



# Humans and Societies in the Age of Artificial Intelligence

*A report by Vladimír Šucha  
and Jean-Philippe Gammel*

### **Getting in touch with the EU**

Europe Direct is a service that answers your questions about the European Union.

You can contact this service:

- by freephone: 00 800 6 7 8 9 10 11  
(certain operators may charge for these calls),
- at the following standard number: +32 22999696 or
- by email via: [https://europa.eu/european-union/contact\\_en](https://europa.eu/european-union/contact_en)

Luxembourg: Publications Office of the European Union, 2021

© European Union, 2021

The views expressed in this report are the sole responsibility of the authors and do not necessarily reflect the views of the European Commission.

Reuse is authorised provided the source is acknowledged and the original meaning or message of the document are not distorted. The European Commission shall not be liable for any consequence stemming from the reuse. Thereuse policy of the European Commission documents is implemented by Commission Decision 2011/833/EU of 12 December 2011 on the reuse of Commission documents (OJ L 330, 14.12.2011, p. 39).

For any use or reproduction of elements that are not owned by the European Union, permission may need to be sought directly from the respective rightholders.

The European Union does not own the copyright in relation to the following elements:

- cover: picture by Henrik5000 (172438113), source: iStock.com,
- page 4: picture by MattL (1200536969), source: iStock.com,
- page 6: picture by Monsitj (1047349788), source: iStock.com,
- page 11: picture by Metamorworks (1169711473), source: iStock.com,
- page 43: picture by Blue Planet Studio (1172642281), source: iStock.com,
- page 51: picture by Whitehouse (1053434168), source: iStock.com,

# **Humans and Societies in the Age of Artificial Intelligence**

A report by Vladimír Šucha and Jean-Philippe Gammel





## Table of contents

<b>Introduction .....</b>	<b>7</b>
<b>Main takeaways .....</b>	<b>11</b>
<b>n°1:</b> In the age of AI, 'Intelligence' will have many different meanings .....	12
<b>n°2:</b> We must not underestimate the cultural dimension of AI .....	15
<b>n°3:</b> Machines already manage our emotions much better than we do .....	17
<b>n°4:</b> Good intentions will not prevent AI from having extremely disruptive effects ..	19
<b>n°5:</b> While we talk a lot about privacy, we continue to give our data away .....	22
<b>n°6:</b> Freedom of choice is being eroded and liberal democracy and free markets are at risk .....	25
<b>n°7:</b> We need to understand if and how the current mental health crisis is linked to the digital transformation .....	28
<b>n°8:</b> Creativity will no longer be a human prerogative .....	31
<b>n°9:</b> AI risks creating the most unequal societies that ever existed .....	35
<b>n°10:</b> Education is our best bet .....	38
<b>Actions.....</b>	<b>43</b>
Strengthening the knowledge base .....	44
Gearing-up policy-making .....	45
Assessing the impacts of AI systems .....	46
Empowering citizens .....	47
Promoting AI based solutions for public good .....	48
Exploring futures .....	49
 Acknowledgements.....	 52
References .....	54
Annex: Questionnaire.....	60





## Introduction

Artificial Intelligence (AI) will radically change our lives and transform our societies. This shift, which has already started, will most probably be the deepest and the fastest humanity has ever experienced. There is general agreement that AI will bring many new opportunities and that most of the changes should be positive. There is also broad recognition that the development of AI entails many risks that need to be handled very carefully. What is however often underestimated is that both the positive and the negative effects of AI will be extremely disruptive. This is why it is very important to go beyond the traditional dialectic of risks and opportunities and to try to better understand what the real impacts of AI on individuals and societies might be.

AI is already able to carry out many tasks better than human intelligence. In the coming years, the number of these tasks will increase. Some day, machines might even reach the level of what many experts call 'Artificial General Intelligence' (AGI); i.e. the ability to achieve most cognitive tasks at least as well as humans. But even if we never reach AGI, the incredibly fast development of AI applications starts to raise fundamental questions for humankind, including what it means to be human in the age of AI. It is therefore important that we engage in a broad conversation about the future of humanity. In his book *Life 3.0*, Max Tegmark calls this the most important conversation of our time [1]. And as this is a conversation about our future, it should not be limited to AI researchers and scientists, but it should also engage policymakers and more generally citizens.

This report looks into some of the potential impacts of AI on creativity, culture, education and more broadly on individuals and societies. It is mainly aimed at policymakers to stimulate a discussion around a number of emerging issues.

### Why this report...

AI has without doubt become the hot topic for discussion. Not a week passes without at least one conference taking place or the publication of several reports aimed at policymakers, which clearly demonstrates the high level of interest in this subject. However, most of the ongoing discussions on AI are focusing on the short and medium-term effects.

What sets us apart from most other reports is that our aim is to have a longer-term approach on this topic. We are convinced that it is important to go beyond the state of the art in technology and beyond the developments that are expected in the next two to three years. The labs developing AI systems are moving extremely fast and some of the elements of what we still consider as being for the longer-term may already be with us in the next 5 to 10 years. If we wake up too late, there will be little room left for policymakers to influence the future with AI.

There is currently a race for AGI going on around the world with a general belief that the winner would take it all as such a breakthrough would further speed up the research effort. The motivation is therefore immense and there is a genuine risk that safety measures might be bypassed. This is strengthened by the fact that the development of AGI has become one of the most hyped scientific challenges and that it attracts some of the brightest minds coming from various disciplines. They usually have an excellent knowledge in mathematics, physics or engineering, but may lack a deep understanding of ethics and the insights provided by social sciences.



We are fully aware that by looking to the longer-term, we enter rather uncomfortable territory for many scientists and policymakers because of the many uncertainties that lie ahead. While most AI researchers agree that there are few physical or technological limits to how far the development of AI will be able to go, there is little consensus about how fast this may happen. However, many stress that in the digital area, a lot of breakthrough innovations came about in a rather unexpected way and often faster than anticipated, that the impacts will be massive, and that it is therefore important to start preparing now. Looking ahead to the longer-term effects is also uncomfortable territory because many of the potential developments of AI are rather counter-intuitive and difficult for people to apprehend from the standpoint of a human being in 2021. The questions that they raise are often rather disconcerting or even disturbing and the easiest way out is to simply discard them as 'science fiction'.

However, whether we like it or not, some of the technologies that are at our door already bring us very close to science-fiction scenarios. These scenarios will of course be very different to those popularised in movies, such as 2001: A Space Odyssey, The Matrix or Terminator and Yuval Noah Harari is right when he writes that 'the future is not what you see in the movies [2].' We must remain very cautious if we want to keep control and make sure that AI will be complementary to human intelligence and that it will fulfil its positive potential and improve our lives, and ultimately benefit humanity. Many experts rightly claim that AI is a product of human intelligence and that it will therefore become what humans want it to be, but at the same time they point out that the more we improve machines' capacity to learn and to be creative, the more it will become difficult for us to understand some of the outcomes. A dystopian future is not impossible. People like Bill Gates and Elon Musk have voiced their concerns, and the late Stephen Hawking already warned us back in December 2014 that 'the development of full AI could spell the end of the human race [3].' Others, like Harari or Le Cun, suggest that Homo sapiens is only one stage of the evolution and that we will maybe soon have to pass on the torch to entirely new types of entities [4, 5]. This should incite us to take the development of AI extremely seriously and to think about what kind of future we want for humankind.

## Methodology

This report is mainly informed by qualitative research and a questionnaire we had prepared (see Annex). We sent this questionnaire to over 70 experts from the digital and AI field, but also from other scientific disciplines, as well as from the cultural and creative sectors.

To frame the topic and to reflect the scope of our reflection, the questionnaire covered 50 questions on a broad range of areas. We did not intend to carry out any quantitative analysis and we therefore asked the respondents to select only the questions that seemed the most relevant to them and to share with us their thoughts, knowledge and opinion in a very free and open form. We also asked them to send us the references to all the books, articles, research projects or sources of data that they considered useful for our research.

Our longer-term approach allowed us to go beyond the current state of the art in technology, so the questionnaire also explored developments that may still sound unrealistic at this stage and may possibly remain unreachable. However, some weak signals are emerging and therefore, we consider that they should already be on the radar of policymakers and part of their reflections.



We collected a total of 34 answers, as well as a long list of references to articles and books. In a number of cases, the written answers were followed-up by videoconferences to clarify some elements or go deeper into certain issues. All the contributors are listed in the annex.

For this report, AI is considered in a broad sense in line with the definition used in the JRC Flagship report on AI: 'AI is a generic term that refers to any machine or algorithm that is capable of observing its environment, learning and based on the knowledge and experience gained, taking intelligent action or proposing decisions [6].' Many different technologies fall under this definition, but it implies some degree of autonomy.

It is important to note that while the focus of this report is on AI, we consider digital tools more generally on a number of issues, including social media, because often they already use some kind of AI and its further technical advancement will accelerate and amplify many of the current trends. Furthermore, we address very broad social transformations for which AI is only one potential factor among many others, including also other scientific developments, such as genetic engineering or nanotechnology.

Our reflection builds on the important work already initiated by the European Commission over the last few years with for example the White Paper on AI published in February 2020 [7], the Ethics Guidelines for Trustworthy AI prepared by the High-Level Expert Group [8], or the recent Digital Services Act [9] and Digital Market Act [10]. These documents extensively cover the challenges that will need to be addressed by the Commission as a matter of priority in the months and in the years to come [11]. Our angle is slightly different and more exploratory as we try to look further ahead.

## What is in the report

Making sense of the huge amount of material we collected has been a very difficult task because of the complexity of the topic, but also because we often felt overwhelmed by the quantity of information that emerged. We sometimes had the impression that every time we asked a question, a whole new world opened-up and we could have gone on forever in our exploration. The order of magnitude is well reflected in the number of scientific publications. In 2019 alone, there were about 10 publications with a topic related to digital and the digital transformation published every hour. Of these, one publication focusing on AI was circulated every 30 minutes [12]. We felt however that it was important to present our initial findings in this report, knowing that we only scratched the surface and that all the aspects we covered will deserve to be looked into in more depth.

Quite paradoxically, despite all the material we collected, it appears that knowledge on the potential impacts of AI is still quite fragmented. Some areas have so far only been very partially covered and there are still a number of gaps for which further research will be needed. Furthermore, the existing literature often covers quite narrow research areas and, in many cases, looks at the problems from the perspective of a single discipline. Also, in some areas experts reach diverging conclusions. This makes it quite difficult to connect the dots and to get a sense of the bigger picture. Meta-studies and transdisciplinary cooperation will therefore be vital in the years to come. We will have to increase multidisciplinary knowledge, we will have to make sense of it, but we will also have to assess and to re-assess constantly all the potential impacts. Science will need to

adapt to this new challenge, but policy-making systems too will need to gear-up to be fit for purpose for the age of AI.

As mentioned above, this report is only a first analysis. We decided to keep it as short and as readable as possible. We will list the 10 most important takeaways from our reflection. For each of these takeaways, we will provide a series of short comments and suggest a few references for further reading for those who are interested in digging more deeply into a given issue. At the end of the report, we propose a list of possible actions in order to move forward. These actions include suggestions for (i) strengthening the knowledge base, (ii) gearing-up policymaking, (iii) assessing the impacts of AI systems, (iv) empowering citizens, (v) promoting AI based solutions for public good, and (vi) exploring futures.

It is important to stress that we intentionally avoid repeating what is already common knowledge. We prefer to focus on elements that are either new, less known or in some cases even counter-intuitive. This is also why we have chosen to dedicate longer sections to areas such as creativity and education that have so far been much less explored from an AI perspective despite their importance in this context.

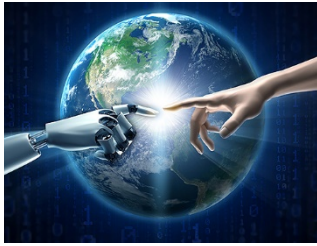
We would also like to clarify that while we refer to a few governmental initiatives, the ambition of this report is in no way to list all the activities that are starting to emerge at national or international level. We recognise that governments will have a key role to play in addressing many of the challenges we have identified. However, an analysis of all the relevant initiatives is outside the scope of this report.

# Part 1



# Main takeaways





### Takeaway n°1:

### **In the age of AI, 'Intelligence' will have many different meanings**

- Despite the progress made thanks to the use of Magnetic Resonance Imaging (MRI) or the recording of brain electrical and magnetic fields (EEG/MEG), we still understand very little about how the human brain functions and what the relation is between brain and mind. Most experts agree that the big discoveries are yet to come. This is why projects like the Human Brain Project [13] or EBRAINS [14], which boost scientific cooperation, support the development of digital research infrastructure and facilitate the sharing of data, are so important.
- At this stage, there is no common understanding of fundamental concepts such as intelligence or consciousness. This makes discussions about human intelligence very complex and comparisons between human and AI even more difficult [15]. There is however a strong interest in such comparisons, and one of the main methods that is used today despite all its limitations and its anthropocentric bias is to decompose human intelligence into its functions and then to compare these specific functions with the abilities of AI.
- AI is very different to human intelligence and it will remain different even if the development of AI will help to better understand the human brain and if this in return will help to further develop technologies, such as deep learning and neural networks. The development of diverse forms of intelligence over time has always been the result of a need to tackle very specific adaptive challenges. Emotions and consciousness have played and still play a fundamental role in the development of human intelligence. Machines do not need emotions and consciousness to develop their own form of intelligence. Many experts do however believe that machines will be able to develop some form of sensations, emotions, consciousness or intuitions in the future and that this will open new perspectives for AI [5, 16, 17].
- Already today, AI is much better than human intelligence at recognising patterns in large sets of data and it is therefore able to outperform humans in a number of tasks, such as playing chess, driving cars or diagnosing diseases. As these tasks are based on a narrow set of goals, this is generally called 'narrow artificial intelligence [1].' What still gives humans an advantage is our flexibility to rapidly adapt to a wide array of different challenges and circumstances. This flexibility builds on a mechanism Yuval Noah Harari calls our narrating-self: our mind constantly tries to impose order in the chaos of our experiences and to make sense of our life by using partial clues in order to shape coherent and plausible stories [4]. These stories help us to develop our simplified representations of the world, which we can then use to build scenarios, simulate the effects of our actions, and develop action plans that enable us to decompose complex tasks into simpler tasks [5]. This model does not prevent humans from making many mistakes, but it is fast and requires only a limited amount of effort [18]. Machines still have a long way to go to reach such flexibility and become as energy-efficient, but that is precisely the aim of AGI.

