

Goos, M., Manning, A., and Salomons, A. 2014. "Explaining Job Polarization in Europe." *American Economic Review*

Goos, M. and Salomons, A., 2019, "The Future of Work: An Agenda", Flexibility@work, Randstad.

Goos et al., 2019, "The Impact of the Digital Transformation on EU Labour Markets", Report of the High-Level Expert Group, European Commission, https://publications.europa.eu/en/publication-detail/-/publication/d9574c3a-67fb-11e9-9f05-01aa75ed71a1.

Helliwell, J. and Huang, H. 2011. "Well-Being and Trust in the Workplace," *Journal of Happiness Studies*, 12(5): 747-767.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Hoffman, M., Kahn, L.B., and Li, D. 2018. "Discretion in Hiring." *The Quarterly Journal of Economics*, 133(2): 765–800.

Hyman, B. 2018. "Can Displaced labour be Retrained? Evidence from Quasi-Random Assignment to Trade Adjustment Assistance." Working Paper.

Ibarra., I. A., Goff, L., Hernandez, D. J., Lanier, J., and Weyl, G. 2018. 'Should We Treat Data as Labor? Moving Beyond "Free"', *American Economic Review Papers & Proceedings*, May.

ILR School & Aspen Institute. https://www.gigeconomydata.org

International Federation of Robotics. 2018. World Robotics 2018 edition.

International Labour Organization. 2019. https://www.ilo.org/ilostat.

Katz, L.F., and Krueger, A.B. 2019. "The Rise and Nature of Alternative Work Arrangements in the United States, 1995–2015." *Industrial & labour Relations Review* 72(2), 382–416.

Katz, L.F., and Krueger, A.B. 2017. "The Role of Unemployment in the Rise in Alternative Work Arrangements." *American Economic Review* 107 (5): 388-92.

Karabarbounis, L., and Neiman, B. 2014. "The Global Decline of the labour Share," *The Quarterly Journal of Economics*, 129(1): 61-103.

Kässi, O., and Lehdonvirta, V. 2018. "Online labour index: Measuring the online gig economy for policy and research." *Technological Forecasting and Social Change* 137: 241–248. https://ilabour.oii.ox.ac.uk/online-labour-index/

Kleinberg, J., Lakkaraju, H., Leskovec, J., Ludwig, J. and Mullainathan, S. 2017. "Human decisions and machine predictions," National Bureau of Economic Research.

Kleinberg, J. Ludwig, J., Mullainathan, S. and Rambachan, A. 2018. "Algorithmic Fairness." *AEA Papers and Proceedings*, 108: 22–27

Kleiner, M. M. and Krueger, A. 2013. "Analyzing the Extent and Influence of Occupational Licensing on the labour Market." *Journal of labour Economics*, 31(2): pp. S173-S202

Kuhn, P. and Skuterud, M., "Internet Job Search and Unemployment Durations." *American Economic Review*, 2004, 94(1), pp.218-232.

Lambrecht, A. and Tucker, C.E. 2016. "Algorithmic Bias? An Empirical Study into Apparent Gender-Based Discrimination in the Display of STEM Career Ads." Working paper.

Leontief, W.W. 1983. "National perspective: The definition of problem and opportunity", in: National Academies, The Long-term Impact of Technology on Employment and Unemployment: A National Academy of Engineering Symposium, June 30, p. 3.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Lopez Sanchez, J.I., Grau Ruiz, M.A., Sanchez-Uran Azaña, M.Y., 2019, "The Impacts of Robotics and Computerization on the Labour Market: Inclusive Insight From a Law and Economics Perspective", Complutense University of Madrid.

Mas, A. and Pallais, A. 2017. "Valuing Alternative Work Arrangements." *American Economic Review*, 107 (12): 3722-59.

Markussen, S. and Road, K. 2019. "Economic Mobility Under Pressure." *Journal of the European Economic Association*, forthcoming.

McKay, S., Jefferys, S., Paraksevopoulou, A. and Keles, J. 2012. 'Study on precarious work and social rights', European Commission, VT/2010/084, p. 78, section 2.6, https://ec.europa.eu/social/BlobServlet?docId=7925&langId=en

McKinsey Global Institute. 2019. Globalization in Transition: The Future of Trade and Value Chains. www.mckinsey.com/mgi.

Mokyr, J., Vickers, C., and Ziebarth, N. L. 2015. "The History of Technological Anxiety and the Future of Economic Growth: Is This Time Different?" *Journal of Economic Perspectives*, 29 (3): 31-50.

OECD. 2019. The Future of Work, Employment Outlook 2019. https://doi.org/10.1787/19991266

OECD. 2019. https://data.oecd.org/emp/hours-worked.htm

OECD. 2017a. OECD Science, Technology and Industry Scoreboard 2017. http://www.oecd.org/sti/oecd-science-technology-and-industry-scoreboard-20725345.htm

OECD. 2017b. "Job Polarisation by Country: Percentage Point Change in Share of Total Employment, 1995 to 2015", in: OECD Employment Outlook 2017, OECD Publishing, Paris, https://doi.org/10.1787/empl_outlook-2017-graph39-en.

OECD. 2016. "Getting Skills Right: Assessing and Anticipating Changing Skill Needs", Chapter 4.

OECD. 2015. In It Together – Why Less Inequality Benefits All. http://www.oecd.org/social/in-it-together-why-less-inequality-benefits-all-9789264235120-en.htm

Oswald, A.J., Proto, E., and Sgroi, D. 2015. "Happiness and Productivity," *Journal of labour Economics* 33(4): 789-822.

Pesole, A., Urzí Brancati, M.C, Fernández-Macías, E., Biagi, F., González Vázquez, I. 2018. Platform Workers in Europe, Publications Office of the European Union, https://bit.ly/2N2TciX.

Posner, E. A. and Weyl, E.G. 2018. "Radical Markets: Uprooting Capitalism and Democracy for a Just Society", Princeton, NJ: Princeton University Press.

Randstad. 2017. "People to Jobs, Jobs to People." Flexibility@Work. https://workforceinsights.randstad.com/hr-research-reports-iza-research-2017

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Randstad 2018a. "Sustainable Growth in the Age of Cities." Flexibility@work. https://workforceinsights.randstad.com/hr-research-reports-flexibilitywork-2018

Randstad. 2018b. "How Will Artificial Intelligence Affect Your Talent Acquisition Strategy?" Workforce Insights, Future of Work. https://www.randstad.com/workforce-insights/hr-tech/how-will-artificial-intelligence-affect-your-talent-acquisition-strategy/

Randstad. 2018c. "Recruitbot: Why AI Will Hire Your Next Employee." Workforce Insights, Future of Work. https://www.randstad.com/workforce-insights/hr-tech/recruitbot-why-ai-will-hire-your-next-employee/

Reuters. 2018. https://uk.reuters.com/article/uk-britain-deliveroo/uk-lawmaker-to-investigate-pay-and-conditions-at-deliveroo-idUKKBN1JE1GL

Rhein, T. and Walwei, U., Forms of Employment in European Comparison, https://www.iab-forum.de/en/forms-of-employment-in-european-comparison/.