- Because AI is highly complex and sometimes counter-intuitive, the general public often has a simplified perception of what it actually is. Many simply tend to reduce it to some type of faster and better performing computing. AI is more than that because it can observe its environment and learn with at least a certain degree of autonomy. Others tend to associate AI with humanoid robots. Robotics is however only one subfield of AI. In line with the definition mentioned in the introduction, AI systems can take many different forms and use many different technologies. Some AI systems have a very high degree of autonomy and can replace humans in certain tasks (e.g. self-driving cars). Other AI systems are hybrid and rely on a permanent interaction with humans.
- One form of hybrid systems that are already in use are the 'AI extenders.' They can be imbedded in tools as simple as smartphones and their purpose is to help humans to complete certain tasks. They do not simply augment our cognitive capacities, they can literally extend them outside our brains and into the tools so that the tools can become constitutive parts of our minds [19, 20]. The AI extenders offer fantastic new opportunities to boost cognitive capacities. For example, they can be used as cognitive prosthetics for people with conditions, such as Alzheimer, learning disorders or autism. However, they also pose extremely complex health, philosophical and ethical challenges, and their reliability and ethical status will need to be regulated and certified [see also Takeaway n°7].
- Until now, humans have been using their natural cognitive capacities given by genetic inheritance and formed by their natural environment. A wide spread of cognitive extenders beyond medical use, may put in question our conception of what humans are capable of and the notion of 'mental normality' [20]. This would also raise the question of the access to such tools. If cognitive extension or enhancement remains reserved to certain 'privileged' groups of people, that would have a huge potential impact on equality and fairness [See also Takeaway n°9].
- There could also be a risk that the use of AI extenders follows the worrying example of the misuse of nootropic drugs (chemical substances enhancing cognitive functions). For example, a study has confirmed that 11% of medical students use these drugs without medical diagnosis [21]. In a survey of the Nature journal, 20% of readers admitted that they had used at least once a chemical substance to enhance their performance [22].
- While AI extenders are probably the most advanced form of hybrid intelligence, it is not the only one and other forms are emerging that are sometimes even more startling. There has been an explosion of more than 11 000 scientific publications about brain – computer interfaces over the past 10 years. Of those, 200 were dedicated to brain implants [12]. The most vocal in this field at this stage is Elon Musk with his start-up Neuralink [23], but at least two more start-ups are working on this topic, Paradromics [24] and Synchron [25], and several big digital companies report activities. All developments and experiments are done currently in the medical domain with the aim to restore lost functionalities of the body and of the brain. If successful, they will certainly move to non-medical domains too. The military is one that springs to mind, but many others may follow. It will be very difficult to prevent the commercial spread of brain implants, since the opportunities linked to

unprecedented human enhancement will be tempting. The challenges related to this are beyond the scope of this report, but should be followed very closely.

- Another specific form of intelligence, which is gaining a momentum with the rise of AI, is collective intelligence. Collective intelligence is generally described as the ability of groups of people to think and work together and tap into collective knowledge and resources when identifying problems, developing solutions, making decisions and learning. If it is well understood and correctly used, it can bring significant improvement to decision-making in both public and private organisations. AI tools can support and reinforce collective intelligence processes in many different ways [26]. For example, they can act as facilitators of debates by making sense of huge amounts of data and knowledge, clustering the relevant elements and therefore help to navigate through the increasing complexity of many problems. They can also act as contributors by providing additional knowledge to a group therefore creating another form of hybridity where the brain-computer interaction does not take place at individual level but at collective level with the emergence of mixed human-machine groups.
- With the emergence of various forms of AI systems, there is an urgent need to better understand the way the intelligence landscape is changing. However, we lack the tools to properly evaluate, compare and classify AI systems, which makes it difficult to determine the future of the field. We also lack the tools to compare AI systems with human intelligence and other forms of biological intelligences. A mapping of different cognitive systems, including humans, non-human animals, AI systems, hybrids and collectives, providing data, analysis and visualisations would be extremely useful. A group of scientists from 12 research institutes has proposed a possible way forward to develop an Atlas of intelligence that could serve as a source of inspiration [27].





## Takeaway n°2:

### **We must not underestimate the cultural dimension of AI**

- The cultural context in which we live strongly influences our values, our identities and our worldviews, both as individuals and as groups. It therefore also affects our perceptions, our behaviours and our decisions. Values and identities are not static: they evolve both over time and through contacts with other individuals and groups.
- Therefore, the cultural context has also a very strong influence on how AI and its applications are designed and deployed in different parts of the world. The understanding of concepts such as privacy, common good or well-being is for example very different in China, in the United States and in Europe [28]. This directly affects the level of acceptance policymakers, AI developers and citizens have concerning the sharing and the transparency of personal data, the intrusiveness of certain technologies, surveillance or the risk of inequalities.
- An analysis of scientific literature has demonstrated that for many years research about the ethical impacts of AI has been lagging behind the development of technological knowledge. This situation has only started to change in 2017. We can now observe a sharp increase in the number of scientific publications on ethics and AI and 90% of the existing knowledge has been published after this date. It is very interesting to note that the main players are the EU, the UK and the US, and that China despite being a superpower in AI technology seems to remain largely uninterested in this field [12].
- The cultural differences will make it very difficult to reach a common understanding at global level on ethical frameworks and to put them into practice effectively. Over the past few years, over 80 sets of ethical principles and guidelines for AI have been developed by civil society, private sector, government, inter-governmental and multi-stakeholder organisations. A number of common themes have emerged including privacy, accountability, safety and security, transparency and explainability, fairness and non-discrimination, human control of technology, professional responsibility, and promotion of human values [29]. However, ethics cannot simply be deduced from codes, interpretation plays a big role and this interpretation will vary greatly according to the cultural context. Some of the codes that were developed mainly based on Western enlightenment principles are already challenged by experts from other parts of the world who claim that they are at odds with Confucian, Hindu or Islamic values [30]. It will therefore be extremely interesting to follow closely UNESCO's efforts to 'globalise' ethical principles for AI by working with 155 countries from across the world.

- If the cultural context has a strong influence on the way in which AI is developed in different parts of the world, AI in return will have a very powerful impact on the way in which our cultures, our values and our identities will evolve. The main risks are either a risk of uniformisation or a risk of cultural hegemony of certain cultures over the others. In both cases, we will have to be very careful to preserve cultural diversity. From the EU cultural perspective, we should be alerted by the fact that most AI related patents are not of European origin. Almost 80% of all filed patents are from the US, Japan and China and of the top 20 companies, 12 are based in Japan, 3 in the US and 2 in China. From a global perspective, we need to point out that the Global South is not present at all in the patenting of AI related technologies [12].
- In this respect, there is an urgent need to better understand the real impacts of the 'recommender algorithms' that guide our choices whenever we want to listen to music, watch a film or a series, or buy a book. While scientists have started to look into this issue, there is at this stage no clear consensus if recommender algorithms have predominantly a positive or a negative effect on cultural diversity [31, 32]. On the one hand, some experts claim that these algorithms keep us in bubbles by suggesting similar choices to what we already know and already like, and that they therefore prevent us from having new and original experiences. This induces a risk of uniformisation or polarisation. They also highlight that there is an imbalance because most of the algorithms are developed and maintained by commercial companies that have mainly economic interests. On the other hand, other experts argue that there is no scientific evidence of the negative effects of algorithms on diversity and they highlight the role played by the diversification and serendipity factors embedded in the algorithms.
- It is very interesting to note that so far algorithms predominantly cover the most widespread cultures and the most spoken languages, but that there is still a long way to go before they will be able to interpret the meaning of signs and languages in widely differing cultures. For example, ongoing projects in Australia or Canada suggest that it may be some time before AI can interpret indigenous systems and the languages based on them. It may preserve these languages and these cultures from uniformisation but at the same time, people speaking these languages will not benefit from the services offered for other languages. This calls for a careful analysis and might require public interventions.
- It is important to stress that cultural diversity is not only based on languages and geography, but also on a wide range of other traits, beliefs and preferences that need to be protected and preserved. Diversity does not only concern groups but also individuals. Atypical, non-conformist or transgressive individuals triggered many of the discoveries and of the 'avant-garde' movements that have allowed humanity to move forward. We need to nurture difference and also to respect what happens in the margins of our societies. In addition, and rather counter-intuitively, many of our steps forward built on chance or initial mistakes or imperfections. One of the challenges in developing AI will therefore be to leave enough room for this type of unpredictability.



### Takeaway n°3:

### **Machines already manage our emotions much better than we do**

- A revolution in the science of emotion has emerged over the last decades. The activity in this field of research is exploding (about 24 000 research papers published over the past 20 years with a sharp increase in recent years). This is creating a paradigm shift in the way in which we understand decision-making processes [33]. This revolution builds on the work of pioneers in the field of behavioural sciences, such as Nobel Prize winners Daniel Kahneman or Richard. D. Thaler [18, 34, 35], but it goes one step further by demonstrating the key role played by emotions as drivers of decision-making and by exposing how powerful they are as predictors of human behaviours [33]. Emotion and reason are not necessarily antagonistic. They are strongly interconnected and complement each other. Our best decisions are actually those that combine reason and emotion [36]. Initial unconscious and emotion-laden processing of information shapes all subsequent phases of thinking. When reasoning then comes in, it is often simply as a verbal justification of already made judgements and to help us build simplified representations of the world [see Takeaway n°1]. Our emotions and intuition can have many undesirable effects, but they are also absolutely necessary because they are powerful shortcuts that allow us to take rapid, automatic and effortless decisions in real time on matters that would exceed the capacity of our reasoning because of their complexity [18, 37].
- Humans are generally very bad at recognising their own emotions, and even worse at understanding the emotions of others. AI is already several steps ahead. Algorithms can read us like open books and detect our emotions and sentiments much better than we do. They can analyse the words that we use, the tone of our voice, our facial expressions, our movement patterns, the micro-spasms of our muscles, etc. On this basis, they can decipher our present emotional state, but also very accurately predict our future behaviours, and even deduce our personality type.
- Because AI is so good at reading and predicting our emotions, it is also extremely powerful at influencing them. It excels at catching and holding our attention, encouraging certain behaviours or even manipulating our decisions [38]. For example, using our emotions to hold our attention is at the basis of the business model of most internet platforms (the 'attention economy') [see also Takeaway n°5]. Jason Davies describes extremely well the model used by TikTok. Users do not even need to specify their preferences when they join the platform. AI algorithms immediately get to work analysing their behaviour and delivering content, as opposed to simply making recommendations. In very little time, they learn enough to make stunningly accurate predictions about which videos will catch a user's interest. In a way, TikTok soon knows users better than they know themselves as behavioural preferences may differ from stated desires. This drives user engagement – exactly the kind of audience advertisers and vendors are after [39]. It is therefore not a surprise that there has been a sharp increase in new publications and new



patenting in the field of emotions/feelings and AI over the past four to five years. The number of filed patents exceeded 200 and one of the most active organisations worldwide is Tencent Technology from China [12].

- Emotions can also be used in a very powerful way for political motivations. For example, anger, anxiety or disgust play an important role in shaping our judgements and attitudes on many political issues [36]. From an evolutionary point of view, they play the role of alarm systems outside conscious awareness to protect us from certain dangers, they are extremely difficult to override. This is something that the disseminators of fake news understood quite early on and use extremely efficiently. Many politicians, mainly from populist parties, have also become experts in using these emotions to the point that the world 'emocracy' has emerged to describe some of our democracies [see also Takeaways n° 5 and 6].
- At this stage, AI systems cannot yet develop or express their own emotions, but they can already mimic emotions and sentiments, such as empathy. These technologies are referred to as 'emotion AI', 'affective computing' or 'artificial emotional intelligence' [40, 41]. Sensations and emotions have a clear evolutionary rationale for human beings, and in principle, AI systems will not need to rely upon emotions to tackle their own 'analogous adaptive challenges.' However, humans may decide to design AI systems that develop emotional reactions for various reasons, such as to facilitate the human - machine interaction, to hybridise AI and human based adaptive strategies, or more prosaically to strengthen the capacity of machines to manipulate people [42]. Technically, it is very likely that these developments will be possible, but they would open many very complex philosophical and ethical questions [5].
- The battle to control, influence or manipulate emotions will without any doubt be one of the key battles of the 21st century. On the one hand, internet platforms will have a strong interest in continuing to develop AI systems that get better and better at reading our emotions and at influencing our behaviours. On the other hand, humans will need to become much better at recognising and managing emotions if they want to protect their free will.
- Emotional literacy and emotional intelligence need to be very strongly reinforced both at individual and collective level. At individual level this goes through strengthening the teaching of socio-emotional skills in all our education systems [see also Takeaway n°10]. At a collective level, it will be extremely important for policymakers to become much more aware of the central role played by emotions in our decision-making and in our behaviours, as well as to learn to acknowledge, integrate and use them. For example, AI systems could help to better sense citizens' concerns and to take the emotional temperature in different geographic locations. This could then contribute to developing policy options that speak to people's emotional needs and values, therefore leading to improved impacts.