Rothwell, J. 2014. Still Searching: Job Vacancies and STEM Skills. Brookings Institution. https://www.brookings.edu/interactives/still-searching-job-vacancies-and-stem-skills/

Spreitzer, G.M., Cameron, L., and Garrett, L. 2017. Alternative Work Arrangements: Two Images of the New World of Work. *Annual Review of Organizational Psychology and Organizational Behavior* 4(1), 473-499.

Stevenson, B. 2008. "The Internet and Job Search," NBER Working Papers 13886, National Bureau of Economic Research, Inc.

The Guardian. 2019. https://www.theguardian.com/technology/2019/mar/06/facial-recognition-software-emotional-science

The New York Times. 2018. https://www.nytimes.com/2018/12/31/technology/human-resources-artificial-intelligence-humu.html

The Wall Street Journal. 2012. "Your Resumé Vs. Oblivion," https://www.wsj.com/articles/SB10001424052970204624204577178941034941330

Tirole, J. 2017. Economics for the Common Good. Princeton University Press, pp. 419-420.

Weil, D. 2014. The Fissured Workplace: Why Work Became So Bad for So Many and What Can Be Done to Improve It. Cambridge, MA: Harvard University Press.West, S.M., Whittaker, M. and Crawford, K. 2019. "Discriminating Systems: Gender, Race and Power in AI." AI Now Institute. https://ainowinstitute.org/discriminatingsystems.html.

Wood, A.J., Lehdonvirta, V. and Graham, M. 2018. "Workers of the internet unite? Online freelancer organization among remote gig-economy workers in six Asian and African countries." *New Technology, Work and Employment*, 33(2): 95-112.

World Employment Confederation. 2018. "Economic Report," https://www.wecglobal.org/fileadmin/media/pdf/WEC_Economic_Report_2018_Edition.pdf

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Yamamoto, I. 2019. "The Impact of AI and Information Technologies on Worker Stress," https://voxeu.org/article/impact-ai-and-information-technologies-worker-stress

a. 2.4. Providing an overview of financial and tax law, labour and social security law

Acemoglu, D., and Restrepo, P. (2017), "Robots and Jobs: Evidence from US Labor Markets," NBER Working Papers 23285, National Bureau of Economic Research, Inc.

Adecco-Cuatrecasas: Informe, (2018), available at https://adecco.es/wp-content/uploads/2018/02/Estudio-cualitativo-sobre-la-percepcio%CC%81n-de-la-robo%CC%81tica-industrial-en-Espan%CC%83a.pdf

Agote Eguizábal, R (2018) 'Inteligencia Artificial, Ser humano y Derecho', Revista Claves de Razón Práctica, No. 237.

Arranz de Andrés, C. (Dir.) (2017) "Aspectos fiscales de la dependencia y la discapacidad", Thomson Reuters Aranzadi, Cizur Menor.

Autor, D.H. (2019): "Work of the Past, Work of the Future", MIT, available at https://economics.mit.edu/files/16724?utm_campaign=Economic%20Studies&utm_source=hs_email&utm_medium=email&utm_content=70315554

De Asís, R. (2014): Una mirada a la Robótica desde los Derechos Humanos, Dykinson, Madrid.

Barrio Andrés, M. (Dir.) (2018) "Derecho de los Robots", La Ley - Wolters Kluwer, Madrid.

Bartel, A., Ichniowski, C., & Shaw, K. (2007). "How Does Information Technology Affect Productivity? Plant-Level Comparisons of Product Innovation, Process Improvement, and Worker Skills". Quarterly Journal of Economics, 122. pp. 1721-1758.

Barua, A., Lee, S. y Whinston, A. (1996).- "The Calculus of Reengineering". Information Systems Research. 7(4), pp.409-428.

Bharadwaj, A.S. (2000).- "A Resource-Based perspective on information technology capability and firm performance: an empirical investigation". MIS Quarterly. Marzo, pp: 160-196

Berman, E., Bound, J. y Griliches, Z (1994).- "Changes in the Demand for Skilled Labour within U.S Manufacturing : Evidence from the Annual Survey of Manufacturers". Quarterly Journal of Economics, 109(2), pp.367-397

Birkinshaw, J.; Mol, M. (2006).- "How management innovation happens.". MIT Sloan Review, 47(4), 81-88.

Black, S. y Lynch, L. (2001).- "How to compete: the impact of workplace practices and information technology on productivity". Working paper 6120. National Bureau of Economic Research. 2001.

Bloom, N.; Garicano, L.; Sadun, R.; Van Reenen, J. (2014).- "The distinct effects of information technology and communication technology on firm organization". Management Science, 60(12), 2859-2885.

Bresnahan, T.E., Brynjolfsson, E. y Hitt, L.M. (2002).- "Information, Technology and Information Worker Productivity: Task Level Evidence". Quarterly Journal of Economics, 2002 - MIT Press 117(1), pp: 339-376.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Brynjolfsson, E, Hitt, L (1996).- "Paradox Lost? Firm-Level Evidence of High Returns to Information Systems Spending". Management Science, 42. pp. 541-558.

Brynjolfsson, E. Yang, S (1996).- "Information Technology and Productivity : A Review of the Literature". Advances in computers, 43. pp. 179-214.

Brynjolfsson, E; Saunders, A. (2009). "Wired for Innovation." Cambridge MA

Brynjolfsson, E, Mcafee, A. (2014): The Second Machine Age: Work, Progress, and Prosperity in a Time of Brilliant Technologies, W. W. Norton & Company, New York.

Calo,, R., Froomkin, A.M. and Kerr, I (Eds.) (2016) "Robot Law", Edward Elgar, Cheltenham.

Calo, R. (2016): 'La robótica y las lecciones del Derecho cibernético', Revista Privacidad y Derecho Digital, No. 2.

Cardona, M., Kretschmer, T. y Strobel, T. (2013).- "ICT and Productivity: conclusions from the empirical literature". Information Economics and Policy. 25(2013). Pp.109-125.

Chouakria–Douzal, A., and Nagabhushan, P. N. (2007). "Adaptive Dissimilarity Index for Measuring Time Series Proximity". Advances in Data Analysis and Classification, 1(1): 5-21.

Comisión Europea (2018): «En ruta hacia la movilidad automatizada: estrategia de la UE para la movilidad del futuro», COM/2018/283 final, Bruselas, 17.5.2018.

Cortina, A.: (2018): 'Ciudadanía digital y dignidad humana', El País, 26 March 2018.

Dauth, W., Findeisen, S., Suedekum, J. and Woessner, N. (2017). "German Robots – The Impact of Industrial Robots on Workers". Discussion Paper. Revised Version December 18, 2017.

Doménech, R., García, J.R., Montañez, M. and Neut, A. (2018). "Afectados por la-revolución digital El caso de España". Papeles De Economía Española, 156: 128-45.

Doménech, R., García, J.R., Montañez, M. and Neut, A. (2018): ¿Cuán vulnerable es el Empleo en España a la revolución digital?, BBVA, Madrid, March 2018.

Avalaible at https://www.bbvaresearch.com/wp-content/uploads/2018/03/Cuan-vulnerable-es-el-empleo-en-Espana-a-la-revolucion-digital.pdf

Fernández Amor, J.A: (2018): «Derecho tributario y cuarta revolución industrial: análisis jurídico sobre aspectos fiscales de la robótica», Nueva Fiscalidad, N.º 1, enero-marzo, pp. 47 y ss.

Frey, C. B., and Osborne, M. A. (2017). "The future of employment: How susceptible are jobs to computerisation?," Technological Forecasting and Social Change, 114: 254-80

FTI Consulting (2018): The global policy response to AI Report, February, available at http://brussels.ftistratcomm.com/wp-content/uploads/sites/5/2018/02/The-global-policy-response-to-AI-snapshot.pdf

García Mexía, P.: 'Entes Capaces No Personales. ¿Hacia una personalidad para los robots', available at https://www.automatas.tech/pablo-garcia-mexia-colaboracion.html

García-Prieto Cuesta, J.: '¿Qué es un Robot?, in BARRIO ANDRÉS, M. (Dir.): Derecho de los Robots, La Ley - Wolters Kluwer, Madrid, 2018.

García Novoa, C. (2018): «Impuestos atípicos en la era post BEPS» (Capítulo 8), en CUBERO TRUYO, A. (Dir.): Tributos asistemáticos del ordenamiento vigente, Tirant lo Blanch, Valencia.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Garicano L (2000) Hierarchies and the organization of knowledge in production. Journal Political Econom. 108(5):874–904.

Garicano, L. Rossi-Handsberg, E. (2006).- "Organization and Inequality in a knowledge economy". NBER Working Papers 11458, National Bureau of Economic Research, Inc.

Graetz, G. and Michaels, G. (2017) "Robots at Work," CEP Discussion Paper No 1335. Revised Version November 11, 2017.

Grau Ruiz, Mª.A. (2017): 'La adaptación de la fiscalidad ante los retos jurídicos, económicos, éticos y sociales planteados por la robótica', Nueva fiscalidad, No. 4.

Gupta, S., Keen, M, Shah, A. and Verdier, G. (Eds.) (2017) "Digital revolutions in public finance", International Monetary Fund, Washington DC, 2017. Available at http://www.elibrary.imf.org/view/IMF071/24304-9781484315224/24304-9781484315224/Other_formats/Source_PDF/24304-9781484316719.pdf

Hawksworth, J., Berriman, R., and Goel, S. (2018). "Will robots really steal our jobs? An international analysis of the potential long term impact of automation". Price Waterhouse Coopers Report.

Holder, C., Khurana, V.,, Hook, J, Bacon, G. and Day, R. (2016): ' Robotics and law: Key legal and regulatory implications of the robotics age (part II of II)', Computer law & Security Review, No. 32, 2016.

IBA, Global Employment Institute (2017): 'Artificial Intelligence and Robotics and Their Impact on the Workplace', April 2017.

ILO: The future of work, available at http://www.ilo.org/global/topics/future-of-work/lang--es/index.htm; a summary of the national dialogues is available at http://www.ilo.org/global/topics/future-of-work/WCMS_591507/lang--es/index.htm

International Federation of Robotics (IFR (2017) "The impact of Robots on Productivity, Employment and Jobs", April 2017; available at https://ifr.org/ifr-press-releases/news/position-paper

International Federation of Robotics (2017). "World Robotics: Industrial Robots 2017", Technical report.

International Monetary Found, (2017): 'Tackling Inequality', IMF Fiscal Monitor, octubre 2017, p. 3-4. Available at http://www.imf.org/en/Publications/FM/Issues/2017/10/05/fiscal-monitor-october-2017

Keynes, J. M. (1930) "Economic Possibilities for our Grandchildren". In Essays in Persuasion (New York: Harcourt Brace): 358-373

Leenes, R., Palmerini, E., Koops, B-J., Bertolini, A., Salvini, P. and Lucivero, F. (2017): 'Regulatory challenges of robotics: some guidelines for addressing legal and ethical issues' , Law, Innovation and Technology Review, Vol. 9, No. 1, 2017

Leontief, W. (1952) "Machines and Man," Scientific American, 187(3): 150-64.

López Sánchez, J.I.; Rodríguez Duarte, A.; Sandulli, F.D. (2010).- "Adopción y Uso de las TIC. Hacia un Modelo para las Pymes españolas", Estudio Núm38, Fundación COTEC, Madrid. ISBN: 978-84-95336-97-2

McKinsey Global Institute - MGI (2017). "Jobs lost, jobs gained: workforce transitions in a time of automation". Technical Report.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

McKinsey Global Institute (2017): 'A future that works: automation, employment and productivity' 2017, available at https://www.mckinsey.com/mgi/overview/2017-in-review/automation-and-the-future-of-work/a-future-that-works-automation-employment-and-productivity

Mckinsey & Company (2016): Car data: paving the way to value-creating mobility. Perspectives on a new automotive business model, Advanced Industries, marzo de 2016. [Documento electrónico disponible en: https://www.mckinsey.com/industries/automotive-and-assembly/our-insights/creating-value-from-car-data]

Mercader Uguina, J.R. (2017): El futuro del trabajo en la era de la digitalización y la robótica; tirant lo blanch, Valencia.

Milgrom, P., Roberts, J. (1995).- "Complementarities and Fit : Strategy, Structure and Organizational Change in Manufacturing". Journal of Accounting and Economics. 19(2), pp. 179-208.

Montero, P. and Vilar, J. A. (2014). "TSclust: An R Package for Time Series Clustering". Journal of Statistical Software, 62 (1): 1-43.

Nelson, R.R. y Winter, S.G. (1982).- "An evolutionary theory of economic change". Cambridge, MA: Belknap Press.

Nieto, J (2017) 'El Futuro del trabajo que queremos y el Derecho del Trabajo', Ius Labor, No. 3, 2017, available at https://www.upf.edu/documents/3885005/140470042/1.+Editorial.pdf/406c3008-6ef9-7ed6-f4ee-8f754c9adc31

Nisa Ávila, J.A. (2016).: 'Robótica e Inteligencia Artificial ¿legislación social o nuevo ordenamiento jurídico', El Derecho.com, Lefebvre, http://tecnologia.elderecho.com/tecnologia/internet_y_tecnologia/Robotica-Inteligencia-Artificial-legislacion-social-nuevo-ordenamiento_11_935305005.html, (consulted in January 2019).

OECD-International Transport Forum (2018): Safer Roads With Automated Vehicles?, 2018 [Documento electrónico disponible en: https://www.itf-oecd.org/safer-roads-automated-vehicles-0]

O'Mahoney M, Van Ark B (2003). "EU productivity and competitiveness: An industry perspective: Can Europe resume the catching-up process?." Office for official publications of the European communities. Luxemburg.