#### Takeaway n°4:

#### **Good intentions will not prevent AI from having extremely disruptive effects**

- As AI will further develop, it will bring huge new opportunities to improve our lives. It can help us to tackle crucial challenges such as climate change or healthcare. It also has considerable potential to increase our well-being more generally. Many experts have developed a very positive vision of what our future could look like. They highlight, for example, that if AI can be used to maximise efficiency and create material abundance, it may allow people to focus less on their material needs and more on human interactions, self-development, reaching their full potential and living happier lives.
- However, our future with AI will depend a lot on our capacity to fully grasp all the implications of the AI-enabled technologies we are developing and of their applications, which is extremely complex. It is important to be aware that AI technologies can have not only intended, but also unintended effects. The intended effects can be based on good intentions such as improving healthcare or improving security or the quality of services. They may also have malicious purposes, such as disinformation or online manipulation [43]. Technologies developed with good intentions can often very easily also be misused for malicious purposes. To complicate things further, technologies developed with good intentions can also have unintended consequences with negative impacts, such as behavioural addiction, strengthening biases, polarisation and radicalisation, or jobs obsolescence and creating technological unemployment [the issue of the future of work will be discussed much more in details under Takeaway n°9]. All these dimensions need to be taken into account when assessing the impacts of AI technologies and applications [29]. When we look back at the history of human inventions, we have to admit that we have always been much better at inventing things than at understanding their real impact and collateral damages.
- All too often, we also underestimate that even the most positive developments enabled by AI will have very disruptive side effects and can radically change our lives as well as what it means to be a human being. Let us just take one example with healthcare. There is general agreement that AI will enable a major leap forward in the coming years and decades. The processing of large quantities of data, including individual genome data and information about people's surrounding environment and variables will not only enable personalised treatments, but also very early preventive interventions. This will eventually raise the question of defining the limit between what is healing and what is upgrading humans with all the ethical and fairness issues this entails. It will also, progressively and significantly, increase the life expectancy of most humans. Life expectancy almost doubled in the 20th century, mainly due to the discovery of antibiotics and the reduction of child mortality. If thanks to technological progress we follow a similar trend in the 21st century, which is considered as a plausible scenario by many scientists, this would revolutionise

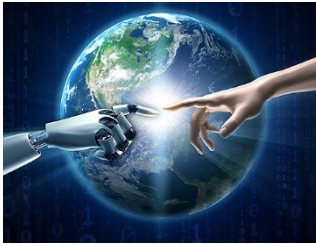
human society. Family structures, marriages, child-parent relationships, but also professional life would be radically transformed with much longer careers [4]. This would raise many extremely complex new challenges, such as the need to find an appropriate way to make place for new generations with their novel ideas and aspirations. The risk of overpopulation of our planet should also not be underestimated. If we let our imaginations go one step further, even immortality, both physical and digital, should no longer be taboo. For certain scientists, death has become nothing more than a technical problem that could and should be solved [4, 44]. If physical immortality ever happens with all its tremendous consequences, this will most probably not be before many decades. However, digital immortality might already be much closer. 'Mind uploading' for example, which should enable the transfer of data from a brain to a computer, has already become more science than fiction, in particular with projects funded by Elon Musk or the Russian billionaire Dimitri Itskov. The aim is to install both memory and behaviour in a machine (or a robot) after death and virtual copies of individuals could continue interacting with their relatives or even managing a company or leading a country forever. The policy implications would be huge and the developments need therefore to be followed very closely. This is actually linked to the already existing and yet unresolved problem of personal digital footprints. Indeed, who should have the right to use and take decisions about the digital content created by individuals over their lifetime?

- Another very complex and much debated issue is to what extent benevolent paternalistic interventions, or in other words nudging for good, should be permitted? One aspect that needs thorough analysis is how far we should go in allowing an AI system to influence a user's behaviour for those individuals' own good. However, an even more controversial practice that is on the horizon is the ethical acceptability of allowing an AI system to nudge a user's behaviour for the good of society. This opens multiple questions, including who should determine what is good for society and how we can agree on any universally accepted 'social goods' taking into account the role cultural variations and tolerances have in this context [45]. A first important step would be that nudging by AI systems should always be based on explicit consent. This is however very complex to implement effectively and far from being enough to ensure that people keep their autonomy [29][see also Takeaway n°6].
- This also brings us to the delicate question of the definition of well-being more generally. As the central aim of AI should be to increase human well-being, we need a framework that is at the same time broadly accepted and concrete enough to be practical. The work carried out by the OECD in this area provides a very solid basis [46]. Some organisations such as the Institute of Electrical and Electronics Engineers (IEEE), which is a global technical professional organisation dedicated to advancing technology for the benefit of humanity, have tried to go one step further by developing metrics that would enable us to assess the impacts of AI on well-being [47]. This is a very interesting avenue to explore, but also quite complicated in its implementation. Cultural variations and differences in interpretations should not be underestimated either. One key issue is various trade-offs among different dimensions of well-being and ultimately also the trade-off between individual freedom and the common good. In this respect, the Chinese position is for example quite different to the European or the American position [28]. Another crucial question is the importance of free will [see also Takeaway n°6]. If well-being is



mainly perceived as minimising suffering, feeling pleasant sensations and being free from unpleasant ones, or avoiding that uncomfortable desires emerge, then free will has a limited importance. In this case, AI would be extremely effective in satisfying our basic needs because it already understands us much better than we understand ourselves. It can manipulate our feelings and decisions according to our needs. However, if we believe that well-being also means that we should be able to give a meaning to our lives, then we should fight to keep as much autonomy as possible. There is a great deal of scientific knowledge arguing that meaning and a sense of purpose are the cornerstones of a healthy mind and well-being.

- It is crucial that we gain a much better understanding of the potential impacts, including the negative effects and the unintended consequences, of all AI systems and applications on individuals and societies. This implies the development of appropriate ex-ante impact assessment as well as ex-post evaluation tools based on rigorous indicators, including where possible well-being indicators, and engaging with all stakeholders concerned. This would enable the professionals who develop and deploy AI systems to better understand the impacts these systems may have. It could contribute to closing the gap between the various ethics codes that are currently being developed and practice [29].
- It will also be necessary to adjust, if not completely change, the way we regulate. Today, most of our regulations are developed in a reactive mode, once the technologies have already been deployed. In the age of AI (and of other exponential technologies), reactive policies will in most cases come too late to have a real effect because of the speed of changes and because of the nature of the potential impacts. New methods will therefore need to be designed in order to follow much more closely the development of technologies and to start elaborating (or updating) the needed regulations in parallel so that these regulations can already be in place when the technologies are deployed. This will require a new type of partnership between policymakers, regulatory scientists and private companies. On the one hand, regulation science, which is today focussing mainly on standardisation tasks, will need to develop much stronger anticipatory and monitoring capacities. On the other hand, trust will need to be built with private companies so that they are ready to open their laboratories to regulatory scientists. This should be of a mutual benefit as good regulation is not only about safety and security, but also about facilitating the market deployment of innovative products. More broadly, public administrations will need to reinforce their technical and scientific skills in order to be able to cope with the new challenges that are emerging, but also to make a better use of the opportunities that AI will provide to enhance policy-making [see also takeaway n°6] and public administration [48, 49]. Transversal cooperation in public administration should also be significantly reinforced and go far beyond the departments responsible for research, the economy or consumer protection since AI is penetrating and influencing all aspects of life.



### Takeaway n°5:

**While we talk a lot about privacy, we continue to give our data away**

- Some experts use the expression 'digital Far West' to describe the context in which we currently live and where data acquisition and manipulation often takes brutal forms, where large parts of our privacy have already been eradicated, and where mass surveillance has become a reality. While facial recognition raises many fears, it is only one of the many forms mass surveillance can take. Every single trace we leave on the internet can be used to deduct not only what we have done, but also who we are and what we are likely to do next.
- In this context, while most people, at least in Europe, remain extremely cautious about giving their personal data to public authorities, they quite paradoxically, have much less hesitation to give it to private companies including the tech giants. This mistrust of public authorities has been illustrated again by all the controversies surrounding the tracing-apps that have been launched by many governments to help monitor the spread of the COVID-19 and around the collection of mobile phone data. While there was a broad recognition that the pooling of contact and location data could be useful for the management of the responses to the crisis, a majority of people preferred not to download the apps and not to participate in the data pooling schemes because of their concerns over the potential misuse of their personal data [50]. To avoid these situations and to mitigate mistrust, reflections about new types of social contracts have started to emerge. The aim of these social contracts would be to enable individuals to share their data in exchange of specific services while having the assurance that this data would be used only on a 'need to know' basis and in full compliance with fundamental rights and freedoms [51, 52].
- One of the sources of this mistrust in public authorities is possibly linked to history. In the 20th century, authoritarian regimes such as the Nazi regime or the communist regimes developed a huge spying capacity. However, despite their attempts to collect all the data centrally, they remained largely unable to process it. As Harari rightly points out, in the 21st century similar regimes would have taken full advantage of the development of AI and algorithms to gain absolute control over their population and to make resistance utterly impossible. Enormous amounts of information collected centrally are no longer a challenge, but can become a decisive advantage [2]. Therefore, one of the key challenges to democracy in the coming decades will be to avoid the emergence of digital dictatorships and one of the possible safeguards will be to keep datasets and their processing as decentralised as possible.

- The internet is however mainly controlled by private corporate algorithms designed to maximise profits by capturing our attention without public accountability [53]. We often say that we are searching 'Google', but truth is that Google is searching us. Two main business models are driving AI and platform development at this stage: advertising and direct sales. Therefore, human attention has become a commercial 'product' especially on social media [see Takeaway n°3]. Social media are designed so that people give their valuable time, attention and data without considering the costs. It is important to note that they pose a privacy risk for users and non-users alike. It extends far beyond what individuals explicitly share with social media sites, because of how much can be inferred from users' activity. It is also important to note that even if they are mainly designed for economic purposes, they are also heavily used for political (and many other) motivations [53].
- There are mainly three reasons why we have so little hesitation to give our precious data away to the tech giants "for free". The first is complexity: users are generally unaware of what data they produce, provide to others and how that data are collected and stored when they perform basic tasks on social media platforms or on the internet in general [53]. The second is that many believe and accept that this is actually the price to pay if they want to take part in modern society and use the various services offered by the tech giants. The third, which we should not neglect, is that many consider that there is no point in doing or experiencing anything if nobody knows about it and that sharing experiences and emotions with other people actually enhances them. People enjoy being part of the dataflow because that makes them part of something much bigger than themselves [4].
- One of the key problems stressed by many experts is the imbalances. Imbalances between the tech giants and smaller companies. Imbalances between private companies and public authorities. Even if there have been some interesting experiments of data sharing during the COVID-19 crisis, the dependency on tech giants for implementing large scale digital solutions in public interest, emphasised the overarching influence of the commercial sector in setting the agenda and in creating the infrastructure used by governments to collect and use data [50]. The way in which Google and Apple managed to impose their views on how governments should build their national tracing-apps in particular has stimulated a lot of debate [54]. Last but not least, imbalances between private companies and customers. The companies see all the data, while customers only see their own. Alexander Peterhaensel gives the example of Facebook, which collects as much data as possible on any internet user – no matter if the user has a Facebook account – and aggregates those data in 'shadow profiles.' Despite Mark Zuckerberg's claims before the US Congress that 'The information that we collect, you can choose to have us not collect. You can delete any of it', it is not possible to access one's 'shadow profile', and therefore also not to monitor the termination of data collection or the deletion of user-associated data [55].

- Several suggestions about how to reduce these imbalances emerged during our conversations with experts. For example, José Hernández-Orallo stressed the need for customers to associate and regain power. He suggested the creation of 'customer-defending entities' supported by public authorities and NGOs that would be given the power to collect and analyse customers' data to be able, if needed, to propose alternative models of behaviour and to counteract the dominant dynamics created by big companies [56]. Marek Havrda stressed that Europe has a tradition of successful public funding which might be a relevant option to consider for AI development. For example, to diversify the business models, incentives could be created for platforms to better align with values that are explicitly chosen by their users. Those values might be for instance time well spent, promoting healthier democratic conversation, reducing environmental footprints, or other goals that come to be recognised in the years ahead [57]. The relevance and feasibility of these schemes and of other similar schemes would deserve to be looked into more carefully.
- It is absolutely crucial to empower citizens. Calls for greater 'media literacy' or 'critical thinking' are extremely important, but unlikely to be enough because of the complexity and the huge number of decisions we have to take about our data every time we surf the internet. The solution might therefore be a technological one: AI has a huge potential to erode people's privacy, it also has the opportunity to safeguard it. Various groups and researchers work on privacy-protecting tools based on AI personal-aid ('personal privacy angels') which would ensure that only the information an individual wants to share is actually shared. Decentralised solutions such as 'decentralised data stores' [58] where access rights are controlled by the individuals are being developed, and prototypes are already being produced [57].