Palmarini, E. (2017): «Robótica y derecho: sugerencias, confluencias, evoluciones en el marco de una investigación europea», Revista de Derecho Privado, Universidad Externado de Colombia, n.º 32, enero-junio de 2017, pp. 53-97. DOI: https://doi.org/10.18601/01234366.n32.03

PwC (2018): Will robots really steal our jobs? An international analysis of the potential long term impact of automation, available at https://www.pwc.es/es/publicaciones/tecnologia/assets/international-impact-of-automation-2018.pdf

Del Rey Guanter, S. (Dtor), (2018): Robótica y su impacto en los Recursos Humanos y en el marco regulatorio de las Relaciones Laborales, La Ley - Wolters Kluwer, Madrid.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Richards, N.M. and Smart, D. (2016): "How sould the Law think about Robots', in CALO et al. : Robot Law, Edward Eldgar, Cheltenham.

Sánchez-Urán Azaña, Mª Y. y Grau Ruiz, Mª A.: «El impacto de la robótica, en especial la robótica inclusiva, en el trabajo: aspectos jurídicos-laborales y fiscales», Eprint UCM, 2018. [Documento electrónico disponible en: https://eprints.ucm.es/47523/]

Sánchez-Urán Azaña, Mª Y y Grau Ruiz, Mª A.: (2018) «Robotics and Work: Labor and Tax regulatory Framework», en International Congress Technological Innovation and Future of Work, 5 y 6 de abril de 2018, Santiago de Compostela. Eprint UCM, 2018. [Documento electrónico disponible en: https://eprints.ucm.es/47718/].

Schwab, K. (editor) (2017). "The Global Competitiveness Report 2017–2018". World Economic Forum Insight Report. ISBN-13: 978-1-944835-11-8.

Segura Alastrué, M. (2018): 'Los robots en el Derecho Financiero y Tributario' (Chapter VII), in BARRIO ANDRÉS, M. (dir.): Derecho de los robots, La Ley -Wolters Kluwer, Madrid.

Slater, S.F. y Naver, J.C. (1995).- "Market Orientation and the learning organization".Journal of Marketing, 59(3), 63-75.

Standing, g. (2013): El Precariado: una nueva clase social, Edit. Pasado y Presente, Barcelona.

United Nations (2017): Trade and Development Report, available at http://unctad.org/en/PublicationsLibrary/tdr2017_en.pdf

Veruggio, G. (2006): The EURON Roboethics Roadmap, available at http://www3.nd.edu/~rbarger/ethics-roadmap.pdf.

VEruggio, G. and Operto, F. (2006) 'Roboethics: a bottom-up interdisciplinary discourse in the field of applied ethics in robotics', International Review of Information Ethics, Vol.6, No.12, 2006.

VEruggio, G. and Operto, F. (2008), "Roboethics: Social and Ethical Implications of Robotics', in Handbook of Robotics, Siciliano, B. y Khatib, O. (eds.), 2008.

Villar Ezcurra, M. (ed.) (2017): State Aids, Taxation and the Energy Sector, Thomson Reuters Aranzadi, Cizur Menor, 2017.

Walsh, J.P; Ungson, G.R. (1991).- "Organizational memory". Academy of Management Review, 16, 57-91

Waters, R. (2017): 'Do androids dream of personal deductions?', The big read, Financial Times, 25 February 2017, available at https://www.ft.com/content/597fff44-fa78-11e6-9516-2d969e0d3b65

Waytz, A. and Norton, M.I.: (2014): '*Botsourcing and Outsourcing: Robot, British, Chinese, and German Workers Are for Thinking - Not Feeling - Jobs*', *Emotion,* Vol. 14, No. 2, ;available at https://www.hbs.edu/faculty/Publication%20Files/waytz%20norton_a358958c-3b94-4f25-bb8c-7d10605738d8.pdf

NATIONAL REPORTS:

- France 2017, Ministère Économique et Financiers: 'France Intelligence Artificielle'.

- Italy 2017, Senato della Republica: 'Impatto sul mercato del lavoro della quarta rivoluzione industriale'.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

- Japan 2015, Headquarters for Japan's economic revitalization: 'New Robot Strategy. Japan's Robot Strategy. Vision, Strategy, Action Plan'.

- Spain 2017, Consejo Económico y Social de España (CES): 'Digitalización de la Economía'.

- United Kigdom 2016, House of Commons: 'Robotics and AI'.

- USA 2016, Executive Office of the President, National Science and Technology Council, Committee on Technology: 'Preparing for the Future of AI'.

2.5. Assessment of human acceptance and ethical concerns of interactive robotics

References on Conceptual Analysis: technology, machine and robot

L. Floridi, Ed., *The Onlife Manifesto: Being Human in a Hyperconnected Era*. Springer International Publishing, 2015.

J. Hernández-Orallo, "The Measure of All Minds by José Hernández-Orallo," *Cambridge Core*. Jan-2017.

S. Lloyd, *Programming the Universe: A Quantum Computer Scientist Takes on the Cosmos*. Vintage Books, 2007

E. Machery, *Philosophy Within Its Proper Bounds*. Oxford University Press, 2017.

A. Mayor, *Gods and Robots: Myths, Machines, and Ancient Dreams of Technology*. Princeton University Press, 2018

AI HLEG, 2019, "High-Level Expert Group on Artificial Intelligence: Ethics Guidelines for Trustworthy AI". *European Commission, 09.04.2019*. https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai

Bryson, Joanna J, 2019, "The Past Decade and Future of Ai's Impact on Society". In Anonymous (ed.), *Towards a New Enlightenment: A Transcendent Decade*. Madrid: Turner - BVVA. https://www.bbvaopenmind.com/en/books/towards-a-new-enlightenment-a-transcendent-decade/

Calo, Ryan, 2018, "Artificial Intelligence Policy: A Primer and Roadmap ". *University of Bologna Law Review, 3* (2), 180-218.

Crawford, Kate, and Ryan Calo, 2016, "There Is a Blind Spot in AI Research". *Nature, 538*, 311-313.

European Group on Ethics in Science and New Technologies, 2018 (09.03.2018). "Statement on Artificial Intelligence, Robotics and 'Autonomous' Systems".

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

European Commission, Directorate-General for Research and Innovation, Unit RTD.01. http://ec.europa.eu/research/ege/pdf/ege_ai_statement_2018.pdf

Floridi, Luciano, Josh Cowls, Monica Beltrametti, Raja Chatila, Patrice Chazerand, Virginia Dignum, . . . Effy Vayena, 2018, "Ai4people—an Ethical Framework for a Good AI Society: Opportunities, Risks, Principles, and Recommendations". *Minds and Machines, 28* (4), 689-707.

Gibert, Martin, 2019, "Éthique Artificielle (Version Grand Public)". In Kristanek, M (ed.), *Encyclopédie Philosophique*. http://encyclo-philo.fr/etique-artificielle-gp/

Giubilini, Alberto, and Julian Savulescu, 2018, "The Artificial Moral Advisor: the "Ideal Observer" Meets Artificial Intelligence". *Philosophy & Technology, 31* (2), 169-188. https://doi.org/10.1007/s13347-017-0285-z

Gomila, Antoni, and Vincent C. Müller, 2012, "Challenges for Artificial Cognitive Systems". *Journal of Cognitive Science, 13* (4), 453-469. http://doi.org/10.17791/jcs.2012.13.4.453

Goodfellow, Ian, Yoshua Bengio, and Aaron Courville, 2016, *Deep Learning*. Cambridge, Mass.: MIT Press.

Greene, Daniel, Anna Lauren Hoffmann, and Luke Stark, forthcoming. "Better, Nicer, Clearer, Fairer: A Critical Assessment of the Movement for Ethical Artificial Intelligence and Machine Learning". *52nd Hawaii International Conference on Systems Science.* http://dmgreene.net/research

IEEE, Global Initiative on Ethics of Autonomous and Intelligent Systems, 2019, "Ethically Aligned Design: A Vision for Prioritizing Human Well-Being with Autonomous and Intelligent Systems". *25.03.2019* (1st ed.), 1-294. https://standards.ieee.org/content/ieee-standards/en/industry-connections/ec/autonomous-systems.html

Jacobs, An, Lynn Tytgat, Michel Maus, Romain Meeusen, and Bram Vanderborght (eds.), 2019, *Homo Roboticus: 30 Questions and Answers on Man, Technology, Science & Art*. Brussels: ASP. http://homo-roboticus.be

Johnson, Deborah G, and Mario Verdicchio, 2017, "Reframing AI Discourse". *Minds and Machines, 27* (4), 575–590. https://link.springer.com/article/10.1007/s11023-017-9417-6 - citeas

Minsky, Marvin, 1985, *The Society of Mind*. New York: Simon & Schuster.

Müller, Vincent C. (ed.), forthcoming, *Oxford Handbook of the Philosophy of Artificial Intelligence*. New York: Oxford University Press.

Shoham, Yoav, Perrault Raymond, Brynjolfsson Erik, Jack Clark, James Manyika, Juan Carlos Niebles, . . . Zoe Bauer, 2018 (December 2018). "The AI Index 2018

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Annual Report". *AI Index Steering Committee, Human-Centered AI Initiative.* http://cdn.aiindex.org/2018/AI Index 2018 Annual Report.pdf

Silver, David, Thomas Hubert, Julian Schrittwieser, Ioannis Antonoglou, Matthew Lai, Arthur Guez, . . . Demis Hassabis, 2018, "A General Reinforcement Learning Algorithm That Masters Chess, Shogi, and Go through Self-Play". *Science, 362* (6419), 1140-1144. http://science.sciencemag.org/content/sci/362/6419/1140.full.pdf

Stahl, Bernd Carsten, Job Timmermans, and Brent Daniel Mittelstadt, 2016, "The Ethics of Computing: A Survey of the Computing-Oriented Literature". *ACM Computing Surveys, 48/4* (55), 1-38.

Stone, Peter, Rodney Brooks, Erik Brynjolfsson, Ryan Calo, Oren Etzioni, Greg Hager, . . . Astro Teller, 2016. "Artificial Intelligence and Life in 2030. One Hundred Year Study on Artificial Intelligence: Report of the 2015-2016 Study Panel". https://ai100.stanford.edu/2016-report

Taddeo, Mariarosaria, and Luciano Floridi, 2018, "How AI Can Be a Force for Good". *Science, 361* (6404), 751-752. http://science.sciencemag.org/content/sci/361/6404/751.full.pdf

Taylor, Steve, Brian Pickering, Michael Boniface, Michael Anderson, David Danks, Asbjørn Følstad, . . . Fiona Wollard, 2018 (June 2018). "Responsible AI – Key Themes, Concerns & Recommendations for European Research and Innovation: Summary of Consultation with Multidisciplinary Experts". https://www.hub4ngi.eu

Trump, Donald J, 2019, "Executive Order on Maintaining American Leadership in Artificial Intelligence". *The White House* (11.02.2019). https://www.whitehouse.gov/presidential-actions/executive-order-maintaining-american-leadership-artificial-intelligence/

Walsh, Toby, 2018, *Machines That Think: The Future of Artificial Intelligence*. Amherst, Mass.: Prometheus Books.

Whittlestone, Jess, Rune Nyrup, Anna Alexandrova, Kanta Dihal, and Stephen Cave, 2019, "Ethical and Societal Implications of Algorithms, Data, and Artificial Intelligence: A Roadmap for Research". 1-59. https://www.adalovelaceinstitute.org/nuffield-foundation-publishes-roadmap-for-ai-ethics-research/

12. Other Internet Resources

Research organizations:

Turing Institute (UK)

https://www.turing.ac.uk/media/news/alan-turing-institute-data-ethics-group/

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

AI Now (at NYU)

https://ainowinstitute.org/

Leverhulme Centre for the Future of Intelligence

http://lcfi.ac.uk/

Future of Humanity Institute

https://www.fhi.ox.ac.uk/

Future of Life Institute

https://futureoflife.org/

Stanford Center for Internet and Society

http://cyberlaw.stanford.edu/

Berkman Klein Center

https://cyber.harvard.edu

Digital Ethics Lab (Oxford)

http://digitalethicslab.oii.ox.ac.uk

Open Roboethics Institute

http://www.openroboethics.org/

Conferences:

Philosophy & Theory of AI

https://www.pt-ai.org/

Ethics and AI 2017

https://philevents.org/event/show/35634

FAT 2018

https://www.fatconference.org

AIES

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

http://www.aies-conference.com/

We Robot 2018

https://conferences.law.stanford.edu/werobot/

Robophilosophy

http://conferences.au.dk/robo-philosophy/

Policy Documents:

EUrobotics TG 'robot ethics' collection of policy documents

http://www.pt-ai.org/TG-ELS/policy

Bibliography:

PhilPapers section 'Robot Ethics'

https://philpapers.org/browse/robot-ethics

PhilPapers section 'Ethics of Artificial Intelligence'

https://philpapers.org/browse/ethics-of-artificial-intelligence

References on Cartography of the Values Involved in Robotics

Clarkson, J., Coleman, R., Keates, S. and Lebbon, C. (eds.) (2003). *Inclusive Design. Design for the whole population*. London: Springer.

J. Echeverría, "Ciencia y valores: Propuestas para una axionomía de la ciencia," *Contrastes. Revista Internacional de Filosofía*, 1998.

J. Echeverría*, Ciencia y valores*. Destino Barcelona, 2002.

J. Echeverría, *La revolución tecnocientífica*. Fondo de Cultura Económica de España, 2003.

M. Á. Quintanilla Fisac, "Engaging Technologies: Criteria for an Alternative Model of Technological Development," in *Spanish Philosophy of Technology: Contemporary Work from the Spanish Speaking Community*, B. Laspra and J. A. López Cerezo, Eds. Cham: Springer International Publishing, 2018, pp. 103–123.

Tewey, B. (1997). *Building Participatory Action Research Partnerships in Disability and Rehabilitation Research*. Washington, DC.: U.S. Department of Education, National Institute on Disability and Rehabilitation Research.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

M. Toboso, "Rethinking disability in Amartya Sen's approach: ICT and equality of opportunity", *Ethics and Information Technology*, 13(2), 2010, pp. 107-118.