### Takeaway n°6:

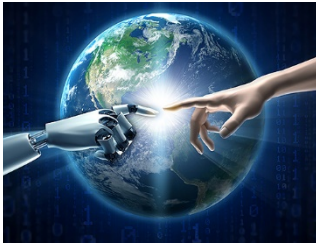
#### **Freedom of choice is being eroded and liberal democracy and free markets are at risk**

- According to Harari, at the close of the 20th century the great ideological battle between fascism, communism and liberalism resulted in the overwhelming victory of liberalism. Liberalism in this context must be understood in a broad sense that goes beyond specific political parties. It is based on the large consensus around the package of individualism, human rights, democracy and free market. The very core of liberalism is the power of individual humans to make free choices. In politics, liberalism believes that the voter knows best and in economics, it maintains that the customer is always right [2, 4].
- However, the capacity of individual humans to actually make free choices is now challenged. Free choices are based on free will and an increasing number of voices claim that free will is an illusion. Most scientists and thinkers do now accept the idea that human brains are some kind of biochemical mechanisms or even biochemical algorithms: our neurons react to electric signals, calculate some probabilities and, based on the results, emit or not further electrical signals [4, 5]. While there is agreement that our capacity to resist these 'natural algorithms' is limited, there is debate about how much room our consciousness and our creative intelligence actually gives us to keep some freedom in our decision-making [44].
- The second factor that challenges our capacity to make free choices is that, because of the way our brains function, AI already knows humans better than they know themselves and this capacity will increase in the years to come. On the one hand, this makes it very easy for AI to influence our decisions and our behaviours [see Takeaway n°3]. On the other hand, AI will become so good at making decisions for us that we will learn to trust it on more and more issues, running the risk that we gradually lose our ability to make decisions for ourselves [2].
- The conclusion Harari draws from this evolution is that democratic elections and free markets will make little sense in the future and that liberalism may collapse. What is indeed the point of having democratic elections when the algorithms know not only how each person is going to vote, but also the underlying neurological reasons to their choice [2, 4].
- The already existing pressure technology puts on democracy such as the creation of filter bubbles and echo chambers, disinformation, polarisation, radicalisation or manipulation by micro-targeting based on individual psychology have been well-documented [53]. Measures to mitigate these pressures have been taken, including at EU level with for example the recent proposals for a European Democracy action plan [59] or the Digital Services Act [9]. However, with the fast development of AI, these pressures will not only accelerate and amplify in the near future, they will also take new forms which are difficult to anticipate. Many experts we consulted therefore stressed that there is a high probability that traditional democratic politics will lose

control of the events, decline and even disappear. What is at stake is therefore less to preserve liberal democracy as we know it, than to be able to reinvent new forms of democracy fit for the age of AI. Concrete suggestions concerning possible ways forward remain however scarce at this stage, but it should not prevent us from starting this important discussion. One possible path worth exploring in this process could be to define or re-define the role of individuals in the technological era. The fear of irrelevance for individuals in the future may become a huge obstacle for democracy.

- Another key question that needs to be raised is how much we should actually accept to rely on AI if it is already and will increasingly become better than we are at making decisions for us. A recent survey on the attitude of Europeans towards technology has shown that one in four favours letting AI make important decisions about the running of their country [60]. Some tech-utopians have already developed the vision for a 'robogov' where AI would be fulfilling the role of political or administrative leadership and decision-making. Some of the claims made in favour of robogov are that it would be less hampered than human leaders by ideological extremism, tunnel vision, egotism and narcissistic tendencies. Therefore, robogov would also be less irrational, better able to capture real world complexity and more predictable [61, 62, 63]. However, this vision is not yet feasible because there remain many uncertainties regarding both the physical world as well as human behaviour. As Havrda also rightly points out, it is also not desirable because these uncertainties should not be dealt with by automated decisions, but need to be the subject to human decision-making, and in democratic societies subject to public debate [61].
- If AI should neither replace public debate nor human decision-making, it has however the capacity to play a major role in informing them. Thanks to AI, we will be able in the coming years to make a major leap forward in the way in which we identify pressing issues, develop scenarios, simulate and visualise the potential impacts of various policy options ex-ante, and evaluate their results ex-post. Deliberations and media reporting could become significantly more fact based. Ultimately, our decision-making and our policy-making could considerably improve. The Arizona State University is already experimenting with a platform to visualise the impacts of potential decisions with its 'Decision Theater.' This could be an interesting example to follow [64].
- Another challenge we will face if democratic elections lose part of their relevance is the need to develop new forms of citizen participation both online and offline. Many experiments are already taking place where citizens are involved either in deliberations, co-creation processes or policy experimentations. Online deliberation and consultation platforms can be very successful when they are carefully designed and skilfully facilitated [53, 65]. Co-creation of policies and measures is mainly based on design practices. Many governments have started to experiment with design thinking methodologies and the European Commission has its own dedicated team in the EU Policy Lab. Design is a fast-evolving discipline. New and more specific areas such as transition design, behaviour design or social design have emerged and can be used to create change and guide its implementation in public and private settings [66, 67]. Citizens can also be involved in policy experimentations. One of

the most advanced experiments carried out so far was probably Experimental Finland launched by the Finnish government and which conducted trials on basic income, digital municipalities or a service voucher system [68, 69].



### Takeaway n°7:

#### **We need to understand if and how the current mental health crisis is linked to the digital transformation**

- There is broad consensus among experts that the digital transformation, and in particular AI, will profoundly change our cognitive and socio-emotional capacities. Many compare this evolution to the revolution humans underwent when they moved from oral cultures to written cultures. Some brain functions that will be less needed will diminish or disappear. Others will augment, including with the help of devices such as brain implants or AI extenders [see Takeaway n°1]. This evolution will be facilitated by our brain's great plasticity and in particular our children's brains. As Pier Luigi Sacco very nicely summarises, from the cognitive viewpoint, we will certainly develop different strategies and heuristics than in a non-digital world. We are at the very beginning of the curve and it is difficult to make predictions. As in all situations, we will probably gain something and lose something at the same time. The important aspect will be to tune the trade-offs optimally to the new adaptive challenges we will have to face. That means building new skills that are relevant in the new environment and drop others that are no longer as needed as in the past [42] [see Takeaway n°10].
- Some first effects of the digital transformation can already be observed. There is wide consensus among experts that we can suffer from attention loss and the inability to concentrate and dive deep into problems because of the multiple disturbances generated by the digital environment. We also suffer from a greater impatience linked to our need of instant reactions and gratification, from reduced memory due to the instant availability of knowledge and recordings online, from a reduced capacity to orient ourselves in space, or from the incapacity to take decisions because we rely more and more on algorithms to decide for us.
- The chaos that is stemming from the deluge of data and information, fake news and disinformation, as well as the manipulation of our emotions creates stress and anxiety. Research has also confirmed a causal link between our growing inability to concentrate and unhappiness [70]. The high level of stress and anxiety goes hand in hand with a rapid increase of loneliness. A recent study by the Commission's Joint Research Centre shows that around 30 million adults in the EU feel lonely and that this is becoming an important feature of our hyper-connected societies [71].
- These various issues increase the risk that more and more people will try to escape real life and seek shelter in a virtual reality without stress and suffering. Becoming an avatar with all the desired mental and physical abilities and playing dream roles in different games and simulations may become increasingly appealing for people struggling to cope with real life situations. The market for virtual reality games or role-playing in online games is growing rapidly and games like Second Life, Active Worlds, Blue Mars or VR Chat are becoming easily accessible to a wide public.



- In this context, we need to pay particular attention to the silent pandemic of mental morbidity among young people around the world [72]. Stress and mental health issues are developing extremely fast and the World Health Organization (WHO) predicts that mental health issues will be the most prevalent diagnosis worldwide by 2030. For example, according to the US Center for Disease Control, there is a significant increase of the depression rate among teenagers aged between 12 and 17. Girls are most heavily hit with an increase from 13% in 2010 to 18% in 2018. The Attention Deficit Hyperactivity Disorder (ADHD) diagnosis increased from 6% of the population in 1998 to 10% in 2016 [73, 74]. In India, 25% of children aged between 13 and 15 suffer from some form of depression or anxiety [75]. In New Zealand, surveys carried out between 2000 and 2019 show a persistent increase of self-reported rates of depression and suicidality. The rates doubled between 2012 and 2019 [72]. Recent surveys in Denmark also showed very similar data and trends [76]. According to a recent report from the WHO, anxiety and stress is at an all-time high with nearly 20% of young people at risk of depression and mental health issues [77].
- There is a strong assumption that social media and technology are among the main determinants of this pandemic of mental morbidity among young people, together with other factors such as bleak futures, existential fear of climate change, use of alcohol/drugs or insufficient development of emotional self-control. However, it is extremely difficult to establish a direct causal link for each of these factors and to weight its relative importance because all the factors are deeply interlinked and they can be analysed in multiple ways. For example, social media can be a source of harassment and cyberbullying. Insufficient development of emotional self-control strengthens the risk of being manipulated online. A fast development of digital technologies creates many uncertainties (e.g. future of work, irrelevance), and these uncertainties greatly contribute to the perception of bleak futures. Furthermore, these factors also have a very complex interplay with a wide range of other developmental, educational, socio-cultural or historical factors that all contribute to determining overall well-being [72].
- The consequence of this complexity is that so far, the research results of individual studies remain mixed and contrasted. For example, in the case of social media it appears that its impact on mental health can be both positive and negative. It seems that the way we use social media is more important than the mere exposure [78]. Similar conclusions emerged in the heated debate on the impact of video gaming on mental health that has finally brought about scientific consensus that the gaming motives (achievement and escapism in particular) and the social context in which gamers play were much more powerful predictors of mental health disorders than raw playtime [79]. If the digital environment is not the only determinant of mental morbidity among young people, there is however a broad perception that it plays an important role in amplifying and accelerating the impacts of other factors.
- The fragmentation of the evidence on the impacts of the digital transformation on cognitive and socio-emotional capacities and on mental health remains an important problem. It makes it difficult to draw clear conclusions. A large meta-study would greatly help to make sense of the existing evidence, better understand the interactions between the various factors and their relative importance, identify knowledge gaps, guide future research and inform policy interventions.

- It is also important to note that while most of the current public debate focuses on the potential negative impacts of the digital transformation on mental health, many AI experts prefer to highlight the omnipotent benefits of using AI-enabled systems to address certain cognitive conditions or mental health issues. Some of these AI-enabled interventions remain quite simple, such as the use of closed-loop digital meditation, which proved to be very effective for individuals with emotional deprivation from childhood [80] or in specific domains such as improvement in attention span [81]. Other interventions are much more complex, such as the use of brain implants or AI extenders [see Takeaway n°1]. AI extenders can offer unique insights into how we understand, assess and treat certain mental health disorders. They have the potential in a near future to bring huge benefits to people suffering for example from Alzheimer, learning disorders, addiction, borderline personality disorders or autistic disorders. They would allow for a high level of personalisation based on the collection and the analysis of large sets of data about an individual (not only biological data, but also information about one's behaviour, social life, screen time, the environments one spends time in, etc.). AI extenders could then directly address the cognitive deficits or mental disorders by compensating for the malfunctioning biological mechanisms. However, using AI extenders for these purposes still poses extremely complex philosophical, ethical and legal challenges linked to safety, reliability, the degree of dependency, responsibility, the risk of interference and manipulation, or even the impact on personal identity [20].



### Takeaway n°8:

## Creativity will no longer be a human prerogative

- There is still a strong belief that the arts, and more generally the creative sector, will be less affected by AI than most other sectors and that they will provide the ultimate sanctuary against automation. The fast development of AI does however challenge this assumption in many ways. Many creative professionals will have to adapt to this new reality, to reshape their practices and to find new forms of complementarities with machines.
- The growing importance of AI in the cultural and creative sectors is the latest development in a larger trend. It started many years ago with the digitisation of a high proportion of creative content, such as images, sound and text, as well as with the deployment of digital tools in a wide variety of contexts including museums or live performances. The COVID-19 crisis has hugely accelerated this trend. The first major consequence of this massive deployment of digital tools is a deeper personalisation and flexibility. People are now able to take part in cultural activities from anywhere in the world at any time. This personalisation is reinforced by the recommendations made by algorithms [see Takeaway n°2]. The second consequence is that the division between cultural content producers and consumers is blurring. This blurring can take various forms. Platforms, such as 'Bandcamp' enable a direct interaction between artists and their fan base at the various stages of the creative process. This can influence artists and improve their connection with the public. In some cases, the interaction can lead to highly original forms of co-creation. In South Korea, series are shot while they are being broadcast, with only a few episodes being ready in advance, so that screenwriters and producers can adjust the storyline depending on the viewers' feedback [82]. Another very interesting example of interactive art is the 'teamLab' in the Tokyo Digital Museum where algorithms, programmers and visitors co-create unique visual experiences [83]. The blurring can go even further because nowadays the new digital content production tools enable just everyone to have access to cheap and easily usable software that can enable the generation of semi-professional or even professional content. Many people elaborate on already existing content through various remixing or curation practices. Others prefer to generate new content that can then also be very broadly made available through social media or other platforms [42, 84]. This is of course particularly true for digital content, but other forms of art are also concerned and nowadays you can produce stone or wood sculptures without any other technical skill than programming.

- Creative industries have been very fast in understanding all the new opportunities that digitisation and now AI is offering them to create new products, content and services and to share their creations. AI is therefore spreading fast in fields as diverse as film, video, TV, games, advertising, architecture, fashion, design or music [85]. To take just one example, in design many companies have fully recognised how AI could help them to generate meaningful products and services and to develop truly personalised solutions. Consequently, many universities in Europe and across the world, such as the Delft Technical University have now launched specific AI and design programmes to build skills in this area [86].
- Beyond creative industries, many individual artists are now also increasingly using AI either as source of inspiration or as a support to create new forms of art. The examples of fascinating interactions between artists and machines are numerous and they can take multiple forms. For instance, David Cope has been, without any doubt, one of the pioneers in this area. The first programme he has developed was named 'EMI' (Experiments in Musical Intelligence) and was specialised in composing chorales in the style of Johann Sebastian Bach, which in the opinion of many listeners sounded more 'Bach than Bach.' He later created 'Annie,' which composes music or haikus in response to input from the outside world. Cope himself has no idea what Annie is going to compose next [87]. Data in itself can become the raw material for artistic creations. For example, Refik Anadol is using huge sets of data and machine learning to create what he calls data sculptures [88]. AI is also finding its place on the arts market. In 2018, Christie's sold Portrait of Edmond Belamy generated by AI for an incredible \$432 500 [89]. AI can also very convincingly write creative texts as demonstrated by GPT-3, Open AI's language generator [90]. Many more examples could be mentioned [91] and what is striking is that it is extremely difficult to make the difference between content generated by AI from content generated by humans. It is even more difficult to clearly delineate the respective roles of the programmer and of the machine at every step of the creative process, which as we will see below will raise extremely complex copyright questions.
- Creativity is not reserved for the artistic disciplines; it can also be a key component of scientific research. For example, in the case of quantum physics an AI system based on machine learning has designed autonomously complex quantum optics experiments. Designing such experiments is generally considered as difficult and it often clashes with human intuition. The AI system has learned to design these experiments without relying on previous knowledge or intuition and proved to be more efficient than the previous approaches. In the process, it also discovered new experimental techniques that are now becoming standard in modern quantum optical experiments. This was not explicitly demanded from the system but emerged through the process of learning. The scientists who led the experiment are convinced that machines could have a significantly more creative role in future research [92].
- AI does not only boost the creativity of artists or scientists and open new territories for them, it is also a fantastic tool to help us to better understand how creative processes actually work [93]. Research is developing fast in this area using big data sets, but also EEG and MRI in controlled experiments. Many different angles are explored, and many scientists are trying to investigate the factors that enable certain creative professionals to achieve success, including 'hot streaks,' i.e. runs of



continued successful creativity over certain periods of time [94]. This can be coupled to big data analysis on the perception of art works by humans, which is highly attractive for creative industries. For example, in music there is huge interest in trying to better understand what components make a song a hit, and the results are increasingly being used by composers and music labels.

- There is broad consensus that AI can help to understand and to boost creativity. However, the question whether AI can really be considered as creative remains much more disputed. Margaret A. Boden who is one of the pioneers and one of the most revered scholars in the field of computer models of creativity believes that this is actually not a scientific question but a philosophical one to which there is no clear answer. According to her, there are three main types of creativity involving different ways of generating novel ideas. The first, 'combinational' creativity, involves novel combinations of familiar ideas. The second, 'exploratory' creativity, involves the generation of novel ideas by the exploration of pre-defined and structured conceptual spaces. The third, 'transformational' creativity, involves the transformation of some dimensions of the conceptual space so that new structures can be generated which could not have arisen before. Computer models of creativity include examples of all three types. So far, those focussing on the exploratory type have been the most successful (this is for example what David Cope used for the creation of his Bach inspired chorales). Where AI still needs to progress is in self-criticism and in the capacity to evaluate the interest and the value of the novel ideas it generates. But this is also not obvious for humans because what counts as 'creative' and what is called 'discovery' depends largely on unarticulated values and the novelties which we are prepared to approve as 'creative', change from culture to culture and from time to time [93, 95]. Research aiming at developing openness, curiosity and creativity in AI systems is currently gaining a lot of traction. The goal of this research is to create systems that are not only 'problem-solving,' but also 'problem-seeking,' or in other words finding solutions where we are not yet aware that there is actually a problem. Many experts believe that, just as for intelligence, machines have already and will continue to develop their own flavour of creativity. The most interesting forms of machine creativity may be the ones that will be very different to the human forms of creativity, and whose outcomes might even be difficult for us to understand. This should of course be handled with great care and boundaries are needed to keep the experiments open-ended and creative AI safe.
- A related and even more disputed question is whether machines can actually be considered as artists or not. As we have seen earlier, when people hear a piece of music, watch a painting or read a text produced by an AI system without being informed about its provenance, it is in general very difficult for them to make the difference with a piece of art produced by a human. However, when they are informed about the provenance, it is very difficult for them to admit that the machine is really 'the creator' and has not simply been providing some kind of assistance to an artist. They are also very quick at declaring that the piece of art lacks something, whether it is soul or an authentic human experience. Once again, this is a very difficult question and the answer depends largely on the definition we give to the word 'art.' There is broad consensus among experts that the artist actually only starts a process and that it is the observer (critic or public) who really gives meaning to the artist's work. The same work in two different contexts can change from

irrelevance to a masterpiece depending on the societal discourse. Pier Luigi Sacco very interestingly notes that the idea of the artist as we know it today is mainly a construct of Western culture. The notion of the artist is already increasingly challenged by the development of new forms of co-creation where the artist is less and less the 'soloist' and more and more the 'conductor'. For him, machines are probably even more able than humans to facilitate processes of massive co-creation and this could become their creative specificity much more than simply creating the machine equivalent of studio art. He believes that machines can be artists, but not necessarily in the conventional sense, and that they should break new ground rather than reproduce traditional human models of creating art [42]. Many other challenging questions emerged in our conversations with experts, such as will machines produce their own definition of art or how will art produced by machines be perceived by other machines? [96]

- We could even go one-step further by asking if we will still need artists in the future. Some scientists and thinkers have started to imagine futures where music could be personalised to fit as closely as possible to the wishes and the needs of the listener. Algorithms could for example read our emotions and our biochemical reactions in real time while we are listening to a song. On this basis, it would be easy for them to select the next songs to be played in order to maintain or change our state of mind. They could also tinker with the songs or melodies themselves and edit or rewrite the parts we like less in order to adjust them as closely as possible to our tastes. They could even create completely personalised melodies based on everything they already know about us [2]. Platforms could also be developed to fulfil all the requests of any listener. Subscribers would have limitless access to individualised music experiences and could for example ask to hear Led Zeppelin performing a particular ABBA song [97]. While these scenarios are still very much science fiction, the underlying technologies could be a reality in a not-too-distant future. They would bring about a complete revolution of the music ecosystem, as we know it. Neither performers, producers or recording engineers, nor distribution systems would be needed any longer because all music would be (re)created locally.
- Without going that far, the growing role played by AI in the arts is already starting to challenge the established norms in copyright. The question of the ownership of material generated by AI systems that may themselves have been trained on copyright protected material is emerging. If we take the case of music, AI is already able to compose songs with limited human intervention. Authorship recognition may therefore require a deep analysis of the operation of the systems and the role of the different actors involved in the process such as the developer, the person behind data generation or the user. It would also require clarifying the degree to which AI is involved in each stage of the process of music composition, performance and production. This will be extremely complex. More generally, copyright in the age of AI generates many questions for which the answers are still far from being clear. We need to consider them urgently as AI technologies continue to develop fast and are applied more and more in domains centred on human creativity [97].



### Takeaway n°9:

#### **AI risks creating the most unequal societies that ever existed**

- AI will bring many opportunities to improve human well-being, but because of the speed at which it is developing and of its disruptive nature, it also risks creating the most unequal societies that ever existed. It is indeed very likely that not all humans will be able to take advantage of AI to the same extent. Some will be able to make a smart use of the new technologies that are developing in order to keep at least a certain level of free will and of autonomy, to take better decisions and to attain their goals. Many others will probably be left behind, and this might unleash unprecedented gaps between the winners and the losers of the AI revolution. Most of the key factors behind this possible divide are not new but might play an even greater role than ever before. They include the social background, the geographic origin, age, gender and ultimately the access to quality education. There is one additional factor that will play a crucial role both at individual and collective level: the access to data.
- We are already observing an unequal data flow from Europe and from the Global South towards the US and Chinese tech giants. History has demonstrated that with each industrial revolution, certain countries (mainly from the West so far) have been able to impose their economic dominance thanks to their early adoption of scientific and technological discoveries. It is very likely that the economic and technological divide will continue or even accelerate in the time of the fourth industrial revolution. It is in particular very difficult to imagine that the very slow technological progress of many developing countries in AI will lead to a fair and equal adoption of new technologies. It is very likely that places with lesser innovation capacity will lag far behind.
- The most discussed aspect so far in relation to potential inequalities has been the future of work. Despite the multiple debates that are organised, the truth is that we do not know much. We have no idea at all what the job market will look like in 2050. One of the most common assumptions is that jobs that require a lot of creativity and imply many human interactions will be less at risk of being automated. We have however seen that AI is already developing some form of creativity, some form of emotions and empathy and some form of intuition. No human job will ever be safe from the threat of future automation. There is therefore very little chance to predict the type of jobs that the children who are in formal education now will have. Therefore, it is also difficult to predict the type of skills that they would need for these jobs.

- If we do not know what the jobs of the future will be, what we can already anticipate is that due to the exponential speed of the changes that are underway, people will most probably have to change the nature of their job many times during their lifetime. In the previous generations, it was still quite common not to change job during one's entire working life. The generation of those who are currently in their 40's or 50's is changing jobs 2 or 3 times on average. The generation of those who are now in elementary or secondary education may need to change the nature of their jobs 5 to 10 times during their lives. Constant change will become the most important feature of life and the key challenge will be the capacity to remain relevant. This will become even more important with a significant increase of life expectancy [see Takeaway n°4]. Change as such is a stressful activity and with age and the decrease of brain neuroplasticity, this will require a lot of energy and a very high mental and psychological balance.
- Already today, jobs in advanced economies involve increasingly unprotected temporary work, freelancing and one-time gigs. It is quite likely that this trend will accelerate in the coming years with all the negative consequences on human well-being. On the other hand, AI may lead to the empowerment of individuals who will be able to take advantage of available technology and build new business models. These individuals could use AI to augment their own capacities, and therefore significantly increase the added-value of human activity. They could also use the full potential of AI systems to develop all aspects from product or service design to marketing and after-sales care. Under these conditions, one could for example imagine that three people using a few online gig-workers (human or AI) could create a new service and ensure its smooth delivery for which a medium-size company would be needed today [57, 98].
- We should however not be too optimistic and simply assume that enough new jobs (even precarious ones) will appear to compensate for all the jobs we will lose. It might also well be that we will move towards post-work societies where there will simply be no jobs for a vast majority of people. This would open extremely complex questions that we should better start to explore now if we want to be prepared in case such a scenario becomes reality. Harari quite provocatively writes that we might witness the rise of a new 'useless' class and he wonders what all the 'superfluous' people will do all day. He also stresses that the social, economic and political models we inherited from the past are inadequate for dealing with such a challenge [2,4]. The first key issue to be addressed would be how to provide for people's basic needs. So far, the elites took care of the masses mainly because they had either a military, an economic or a political relevance. However, this relevance might vanish, and it would be an important change of mind-set for most of the elites to have to take care of the masses even if they do not need them any longer. Public authorities would have a crucial role to play in switching from protecting the jobs to protecting the workers. One interesting model that is already being explored by some countries is 'universal basic income.' This model would certainly deserve to be further analysed and on a larger scale. The second key issue is how to protect people's social status and self-worth. New forms of social status independent from work would have to be (re)discovered and valued, and the additional time people would have available by working less could become a blessing if it is used for self-development and human flourishing. This is however very far from being self-evident and is deeply linked to education [see Takeaway n° 10].



- The risks of future inequalities go however far beyond work. As already hinted above, AI will provide or facilitate multiple ways to augment the physical or cognitive capacities of humans through AI extenders, implants, genetic re-engineering or other forms of bioengineering [see Takeaways n°1 and 4]. If these interventions prove to be expensive, which most of them will certainly be at least in a first stage, humankind might split into a small and privileged class of upgraded humans (or 'superhuman') and a massive underclass that would be left behind. Throughout history, the elites always liked to imagine that they had superior skills to everybody else and that this was why they were in control. Evidence generally showed however that in most cases they owed their privileges rather to unjust legal and economic discrimination. In the future however, the rich might really become more talented, more creative and more intelligent than the poor thanks to the 'upgrades' they will be able to afford [2]. These inequalities could start before birth if we move towards genetic child catalogues [4] and they could continue long after death through 'mind uploading' technologies without even speaking about physical immortality.



### Takeaway n°10:

### Education is our best bet

- The debate about the digital transformation and education is still too often reduced to the need for people to develop the relevant digital skills for the jobs of tomorrow. However, education for the digital age is far more than that. The Commission, in its Digital Education action plan, adopted in January 2018 [99] and updated in September 2020 [100], stressed that it is crucial to equip as many people as possible with 'basic digital skills' and to give European students the opportunity to acquire the latest 'advanced digital skills.' But the Commission also highlighted that it is at least as important to educate people at all ages about how to navigate safely in a digital environment, about how to approach information critically and about the potential impacts of digital technology on well-being. This becomes even more important with the fast development of AI. While the development of AI systems will remain reserved to a selected few, AI will affect the lives and opportunities of everyone. Everyone should therefore have at least a basic understanding of what AI is and of how algorithms, data, statistics and probabilities work in order to make sense of what is happening around them and to be able to survive in this new environment.
- Digital education is only one tiny sub part of the education we will need for the age of AI. To remain relevant, our education systems will need a radical change in their purpose, form and content. As education is a sector that traditionally changes very slowly, the adaptation process must start now. The reasons behind the need for such a radical change are multiple. First, our education systems still focus mainly on the acquisition of knowledge and skills for professional success, but as seen earlier [see Takeaway n°9], it is incredibly difficult to predict what the jobs of the next decades will be. Therefore, much of what kids learn today will likely become irrelevant. Rather than to focus too much on the immediate needs of the economic system, we should therefore have a longer-term approach enabling people to be ready also for unknown futures. Second, the fundamental assumptions underlying the current education systems are that economic growth measured by GDP is determinant to human well-being and that human capital is a necessity for economic growth. This approach has resulted in an increase in material wealth, however recent studies have indicated a disconnect between economic growth and well-being. Covering material needs is an undeniable pre-requisite for well-being. However, the sole emphasis on the material side of life has come at the cost of increasing levels of anxiety, depression, insecurity and poor interpersonal relationships [101][see Takeaway n°7]. Third, there is a high probability that people will work less in the coming decades and that they will have more free time. It is therefore important to equip them with the tools to turn this into an opportunity to make their lives better. The concept of 'lifelong learning' should evolve into 'lifelong self-development' and individuals should have the freedom to choose how they want to develop. Self-awareness, creativity, the curiosity to learn new things and improved interpersonal relations could become

some of the building blocks of this new approach. Fourth, humans develop their cognitive and their socio-emotional capacities in order to tackle adaptive challenges. Our capacities as they stand are the result of a successful adaptation to a pre-digital environment. The speed of the digital/technological transformation and the fast-growing interaction between humans and machines will increasingly generate new adaptive challenges. We will therefore need to develop different strategies and heuristics than in a non-digital world. We are however still at the very beginning of this curve and it is therefore difficult to predict what precisely these new strategies will be. What is already clear is that socio-emotional stability and resilience, as well as the capacity to reinvent yourself will play a crucial role in this evolution [see Takeaway n°7].