L. Winner, "Do Artifacts Have Politics?" *Daedalus*, vol. 109, no. 1, pp. 121–136, 1980.

References on Discursive Frameworks for the Development of Inclusive Robotics

Convention on the Rights of the Child, New York, UN, 1989.

Convention on the Rights of Persons with Disabilities. New York, UN, 2006.

A. Cortina, *Ciudadanos del mundo. Hacia una teoría de la ciudadanía,* Madrid: Alianza Editorial, 2005.

European Group on Ethics in Science and New Technologies, "Artificial Intelligence, Robotics and 'Autonomous' Systems", Luxembourg: Publications Office of the European Union, 2018.

J. Habermas, "El concepto de dignidad humana y la utopía realista de los derechos humanos" en *La constitución de Europa*, Madrid: Trotta, 2012.

A. Honneth, *El derecho de la libertad. Esbozo de una eticidad democrática.* Madrid: Katz, 2014.

E. F. Kittay, *Love's Labor: Essays on Women, Equality, and Dependency*, Nueva York, Routledge, 1999.

M. Nussbaum, *Las fronteras de la justicia*, Barcelona: Paidós, 2007.

M. C. Nussbaum & A. Sen (Eds.), *The quality of life,* New York: Oxford University Press, 1993.

A. Ortega, *La imparable marcha de los robots,* Madrid: Alianza Editorial, 2016.

A. Palacios, *El modelo social de la discapacidad: orígenes, caracterización y plasmación en la Convención Internacional sobre los Derechos de las Personas con Discapacidad*, Madrid: Ediciones Cinca, 2008.

L. Rojas Marcos, *Convivir. El laberinto de las relaciones de pareja, familiares y laborales,* Madrid, Aguilar, 2008.

J. Romañach y M. Lobato, "Functional diversity. A new term in the struggle for dignity in the diversity of the human being", Foro de Vida Independiente y Divertad (Spain), 2005. https://disability-studies.leeds.ac.uk/library/

J. Tronto, *Caring Democracy. Markets, Equality, and Justice*, New York University Press, 2013.

Unesco, "Report_of_COMEST_on_robotics_ethics", 2017.

L. Winner, "Is there a right to shape technology?", *Argumentos de Razón Técnica*, nº 10, 2007, pp. 305–328.

References on Socially assistive robots for people with intellectual and developmental disabilities

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Allen, A. (2016). Privacy and Medicine. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Winter 2016). Retrieved from https://plato.stanford.edu/archives/win2016/entries/privacy-medicine/

Beauchamp, T. L., & Childress, J. F. (2009). *Principles of Biomedical Ethics* (6th ed.). Oxford: Oxford University Press.

Borenstein, J., & Arkin, R. (2016). Robotic Nudges: The Ethics of Engineering a More Socially Just Human Being. *Science and Engineering Ethics*, *22*(1), 31–46. https://doi.org/10.1007/s11948-015-9636-2

Bradshaw, J., Saling, M., Hopwood, M., Anderson, V., & Brodtmann, A. (2004). Fluctuating cognition in dementia with Lewy bodies and Alzheimer's disease is qualitatively distinct. *Journal of Neurology, Neurosurgery, and Psychiatry*, *75*(3), 382–387. https://doi.org/10.1136/jnnp.2002.002576

Chalghoumi, H., Cobigo, V., Dignard, C., Gauthier-Beaupré, A., Jutai, J., Lachapelle, Y., … Perrin, M. (2017). Information Privacy for Technology Users with Intellectual and Developmental Disabilities: Why Does It Matter? *Ethics & Behavior*.

Charland, L. C. (2015). Decision-Making Capacity. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2015). Retrieved from https://plato.stanford.edu/archives/fall2015/entries/decision-capacity/

Coeckelbergh, M. (2012). Are Emotional Robots Deceptive? *IEEE Transactions on Affective Computing*, *3*(4), 388–393. https://doi.org/10.1109/T-AFFC.2011.29

Coeckelbergh, Mark, Pop, C., Simut, R., Peca, A., Pintea, S., David, D., & Vanderborght, B. (2016). A Survey of Expectations About the Role of Robots in Robot-Assisted Therapy for Children with ASD: Ethical Acceptability, Trust, Sociability, Appearance, and Attachment. *Science and Engineering Ethics*, *22*(1), 47–65. https://doi.org/10.1007/s11948-015-9649-x

Conti, D., Di Nuovo, S., Buono, S., & Di Nuovo, A. (2017). Robots in Education and Care of Children with Developmental Disabilities: A Study on Acceptance by Experienced and Future Professionals. *International Journal of Social Robotics*, *9*(1), 51–62. https://doi.org/10.1007/s12369-016-0359-6

Council of Europe. (1997). *Convention for the protection of Human Rights and Dignity of the Human Being with regard to the Application of Biology and Medicine: Convention on Human Rights and Biomedicine*. Retrieved from http://www.coe.int/web/conventions/full-list

Diehl, J. J., Schmitt, L. M., Villano, M., & Crowell, C. R. (2012). The Clinical Use of Robots for Individuals with Autism Spectrum Disorders: A Critical Review. *Research in Autism Spectrum Disorders*, *6*(1), 249–262. https://doi.org/10.1016/j.rasd.2011.05.006

Esteban, P. G., Baxter, P., Belpaeme, T., Billing, E., Cai, H., Cao, H.-L., … Ziemke, T. (2017). How to Build a Supervised Autonomous System for Robot-Enhanced Therapy for

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Children with Autism Spectrum Disorder. *Paladyn, Journal of Behavioral Robotics*, *8*(1), 18–38. https://doi.org/10.1515/pjbr-2017-0002

European Union. (2012). *Charter of Fundamental Rights of the European Union*. European Union.

Eyal, N. (2012). Informed Consent. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2012). Retrieved from http://plato.stanford.edu/archives/fall2012/entries/informed-consent/

Feil-Seifer, D., & Matarić, M. J. (2009). Toward Socially Assistive Robotics for Augmenting Interventions for Children with Autism Spectrum Disorders. In O. Khatib, V. Kumar, & G. J. Pappas (Eds.), *Experimental Robotics* (pp. 201–210). Springer Berlin Heidelberg.

Fields, L. M., & Calvert, J. D. (2015). Informed consent procedures with cognitively impaired patients: A review of ethics and best practices. *Psychiatry and Clinical Neurosciences*, *69*(8), 462–471. https://doi.org/10.1111/pcn.12289

Jaworska, A. (2017). Advance Directives and Substitute Decision-Making. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Summer 2017). Retrieved from https://plato.stanford.edu/archives/sum2017/entries/advance-directives/

Lehmann, H., Iacono, I., Dautenhahn, K., Marti, P., & Robins, B. (2014). Robot companions for children with down syndrome: A case study. *Interaction Studies*, *15*(1), 99–112. https://doi.org/info:doi/10.1075/is.15.1.04leh

Liu, C., Conn, K., Sarkar, N., & Stone, W. (2008). Online affect detection and robot behavior adaptation for intervention of children with autism. *IEEE Transactions on Robotics*, *24*(4), 883–896. https://doi.org/10.1109/TRO.2008.2001362

MacIntyre, A. (1999). *Dependent Rational Animals: Why Human Beings Need the Virtues*. London: Duckworth.