- An education fit for the age of AI should therefore go beyond just materialistic wealth as the primary driver for human well-being, and include the notion of human flourishing. Human flourishing is traditionally defined as the state wherein an individual experiences positive emotion, acquires the abilities to achieve their own goals within themselves and with the external environment, and is able to interact positively with their social environment [101]. Furthermore, our future should be characterised by a smooth human-AI cooperation and complementarity rather than competition. What we should aim at is not to try to run faster than the algorithms as is sometimes suggested, but to co-evolve and to grow together harmoniously.
- Recent research from the neurosciences has demonstrated that to be able to achieve human well-being and flourishing, education must be based on a 'whole brain' approach and nurture both the cognitive and socio-emotional parts of the brain. The curricula today need therefore to be redesigned to build a range of abilities that include not only skills of literacy and numeracy, but also critical reasoning to develop intellectual intelligence and socio-emotional skills of empathy, emotional regulation and compassion for emotional intelligence. The whole brain approach must also take fully into account our amazing 'neuroplasticity,' or in other words, the capacity of our brains to form and reorganise new connections [101].
- So far, social and emotional learning remain neglected in most of our education systems and it is crucial to expand them significantly. Social and emotional connections and skills are a pre-requisite for higher-order cognition. In other words, without being emotionally stable, positive and socially connected, humans are unable to learn efficiently and work intellectually. A growing body of scientific research indicates that students' social and emotional competence not only predicts their school success, but also graduation at university, employment, financial stability, physical health, and overall mental health and well-being [68]. Emotional literacy and stability are also key to be able to navigate safely on the AI powered internet and to mitigate the risks of manipulation. UNESCO has carried out a major review of the existing social and emotional learning frameworks for education systems worldwide. The key competencies that emerged across this review can be categorised under three tiers. The first, 'the self,' sounds almost like a survival kit for the age of AI and encompasses attention regulation (the ability to concentrate and focus in the present), self-regulation (the ability to identify and recognise one's own emotions, thoughts and influences on behaviour), emotional regulation (the ability to regulate one's emotions, thoughts and behaviours effectively) and critical inquiry (a

process of collecting and analysing information and undertaking a critical analysis of the internal consistency in arguments, facts, data and conclusions). The second tier, 'the other and society' covers empathy, social awareness and relationship skills. The third, 'agency, behavioural change and action,' encompasses compassion, cooperation and responsible decision-making [75].

- The development of AI will bring a cascade of ever-bigger disruptions. The nature of work will change fast, many jobs will disappear and we might even move towards post-work societies. People will therefore need to repeatedly learn new skills, reinvent themselves and try to maintain their relevance. Change is however always stressful and the hectic world of the early 21st century has already started to produce a global pandemic of stress [2]. Strengthening people's resilience and their ability to deal with change and to preserve their mental balance and motivation in unfamiliar situations will therefore be key. There has been a steady increase of knowledge produced on psychological resilience over the past 10 years with a significant peak over the past 2-3 years. This research field is however dominated by the Anglo-Saxon world and the EU remains a minor player. Furthermore, the link between AI and psychological resilience has not really been covered so far with only 87 papers published over the past few years [12]. Beyond the above-mentioned socio-emotional competencies, creativity should become a key building block of our education systems to strengthen our resilience and our capacity to adapt to new challenges. Being curious and open minded and having the capacity to find new approaches to unexpected situations should increasingly become a core skill set for everyone. In Europe, Aalto University is one of the frontrunners in this area and has introduced 'radical creativity' as a cross-cutting approach for everyone in their community [102]. Together with some other EU countries, Finland is also one of the pioneers of an open-ended and phenomena-driven learning culture. The aim is to train all individuals to be able to grasp a challenge and to find new viewpoints, approaches and solutions to it. Rather than simply relying on existing knowledge, which is typically the kind of knowledge that will become obsolete with more AI, people are encouraged to think and act creatively [86, 103]. Last but not least, well-being should also be considered as a skill as such and become an integral part of curricula. The latest neuroscientific knowledge shows that well-being is a combination of four fundamental constituents: (i) individual resilience, (ii) stimulation of positive emotions, (iii) generosity and (iv) attention [104, 105].
- AI will not only be a challenge for our education systems, it will also offer many new opportunities to improve the ways in which we teach and learn. AI combined with neuroscience research is already starting to unlock the mysteries of how learning takes place. It will also progressively enable a true individualisation of learning, which is important because no two brains are identical and each brain learns in a different way. Individualised learning is not a new concept, but it has so far been very difficult to implement by a teacher in a classroom. AI powered platforms will therefore be able to provide a very useful complement to in-classroom instruction. These platforms will be based on algorithms that will not only teach maths or history, but also study each pupil, and dynamically measure their level of attention, assess learning, address gaps and track outcomes. Learning will be tailored to individual learning styles and content will be presented in a way that is interesting for the learner. This will progressively strengthen the intrinsic motivation of learners and



diminish the importance of the currently predominant extrinsic motivation system measured through grades and marks. Education will also progressively become transformative rather than just transmissive. While many different learning strategies can be used by these platforms, games in particular are very promising because they allow a focus on multiple skills at once: while learners work to improve their understanding of core concepts, they can also develop skills such as creativity, curiosity and persistence in the process [101]. Various models of AI powered digital platforms are currently being developed worldwide, including by many public bodies. The new European Digital Education Hub launched under the Digital Education action plan should play an ambitious role in this context and act as a laboratory in which the most promising models could be tested. Best practice should then be promoted across the EU and support for their rollout provided [100]. Specific attention needs to be paid to the privacy aspects and to the ethical dimension of individualised learning platforms. These issues are extremely complex because of the speed at which AI tools are developing and because many platforms gather at least some personal data that often enables a learner to be identified. Therefore, the risks need to be better understood and adequate safeguards need to be put in place [101]. It is also important to ensure that AI learning tools fully respect cultural differences and do not endanger cultural diversity.



## Part 2



## Actions

As mentioned in the introduction, the potential impacts of AI on society are huge and difficult to predict. One could propose many actions at different levels. Instead, we opted for a comprehensive and feasible set of actions that could be implemented at EU level and may also serve as inspiration for similar actions to be carried out at national or regional levels.



## Strengthening the knowledge base

### ➤ **Action 1 – Intelligence landscape**

Launch a mapping of different cognitive systems, including humans, non-human animals, AI systems, hybrids and collectives. This mapping should provide data, analysis and visualisations. It should serve as a tool to properly evaluate, compare and classify AI systems and to compare AI systems with human intelligence and other forms of biological intelligences. The preparatory work already carried out for an Atlas of intelligence could serve as a source of inspiration [27].

### ➤ **Action 2 – Mental health**

Launch a meta-study on the impacts of the digital transformation, and in particular of AI, on cognitive and socio-emotional capacities and on mental health in order to collect and make sense of the existing evidence, identify knowledge gaps, guide future research and inform policy interventions.

### ➤ **Action 3 – Values and well-being**

Launch a series of studies to build a deeper understanding of values and of well-being, of how they are interpreted across different cultures, and of how and why they change through time. The 'Science of values' research project launched by the Commission's Joint Research Centre [36], as well as the work carried out by the OECD on well-being [46] could serve as starting points. AI systems are already and will increasingly be involved in moral situations that directly affect the well-being of humans and where they have to make decisions autonomously. Therefore, these studies should also aim to better understand how AI developers across the world do already teach values to these AI systems. A key challenge for the future will be to reach a broad consensus at global level about which sets of values to use in AI systems, as well as how to interpret them. The endeavour of the UNESCO to develop universal ethical impact assessment principles is therefore extremely important and should be strongly supported by the EU.



#### ➤ **Action 4 – Recommender algorithms**

Launch a study on the positive and negative impacts of recommender algorithms on cultural diversity. This study should analyse in particular how effective the diversification and serendipity factors that are embedded in these algorithms are, how they could be improved if needed, as well as how they could be more broadly rolled out and by whom. It should also assess if in line with the existing practice in broadcasting, where we have 'a public broadcaster' acting in the defined public interest, it would be relevant to introduce 'public recommending algorithms' aiming at a similar result in the virtual environment.

#### ➤ **Action 5 - Copyright**

Launch a study on all the potential impacts of AI on copyright to identify more precisely the challenges we are already facing and to anticipate those that may emerge in the coming years.



### **Gearing-up policy-making**

#### ➤ **Action 6 – Regulatory and policy-making practices**

Start reflecting about new ways of developing regulations where policymakers, with the support of regulatory scientists, could follow much more closely the development of new technologies and work in parallel at the elaboration of the needed regulations. This effort should be at least EU wide and even worldwide since the deployment of most new technologies has become instantaneous and global. Feasibility studies and inter-institutional working groups should be established in this respect.

#### ➤ **Action 7 – Public administration**

Develop a new set of skills for public administrators in order to enable them to better use the opportunities and better address the challenges that will emerge with AI. This should go hand in hand with an effort to increase significantly transversal cooperation between all departments of public administration since AI is penetrating and influencing all aspects of life. The work on the new skills for public administrators should start as soon as possible in close cooperation with the EU Member states and international organisations such as the OECD and the UNESCO.

### ➤ **Action 8 – AI empowered policy-making**

Launch pilot projects to test a more systematic use of AI systems for ex-ante policy impact assessments and ex-post evaluations. Under this framework, an experiment should also be carried out in close cooperation with video game developers to develop tools that would translate dry technical descriptions of the potential impacts of policy options and provide visual experiences of these impacts on countries, regions and communities to policymakers as well as to citizens in order to facilitate debate and decision-making [61].



### **Assessing the impacts of AI systems**

### ➤ **Action 9 – AI cognitive extenders**

Classify AI cognitive extenders into dynamic categories according to their actual and potential impacts on well-being and on this basis design a regulatory framework. Whenever relevant, the regulatory conditions for AI extenders should be no less stringent than for any other medical device or software. Their safety and their mechanical and physical reliability should be assessed, but also their ethical behaviour as well as their potential for addiction (like for addictive pharmaceutical drugs). Ethical guidelines and a possible ethical impact assessment methodology for AI cognitive extenders should be widely discussed with experts as well as with citizens.

### ➤ **Action 10 - Ex-ante impact assessment tools**

Develop ex-ante impact assessment tools in order to evaluate both the intended and non-intended impacts of various applications of AI at the level of individual users and at the level of societies as a whole. These tools should be based on rigorous indicators and include as far as possible also well-being indicators. They should at least be co-created by public actors as most businesses will have potential conflicts of interest due to the prevailing business models, and engage all stakeholders. The work already carried out by the organisation 'GoodAI' could serve as one of the sources of inspiration for future work at EU level [29]. Besides developing new metrics and assessment methods, it would also be important to create reference testing centres or other types of testing infrastructures that would promote a comprehensive approach and provide the necessary expertise, and that could be used by companies or other relevant actors.



## Empowering citizens

### ➤ **Action 11 – New model for education**

Support EU Member States for the development of a new model for education that goes beyond just materialistic wealth as the primary driver for human well-being and includes the notion of human flourishing. This new model of education should be based on a 'whole brain' approach and include a strong component of social and emotional learning. A strong attention should be paid to emotional and psychological resilience, creativity and people's ability to adapt to change. Well-being should be considered as a skill and an appropriate framework should be established in order to become an integral part of curricula.

### ➤ **Action 12 - Explicit consent and opt-in mechanisms**

Explore the feasibility and effectiveness of various measures, tools and methods that could ensure that nudging by AI systems, including nudging for good, would only be possible based on explicit consent and opt-in mechanisms.

### ➤ **Action 13 - Customer-defending entities**

Examine the relevance and feasibility of supporting the development of 'customer-defending entities' that would inter-alia analyse customer behaviour with AI techniques to counterbalance the analytic power of corporations having access to the data of millions of customers. Examine also the relevance and feasibility of creating incentives for platforms that would be better aligned with European values and the public interest.

### ➤ **Action 14 – Participatory policy experimentations**

Launch policy experimentations similar to Experimental Finland at EU level to test innovative policies in close cooperation with stakeholders and citizens.



## Promoting AI based solutions for public good

### ➤ **Action 15 – Emotional temperature**

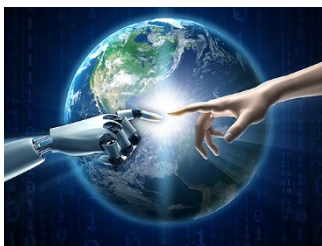
Develop new tools to take the emotional temperature and to sense citizen's concerns, fears, hopes and suffering more effectively in order to provide important new information to guide policy choices. For example, emotions indicators and maps could be developed by using text-mining and text-monitoring techniques of media and social media platforms. This could detect and classify emotions present in the media and aggregate them to obtain general levels of emotions expressed by citizens in different geographic locations [36].

### ➤ **Action 16 – Privacy protection tools**

Take stock of the technological and/or AI based solutions that are emerging to protect privacy online and analyse if any further support or incentives are needed at EU level. The relevance of supporting the development of open-source instruments in particular should be explored. These instruments could then be deployed by various developers with the aim to be used both in private and public sector applications, to help maintain and safeguard privacy.

### ➤ **Action 17 – Personalised education**

Identify all relevant AI powered tools for individualised learning, test them and promote best practice. The new EU Digital Education Hub should play a central role in this process. Particular attention should be paid to the privacy aspects and the ethical dimension of these tools.



## Exploring futures

### ➤ **Action 18 – Life expectancy and mind uploading**

Launch a foresight exercise to explore various scenarios of life expectancy increases and examine their potential impacts on family structures, the future of work or welfare. This exercise should cover also mind-uploading and digital immortality. It could thus help to pave the way for the design of a policy dealing with the rights linked to the digital footprint after death.

### ➤ **Action 19 – Liberal democracy**

Launch a foresight exercise on the future of liberal democracy. This exercise could take place in the context of the reflection about the 'Future of Europe' and could build on some of the work already carried out by the Commission's Joint Research Centre in this area [53, 65].

### ➤ **Action 20 – Post-work society**

Launch a foresight exercise to explore several scenarios for a post-work society and examine their potential impacts on wealth concentration or redistribution. Various dimensions should be covered including potential sources for public funds, different ways to provide for people's basic needs, new forms of social status independent from work, as well as new ways to create a sense of purpose and of self-worth and to better value non-economic activities.

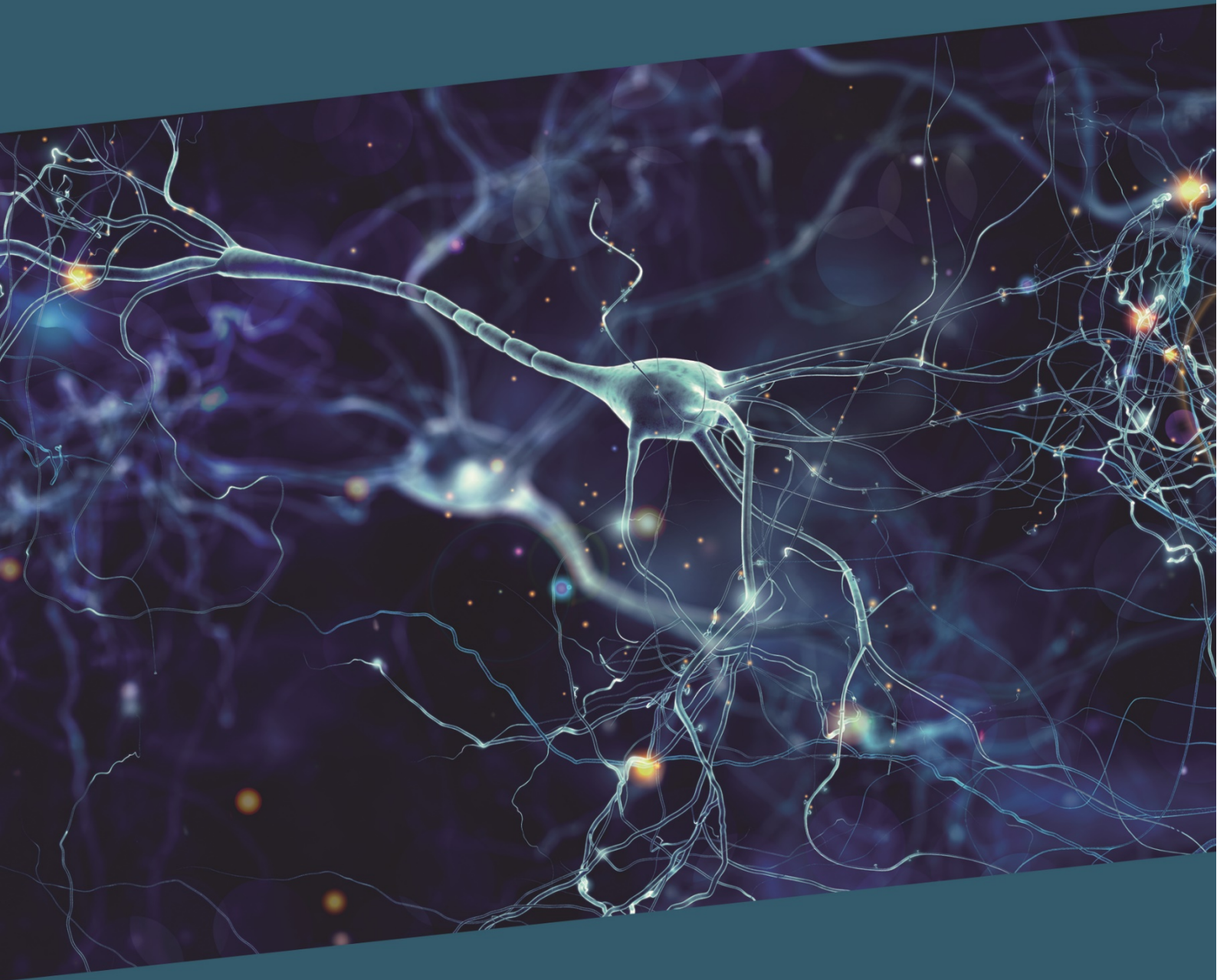
### ➤ **Action 21 – Geographical divide**

Launch a study or a foresight exercise on the possible impacts of a different pace in the adoption of technology in different parts of the world.





# **Part 3**



**Acknowledgements**  
**References**  
**Annex**



## Acknowledgements

We would first and foremost like to thank all those who kindly took the time to answer our questionnaire, to send us contributions and to share references of relevant books and articles with us. Without them, this report would not have been possible. Their names are listed below.

We also had the privilege to have a series of conversations at various stages of the preparation of this report with Alessandro Annoni, Peter Gluckman and Kristiann Allen, Emilia Gomez-Gutierrez, Marek Havrda, José Hernández-Orallo, Alexander Peterhaensel, Pier Luigi Sacco, Roy Saurabh, Paweł Świeboda and Anna Valtonen. Their precious advice has guided our reflections and their comments on the manuscript have greatly helped us to improve the text.

Finally, we would like to thank Alexandra Balahur, Laura Cassio, Marion Dewar, Gianluca Misuraca and Andrea Saltelli who reviewed the text and provided many thoughtful suggestions.

### *List of Contributors*

1. Alessandro Annoni
2. Maria Bielikova (Kempelen Institute of Intelligent Technologies)
3. Tommaso Calarco and Daniele Binosi (Forschungszentrum Jülich)
4. Stephane Chaudron (JRC)
5. Max Craglia (JRC)
6. Kate Crawford (Eviva)
7. Béatrice D'Hombres (JRC)
8. Paul Dujardin, Anna Vondracek, Emma Dumartherey (Bozar)
9. European University Institute – Florence
10. Peter Gluckman and Kristiann Allen (University of Auckland, Kōi Tū: The Centre for Informed Futures)
11. Emilia Gomez-Gutierrez and the Team of the HUMAINT Project (JRC)
12. Heikki Hämmäinen (Aalto University)
13. Marek Havrda, Olga Afanasjeva and Will Millership (GoodAI)
14. José Hernández-Orallo (Universitat Politècnica de València, Leverhulme Centre for the Future of Intelligence)
15. Akitoshi Honda (CHORA – Conscious City Lab, Berlin)

16. Vishal Kumar (Cultural Data Scientist)
17. Juha-Matti Liukkonen, Johan Himberg and Tommi Pajala (Reaktor Innovations Oy, Finland)
18. Markku Markkula and Taina Tukiainen (Committee of the Regions, Aalto University)
19. Valentina Montalto (JRC)
20. Rubén Moreno-Bote (Universitat Pompeu Fabra Barcelona)
21. Geoff Mulgan (University College London)
22. Ilkka Niemelä and Aalto University
23. Ray Perrault (SRI International)
24. Alexander Peterhaensel (Technische Hochschule Brandenburg, University of Applied Sciences Brandenburg, artist)
25. Pier Luigi Sacco (IULM University, MetaLab Harvard)
26. Michaela Saisana (JRC)
27. Roy Saurabh (UNESCO)
28. Isabelle Schwarz and André Wilkens (European Cultural Foundation)
29. Richard Simpson (Meta Moto)
30. Bill Simpson-Young (Gradient Institute, University of Sydney)
31. Paweł Świeboda (Human Brain Project & EBRAINS AISBL)
32. Anna Valtonen (Aalto University)
33. Paul Wiersbinski (Lacuna Lab, artist)
34. David Wortley (International Society of Digital Medicine)



## References

- [1] Max Tegmark. *Life 3.0. Being Human in the Age of Artificial Intelligence*. Allen Lane, London, 2017.
- [2] Yuval Noah Harari. *21 Lessons for the 21st Century*. Jonathan Cape, London, 2018.
- [3] <https://www.bbc.com/news/technology-30290540>
- [4] Yuval Noah Harari. *Homo Deus. A Brief History of Tomorrow*. Vintage, London, 2017.
- [5] Yann Le Cun. *Quand la machine apprend*. Odile Jacob, Paris, 2019.
- [6] Joint Research Centre. *Artificial Intelligence. A European Perspective*. European Commission, 2018: <https://ec.europa.eu/jrc/en/publication/artificial-intelligence-european-perspective>
- [7] European Commission. *White Paper on Artificial Intelligence: a European approach to excellence and trust*, 2020: [https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust\\_en](https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust_en)
- [8] High-Level Expert Group on AI. *Ethics Guidelines for Trustworthy Artificial Intelligence*, 2019: <https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>
- [9] European Commission. *Proposal for a Regulation of the European Parliament and of the Council on a Single Market For Digital Services (Digital Services Act) and amending Directive 2000/31/EC*, 2020: [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-services-act-ensuring-safe-and-accountable-online-environment_en)
- [10] European Commission. *Proposal for a Regulation of the European Parliament and of the Council on contestable and fair markets in the digital sector (Digital Markets Act)*, 2020: [https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/europe-fit-digital-age/digital-markets-act-ensuring-fair-and-open-digital-markets_en)
- [11] The HUMAINT project in the European Commission's Joint Research Centre has analysed some of these challenges. See for example, Joint Research Centre. *HUMAINT: Understanding the impact of Artificial Intelligence on human behaviour*. European Commission, 2021. <https://ec.europa.eu/jrc/communities/en/community/humaint/document/humaint-brochure>
- [12] For all the analyses of scientific research and patents in the field of AI, we used *Tools for Innovation Monitoring* (TIM) developed by the Joint Research Centre: [https://knowledge4policy.ec.europa.eu/text-mining/topic/tim\\_analytics\\_en](https://knowledge4policy.ec.europa.eu/text-mining/topic/tim_analytics_en)
- [13] <https://www.humanbrainproject.eu/en/>
- [14] <https://ebrains.eu/>
- [15] José Hernández-Orallo. *The Measure of All Minds. Evaluating Natural and Artificial Intelligence*. Cambridge University Press, 2017.
- [16] Stanislas Dehaene. *Le Code de la conscience*. Odile Jacob, Paris, 2014.



- [17] Stanislas Dehaene, Yann Le Cun, Jacques Girardon. *La Plus belle histoire de l'intelligence. Des Origines aux neurones artificiels : vers une nouvelle étape de l'évolution*. Robert Laffont, Paris, 2018.
- [18] Daniel Kahneman. *Thinking, Fast and Slow*. Allen Lane, London, 2011.
- [19] José Hernández-Orallo, Karina Vold. AI Extenders: the ethical and societal implications of humans cognitively extended by AI. *Proceedings of the 2019 AAAI/ACM Conference on AI, Ethics, and Society*, 507-513, 2019: <https://dl.acm.org/doi/10.1145/3306618.3314238>
- [20] José Hernández-Orallo, Karina Vold. AI Extenders and the ethics of mental health. In Marcello Ienca and Fabrice Jotterand (editors), *Ethics of Artificial Intelligence in Brain and Mental Health*, Springer, forthcoming: <https://philpapers.org/archive/VOLAEA.pdf>
- [21] John B. Bossaer, Jeffrey A. Gray, Stacy E. Miller, Gavin Enck, Vamsi C. Gaddipati, Robert E. Enck. The Use and misuse of prescription stimulants as "cognitive enhancers" by students at one academic health sciences center. *Academic Medicine*, 88 (7), 967-971, 2013: [https://journals.lww.com/academicmedicine/Fulltext/2013/07000/The\\_Use\\_and\\_Misuse\\_of\\_Prescription\\_Stimulants\\_as.26.aspx](https://journals.lww.com/academicmedicine/Fulltext/2013/07000/The_Use_and_Misuse_of_Prescription_Stimulants_as.26.aspx)
- [22] <https://www.nature.com/news/2008/080409/full/452674a.html>
- [23] <https://neuralink.com>
- [24] <https://paradromics.com/>
- [25] <https://www.synchronmed.com/>
- [26] NESTA. *The Future of Minds and Machines: How artificial intelligence can enhance collective intelligence*, 2020: <https://www.nesta.org.uk/report/future-minds-and-machines/>
- [27] Sankalp Bhatnagar et al. Mapping Intelligence: Requirements and possibilities. *Proceedings of the 3<sup>rd</sup> Conference on Philosophy and Theory of Artificial Intelligence*, Springer, 2017: [https://doi.org/10.1007/978-3-319-96448-5\\_13](https://doi.org/10.1007/978-3-319-96448-5_13)
- [28] Gaspard Koenig. *La Fin de l'individu. Voyage d'un philosophe au pays de l'intelligence artificielle*. Editions de l'Observatoire, Paris, 2019.
- [29] Marek Havda, Bogdana Rakova. Enhanced well-being assessment as basis for the practical implementation of ethical and rights-based normative principles for AI. *2020 IEEE International Conference on Systems, Man, and Cybernetics*, 2754-2761, 2020: <https://ieeexplore.ieee.org/abstract/document/9283137>
- [30] NESTA. *AI ethics and the limits of code(s)*, 2019: <https://www.nesta.org.uk/blog/ai-ethics-and-limits-codes/>
- [31] <https://recsys.acm.org/recsys20/>
- [32] Lorenzo Porcaro, Carlos Castillo, Emilia Gomez. Music recommendation diversity: a tentative framework and preliminary results. *Proceedings of the 1st Workshop on Designing Human-Centric MIR Systems*, 2019: [https://repositori.upf.edu/bitstream/handle/10230/43113/porcaro\\_hcmir19\\_music.pdf?sequence=1&isAllowed=y](https://repositori.upf.edu/bitstream/handle/10230/43113/porcaro_hcmir19_music.pdf?sequence=1&isAllowed=y)
- [33] Jennifer S. Lerner, Ye Li, Piercarlo Valdesolo, Karim S. Kassam. Emotion and decision-making. *Annual Review of Psychology*, 66, 799-823, 2015: <https://www.annualreviews.org/doi/full/10.1146/annurev-psych-010213-115043>