Maulik, P. K., Mascarenhas, M. N., Mathers, C. D., Dua, T., & Saxena, S. (2011). Prevalence of intellectual disability: a meta-analysis of population-based studies. *Research in Developmental Disabilities*, *32*(2), 419–436. https://doi.org/10.1016/j.ridd.2010.12.018

Ménard, J.-F. (2010). A 'Nudge' for Public Health Ethics: Libertarian Paternalism as a Framework for Ethical Analysis of Public Health Interventions? *Public Health Ethics*, *3*(3), 229–238. https://doi.org/10.1093/phe/phq024

Moyle, W., Cooke, M., Beattie, E., Jones, C., Klein, B., Cook, G., & Gray, C. (2013). Exploring the effect of companion robots on emotional expression in older adults with dementia: a pilot randomized controlled trial. *Journal of Gerontological Nursing*, *39*(5), 46–53. https://doi.org/10.3928/00989134-20130313-03

Mueller, V. (Forthcoming). *Ethics of Artificial Intelligence and Robotics*.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Nissenbaum, H. (2009). *Privacy in Context: Technology, Policy, and the Integrity of Social Life*. Stanford, CA, USA: Stanford University Press.

O'Brolcháin, F. (2017). Robots and people with dementia: Unintended consequences and moral hazard. *Nursing Ethics*, 0969733017742960. https://doi.org/10.1177/0969733017742960

O'Brolcháin, F. (2018). Autonomy Benefits and Risks of Assistive Technologies for Persons With Intellectual and Developmental Disabilities. *Frontiers in Public Health*, *6*. https://doi.org/10.3389/fpubh.2018.00296

Peca, A., Simut, R., Pintea, S., Costescu, C., & Vanderborght, B. (2014). How do typically developing children and children with autism perceive different social robots? *Computers in Human Behavior*, *41*, 268–277. https://doi.org/10.1016/j.chb.2014.09.035

Rabbitt, S. M., Kazdin, A. E., & Scassellati, B. (2015). Integrating socially assistive robotics into mental healthcare interventions: Applications and recommendations for expanded use. *Clinical Psychology Review*, *35*, 35–46. https://doi.org/10.1016/j.cpr.2014.07.001

Robot Revolution - Alzheimer's and Dementia. (2014, April 22). Retrieved 14 December 2016, from Alz Live website: http://alzlive.com/elder-care/home/robot-roundup/

Robson, A. (2018). Intelligent machines, care work and the nature of practical reasoning. *Nursing Ethics*, 0969733018806348. https://doi.org/10.1177/0969733018806348

Sartorato, F., Przybylowski, L., & Sarko, D. K. (2017). Improving therapeutic outcomes in autism spectrum disorders: Enhancing social communication and sensory processing through the use of interactive robots. *Journal of Psychiatric Research*, *90*, 1–11. https://doi.org/10.1016/j.jpsychires.2017.02.004

Scassellati, B., Admoni, H., & Matarić, M. (2012). Robots for use in autism research. *Annual Review of Biomedical Engineering*, *14*, 275–294. https://doi.org/10.1146/annurev-bioeng-071811-150036

Shamsuddin, S., Yussof, H., Mohamed, S., & Hanapiah, F. A. (2014). Design and Ethical Concerns in Robotic Adjunct Therapy Protocols for Children with Autism. *Procedia Computer Science*, *42*, 9–16. https://doi.org/10.1016/j.procs.2014.11.027

Tapus, A., Fasola, J., & Mataric, M. J. (2008). Socially assistive robots for individuals suffering from dementia. *In ACM/IEEE 3rd Human-Robot Interaction International Conference, Workshop on Robotic Helpers: User Interaction, Interfaces and Companions in Assistive and Therapy Robotics*.

Thaler, R. H., & Sunstein, C. R. (2008). *Nudge: Improving Decisions about Health, Wealth, and Happiness*. Yale University Press.

The United Nations. (1948). *Universal Declaration of Human Rights*. United Nations.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Trachsel, M., Hermann, H., & Biller-Andorno, N. (2015). Cognitive fluctuations as a challenge for the assessment of decision-making capacity in patients with dementia. *American Journal of Alzheimer's Disease and Other Dementias*, *30*(4), 360–363. https://doi.org/10.1177/1533317514539377

UNESCO, (United Nation's Educational Scientific and Cultural Organization). (2005). *Universal Declaration on Bioethics and Human Rights*. Retrieved from http://portal.unesco.org/en/ev.php-URL_ID=31058&URL_DO=DO_TOPIC&URL_SECTION=201.html

United Nations. (2006a). *Convention on the Rights of People with Disabilities*. New York: United Nations.

United Nations. (2006b). *Convention on the rights of persons with disabilities*. Retrieved from http://www.un.org/disabilities/convention/conventionfull.shtml

Valentí Soler, M., Agüera-Ortiz, L., Olazarán Rodríguez, J., Mendoza Rebolledo, C., Pérez Muñoz, A., Rodríguez Pérez, I., … Martínez Martín, P. (2015). Social robots in advanced dementia. *Frontiers in Aging Neuroscience*, *7*. https://doi.org/10.3389/fnagi.2015.00133

Vitelli, R. (2013). Can Robots Help Care for the Elderly? Retrieved 22 March 2017, from Psychology Today website: http://www.psychologytoday.com/blog/media-spotlight/201306/can-robots-help-care-the-elderly

Wagner, A. R., & Arkin, R. C. (2011). Acting Deceptively: Providing Robots with the Capacity for Deception. *International Journal of Social Robotics*, *3*(1), 5–26. https://doi.org/10.1007/s12369-010-0073-8

Wasserman, D., Asch, A., Blustein, J., & Putnam, D. (2013). Cognitive Disability and Moral Status. In E. N. Zalta (Ed.), *The Stanford Encyclopedia of Philosophy* (Fall 2013). Retrieved from http://plato.stanford.edu/archives/fall2013/entries/cognitive-disability/

World Health Organisation, (WHO). (2019). Definition: intellectual disability. http://www.euro.who.int/en/health-topics/noncommunicable-diseases/mental-health/news/news/2010/15/childrens-right-to-family-life/definition-intellectual-disability

Yun, S.-S., Kim, H., Choi, J., & Park, S.-K. (2016). A robot-assisted behavioral intervention system for children with autism spectrum disorders. *Robotics and Autonomous Systems*, *76*, 58–67. https://doi.org/10.1016/j.robot.2015.11.004

Robotics as an instrument for social mediation

References

R. de Asís Roig, "Ethics and Robotics. A First Approach", *The Age of Human Rights Journal*, Nº 2, 2014, pp. 1-24.

J. Echeverría, *Innovation and Values: A European Perspective*, Reno, NV, University of Nevada Reno, Center of Basque Studies, 2014.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

M. Goos, "The impact of technological progress on labour markets: policy challenges", *Oxford Review of Economic Policy*, Volume 34, Issue 3, 2018, pp. 362–375.

M. C. Nussbaum & A. Sen (Eds.), *The Quality of Life,* New York: Oxford University Press, 1993.

J. Romero, "CiberÉtica como ética aplicada: una introducción", *Dilemata. Revista Internacional de Éticas Aplicadas*, 24, 2017, pp. 45-63.

J. A. Schumpeter, *Theorie der wissenschaftlichen Entwicklung*, Berlin, Duncker & Humblot, 1912: english translation, *The Theory of Economic Development*, Boston, Harvard University Press, 1934.

A. Sen, "The Standard of Living", in G. Hawthorn (Ed.), *The Standard of Living*. Cambridge: Cambridge University Press, 1987.

A. Sen, *Inequality Reexamined,* Harvard: Harvard University Press, 1998.

M. Toboso, "Rethinking disability in Amartya Sen's approach: ICT and equality of opportunity", *Ethics and Information Technology*, 13(2), 2010, pp. 107-118.

Bibliography on Environmental Issues

Arendt, H. (1958). *The Human Condition: Second Edition*. University of Chicago Press.

Beaumont, P. (2019, April 24). Rotten eggs: e-waste from Europe poisons Ghana's food chain. *The Guardian*. Retrieved from https://www.theguardian.com/global-development/2019/apr/24/rotten-chicken-eggs-e-waste-from-europe-poisons-ghana-food-chain-agbogbloshie-accra

Hern, A. (2018, January 17). Bitcoin's energy usage is huge – we can't afford to ignore it. *The Guardian*. Retrieved from https://www.theguardian.com/technology/2018/jan/17/bitcoin-electricity-usage-huge-climate-cryptocurrency

IPCC. (2014). *Climate Change 2014: Synthesis Report. Contribution of Working Groups I, II and III to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* (p. 151). Retrieved from IPCC website: https://www.ipcc.ch/

Lee, T. B. (2018, May 17). New study quantifies bitcoin's ludicrous energy consumption. Retrieved 23 April 2019, from Ars Technica website: https://arstechnica.com/tech-policy/2018/05/new-study-quantifies-bitcoins-ludicrous-energy-consumption/

Müller, V. (Forthcoming). *Ethics of Artificial Intelligence and Robotics*. More info needed

Petrlik, J., Adu-Kumi, B., Hogarh, J., Akortia, E., Kuepouo, G., Behnisch, P., … DiGangi, J. (2019). *Persistent Organic Pollutants (POPs) in Eggs: Report for Africa*. Retrieved from IPEN website: https://ipen.org/documents/pops-eggs-report-africa

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

van Wynsberghe, A., & Donhauser, J. (2018). The Dawning of the Ethics of Environmental Robots. *Science and Engineering Ethics*, *24*(6), 1777–1800. https://doi.org/10.1007/s11948-017-9990-3

World Wildlife Fund. (2018). Climate change could imperil half of plant and animal species in the world's most naturally rich areas. Retrieved 19 March 2018, from World Wildlife Fund website: https://www.worldwildlife.org/stories/climate-change-could-imperil-half-of-plant-and-animal-species-in-the-world-s-most-naturally-rich-areas

References RRI

Funtowicz, S. and Ravetz, J. (1993). "Science for the post-normal age", *Futures*, 31(7): 735-755.

Funtowicz, S. O., and Ravetz, J. R. (1990). "Post-normal science: A new science for new times", *Scientific European*, October 1990, 20-22.

Reber, B. (2018). "RRI as the inheritor of deliberative democracy and the precautionary principle", *Journal of Responsible Innovation*, 5:1, 38-64.

Stilgoe, J., Owen, R., Macnaghten, P. (2013). "Developing a framework for responsible innovation", *Research Policy* 42(9), 1568-1580.

Sunstein, C.R. (2005). *Why Societies Need Dissent*, Harvard University Press.

van den Hoven, M. J., Lokhorst, G. J. C. and van de Poel, I. (2012) "Engineering and the problem of moral overload", *Science and Engineering Ethics*, 18, 1–13

von Schomberg, R. (2013). "A Vision of Responsible Research and Innovation?, in R. Owen, J. Bessant and M. Heintz (eds.), *Responsible Innovation: Managing the Responsible Emergence of Science and Innovation in Society*, Chichester, UK: Wiley, 51-74.

Winner, L. (1980). "Do Artifacts Have Politics?" *Daedalus*, Vol. 109, No. 1: 121-136. Reprinted in Donald A. MacKenzie and Judy Wajcman (eds.), *The Social Shaping of Technology*, London: Open University Press, 1985; second edition 1999.

References AI and discrimination

Angwin, J., Larson J., Mattu, S. y Kirchner, L.,"Machine Bias", ProPublica, 23 de mayo de 2016, https://www.propublica.org/article/machine-bias-risk-assessments-in-criminal-sentencing

Brynjolfsson, E. y McAfee, A., *The Second Machine Age*, Nueva York; Norton,W.W, 2016.

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

Bughin J. and van Zeebroeck, N., "The promise and pitfalls of Ai", *Project Syndicate*, September 6, 2018, https://www.project-syndicate.org/commentary/artificial-intelligence-digital-divide-widens-inequality-by-jacques-bughin-and-nicolas-van-zoeebroeck-2018-09.

Cave, J., *The ethics of data and of data science: an economist's perspective*, Published:28 December 2016, https://doi.org/10.1098/rsta.2016.0117

Chander, A. "The Racist Algorithm?," *Mich Law Review* 115, no. 6 (2017): 1023, http://michiganlawreview.org/wp-content/uploads/2017/04/115MichLRev1023_Chander.pdf.

Dastin, J. "Amazon scraps secret AI recruiting tool that showed bias against women", *Reuters,* October 10, 2018, reuters.com/article/amazoncom-jobs-automation/rpt-insight-amazon-scraps-secret-ai-recruitingtool-th-showed-bias-against-women-idUSL2N1WP1RO .

McKinsey Global Institute, *Jobs lost, Jobs gained: Workforce transitions in a time of automation*, December 2017;

McKinsey, "Notes from the Frontier. Applying AI for Social Good", 2018.

McKinsey Global Institute*, Skill Shift, Automation and the future of the workforce*, McKinsey Global Institute, May 2018.

European Commission's High-Level Expert Group on Artificial Intelligence (AI HLEG),*Draft Ethics guidelines for trustworthy AI*, 18th December 2018.

Tenner, E., "The Efficiency Paradox: What Big Data Can't Do", Nueva York: Alfred A. Knopf, 2018; Centro Berkman Klein para Internet y Sociedad, "AI and inclusion", consultado el 22 de junio de 2018, https://aiandinclusion.org/

The Toronto Declaration: Protecting the rights to equality and non-discrimination in machine learning systems, May 16, 2018, Preamble, accesible at https://www.accessnow.org/cms/assets/uploads/2018/05/Toronto-Declaration-D0V2.pdf

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073

INBOTS

Inclusive Robotics for a better Society

www.inbots.eu

@INBOTS_CSA

@INBOTS_CSA

@INBOTS_CSA

This project has received funding from the European Union's Horizon 2020 research and innovation program under grant agreement No 780073