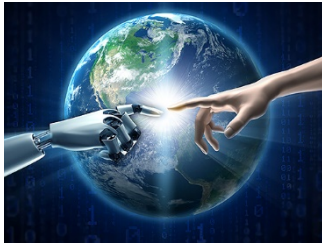
- [34] Richard H. Thaler, Cass R. Sunstein. *Nudge. Improving decisions about health, wealth and happiness*. Penguin Books, London, 2009.
- [35] Richard H. Thaler. *Misbehaving. The Making of behavioral economics*. Norton, New-York, 2015.
- [36] Joint Research Centre. *Understanding our Political Nature. How to put knowledge and reason at the heart of political decision-Making*. European Commission, 2019: <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/understanding-our-political-nature-how-put-knowledge-and-reason-heart-political-decision>
- [37] Antonio Damasio. *Descartes' Error: Emotion, reason and the human brain*. Putnam Publishing, New-York, 1994.
- [38] See for example *The Social Dilemma* on Netflix: <https://www.netflix.com/lu-en/title/81254224>
- [39] <https://knowledge.insead.edu/entrepreneurship/the-tiktok-strategy-using-ai-platforms-to-take-over-the-world-11776>
- [40] <https://mitsloan.mit.edu/ideas-made-to-matter/emotion-ai-explained>
- [41] <https://people.ict.usc.edu/~gratch/>
- [42] Conversation with Pier Luigi Sacco.
- [43] <https://maliciousaireport.com/>
- [44] Antonio Damasio. *The Strange Order of Things. Life, Feeling and the Making of Cultures*. Pantheon Books, New-York, 2018.
- [45] Jason Borenstein, Ronald C. Arkin. Nudging for good: robots and the ethical appropriateness of nurturing empathy and charitable behavior. *AI & Soc*, 32, 499–507, 2017 : <https://doi.org/10.1007/s00146-016-0684-1>
- [46] <https://www.oecd.org/statistics/better-life-initiative.htm>
- [47] <https://beyondstandards.ieee.org/ieee-7010-2020-launch-prioritizes-human-well-being-and-environmental-sustainability-via-technology/>
- [48] OECD. *Hello, World! Artificial Intelligence and its Use in the Public Sector*. 2019. <https://www.oecd.org/governance/innovative-government/working-paper-hello-world-artificial-intelligence-and-its-use-in-the-public-sector.htm>
- [49] Joint Research Centre. *AI Watch - Artificial Intelligence in public services*. European Commission, 2020. <https://publications.jrc.ec.europa.eu/repository/handle/JRC120399>
- [50] Joint Research Centre. *Artificial Intelligence and Digital Transformation: Early lessons from the COVID-19 crisis*. European Commission, 2020: <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/artificial-intelligence-and-digital-transformation-early-lessons-covid-19-crisis>
- [51] <https://www.bbvaopenmind.com/en/humanities/beliefs/the-social-contract-2-0-big-data-and-the-need-to-guarantee-privacy-and-civil-liberties/>
- [52] <https://www.weforum.org/agenda/2019/03/why-we-need-a-new-social-contract-for-data-in-healthcare/>

- [53] Joint Research Centre. *Technology and Democracy: Understanding the influence of online technologies on political behaviour and decision-making*. European Commission, 2020: <https://ec.europa.eu/jrc/en/publication/eur-scientific-and-technical-research-reports/technology-and-democracy>
- [54] <https://www.businessinsider.com/opinion-google-apple-contact-tracing-app-troubling-governments-2020-6>
- [55] Alexander Peterhaensel. Smile to vote: towards political physiognomy analytics - predicting electoral behavior from live video. *Proceedings of the 24th International Symposium on Electronic Art (ISEA)*, 79-85, 2018: <http://www.isea-international.org/isea2018/wp-content/uploads/2018/07/ISEA-proceedings-final.pdf>
- [56] Conversation with José Hernández-Orallo.
- [57] Conversation with Marek Havrda.
- [58] <https://solidproject.org/>
- [59] European Commission. *Communication on the European Democracy Action Plan*, 2020: [https://ec.europa.eu/info/strategy/priorities-2019-2024/new-push-european-democracy/european-democracy-action-plan\\_en](https://ec.europa.eu/info/strategy/priorities-2019-2024/new-push-european-democracy/european-democracy-action-plan_en)
- [60] The Center for the Governance of Change. *European Tech Insights 2019*: <https://docs.ie.edu/cgc/European-Tech-Insights-2019.pdf>
- [61] Marek Havrda. Artificial intelligence's role in community engagement within the democratic process. *International Journal of Community Well-Being*, 3, 437-441, 2020: <https://doi.org/10.1007/s42413-020-00100-8>
- [62] <https://www.weforum.org/agenda/2018/10/could-robot-government-lead-better-current-politicians-ai/>
- [63] <https://dataconomy.com/2017/09/artificial-intelligence-political-systems/>
- [64] <https://dt.asu.edu/>
- [65] Joint Research Centre. *The Future of Government 2030+. A Citizen centric perspective on new government models*. European Commission, 2019. <https://ec.europa.eu/jrc/en/publication/future-government-2030>
- [66] Anna Valtonen. Approaching Change with and in Design. *She Ji: The Journal of Design, Economics, and Innovation*, 6 (4), 505-529, 2020: <https://www.sciencedirect.com/science/article/pii/S2405872620300642>
- [67] Terry Irwin. The Emerging Transition Design Approach. 2018: [https://www.researchgate.net/publication/329155155\\_The\\_Emerging\\_Transition\\_Design\\_Approach](https://www.researchgate.net/publication/329155155_The_Emerging_Transition_Design_Approach)
- [68] [https://kokeilevasuomi.fi/en/frontpage?p\\_p\\_id=fi\\_yja\\_language\\_version\\_tool\\_web\\_portlet\\_LanguageVersionToolMissingNotificationPortlet&\\_fi\\_yja\\_language\\_version\\_tool\\_web\\_portlet\\_LanguageVersionToolMissingNotificationPortlet\\_missingLanguageVersion=1](https://kokeilevasuomi.fi/en/frontpage?p_p_id=fi_yja_language_version_tool_web_portlet_LanguageVersionToolMissingNotificationPortlet&_fi_yja_language_version_tool_web_portlet_LanguageVersionToolMissingNotificationPortlet_missingLanguageVersion=1)
- [69] [https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/161281/2-2019-KOKSU\\_raportti\\_.pdf?sequence=1&isAllowed=y](https://julkaisut.valtioneuvosto.fi/bitstream/handle/10024/161281/2-2019-KOKSU_raportti_.pdf?sequence=1&isAllowed=y)
- [70] <https://news.harvard.edu/gazette/story/2010/11/wandering-mind-not-a-happy-mind/>

- [71] Joint Research Centre. *Loneliness – an unequally shared burden in Europe*, 2018: [https://ec.europa.eu/jrc/sites/jrcsh/files/fairness\\_pb2018\\_loneliness\\_jrc\\_i1.pdf](https://ec.europa.eu/jrc/sites/jrcsh/files/fairness_pb2018_loneliness_jrc_i1.pdf)
- [72] Rochelle Menzies, Peter Gluckman, Richie Poulton. *Youth mental health in Aotearoa New Zealand: Greater urgency required*. Kōi Tū: the Centre for Informed Futures, University of Auckland, 2020: <https://informedfutures.org/youth-mental-health-in-aotearoa-nz/>
- [73] <https://www.cdc.gov/ncbddd/adhd/>
- [74] Melissa L. Danielson, Rebecca H. Bitsko, Reem M. Ghandour, Joseph R. Holbrook, Michael D. Kogan, Stephen J. Blumberg. Prevalence of Parent-Reported ADHD Diagnosis and Associated Treatment Among U.S. Children and Adolescents. *Journal of Clinical Child & Adolescent Psychology*, 47 (2), 199-212, 2018: <https://www.tandfonline.com/doi/full/10.1080/15374416.2017.1417860>
- [75] UNESCO MGIEP. *Rethinking Learning. A Review of social and emotional learning for education systems*. 2020: <https://mgiep.unesco.org/rethinking-learning>
- [76] National Board of Health Denmark. *Danes's Health – The National Health Profile 2017*. Copenhagen, 2017.
- [77] World Health Organisation. *Depression and other Common Mental Health Disorders. Global Health Estimates*. Geneva, 2017: [https://www.who.int/mental\\_health/management/depression/prevalence\\_global\\_health\\_estimates/en/](https://www.who.int/mental_health/management/depression/prevalence_global_health_estimates/en/)
- [78] Chloe Berryman, Christopher J. Ferguson, Charles Negy. Social Media Use and Mental Health among Young Adults. *Psychiatric Quarterly*, 89, 307–314, 2018: <https://link.springer.com/article/10.1007%2Fs11126-017-9535-6>
- [79] Marian Sauter, Tina Braun, Wolfgang Mack. Social context and gaming motives predict mental health better than time played: an exploratory regression analysis with over 13,000 video game players. *Cyberpsychology, behavior and social networking*, 24 (2), 94-100, 2020: <https://www.liebertpub.com/doi/10.1089/cyber.2020.0234>
- [80] Jyoti Mishra, Rajesh Sagar, Sana Parveen, Senthil Kumaran, Kiran Modi, Vojislav Maric, David Ziegler, Adam Gazzaley. Closed-loop digital meditation for neurocognitive and behavioral development in adolescents with childhood neglect. *Translational Psychiatry*, 10, 153, 2020: <https://www.nature.com/articles/s41398-020-0820-z>
- [81] David A. Ziegler, Alexander J. Simon, Courtney L. Gallen et al. Closed-loop digital meditation improves sustained attention in young adults. *Nature Human Behaviour*, 3, 746–757, 2019: <https://www.nature.com/articles/s41562-019-0611-9>
- [82] <https://koreancultureblog.com/2016/10/03/k-drama-series-live-shoot-system/>
- [83] <https://borderless.teamlab.art/>
- [84] Pier Luigi Sacco, Guido Ferilli, Giorgio Tavano Blessi. From Culture 1.0 to Culture 3.0: Three socio-technical regimes of social and economic value creation through culture, and their Impact on European cohesion policies. *Sustainability* 10 (11) 3923, 2018: <https://www.mdpi.com/2071-1050/10/11/3923>
- [85] NESTA. *The Art in the Artificial: the Implications of AI for the creative industries and how we should respond*, 2020: <https://www.nesta.org.uk/blog/art-artificial/>
- [86] Conversation with Anna Valtonen.
- [87] <https://open.spotify.com/artist/2e4gmURRiGni4WWyuiXI0?si=Ffqe7QcRQESLtuPN7D9OCQ>

- [88] <https://www.youtube.com/watch?v=I-EIVIHvHRM>
- [89] <https://www.christies.com/features/A-collaboration-between-two-artists-one-human-one-a-machine-9332-1.aspx>
- [90] <https://www.theguardian.com/commentisfree/2020/sep/08/robot-wrote-this-article-gpt-3>
- [91] More examples at: <https://aiartists.org/> or <https://www.artistinthemachine.net/about/>
- [92] Alexey A. Melnikov, Hendrik Poulsen Nautrup, Mario Krenn, Vedran Dunjko, Markus Tiersch, Anton Zeilinger, Hans J. Briegel. Active machine learning for quantum experiments. *Proceedings of the National Academy of Sciences*, 115 (6), 1221-1226, 2018: <https://www.pnas.org/content/115/6/1221>
- [93] Margaret A. Boden. Creativity and artificial intelligence. *Artificial Intelligence*, 103, 347-356, 1998 : <https://www.sciencedirect.com/science/article/pii/S0004370298000551>
- [94] Lu Liu, Yang Wang, Roberta Sinatra, C. Lee Giles, Chaoming Song, Dashun Wang. Hot streaks in artistic, cultural and scientific careers. *Nature*, 559, 396-399, 2018: <https://www.nature.com/articles/s41586-018-0315-8/>
- [95] Margaret A. Boden. Computer Models of Creativity. *AI Magazine*, 30(3), 23-34, 2009: <https://doi.org/10.1609/aimag.v30i3.2254>
- [96] Conversation with Alexander Peterhaensel.
- [97] Bob L.T. Sturm, Maria Iglesias, Oded Ben-Tal, Marius Miron, Emilia Gómez. Artificial Intelligence and Music: Open questions of copyright law and engineering praxis. *Arts*, 8(3), 115, 2019: <https://www.mdpi.com/2076-0752/8/3/115>
- [98] <https://medium.com/goodai-news/ai-and-work-a-paradigm-shift-7b314268bf05>
- [99] European Commission. *Digital Education Action Plan (2018-2020)*, 2018: <https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=COM%3A2018%3A22%3AFIN>
- [100] European Commission. *Digital Education Action Plan (update 2021-2027)*, 2020: [https://ec.europa.eu/education/education-in-the-eu/digital-education-action-plan\\_en](https://ec.europa.eu/education/education-in-the-eu/digital-education-action-plan_en)
- [101] Roy Saurabh, Nandini Singh, Anantha Duraipappah. The Alternate Education for the 21st Century. *2019 IEEE 18th International Conference on Cognitive Informatics & Cognitive Computing (ICCI\*CC)*, 265-271, 2019: <https://ieeexplore.ieee.org/document/9146061>
- [102] <https://www.aalto.fi/en/our-strategy/radical-creativity>
- [103] <https://www.teachermagazine.com.au/articles/what-is-finlands-phenomenon-based-learning-approach>
- [104] Cortland J. Dahl, Christine D. Wilson-Mendenhall, Richard J. Davidson. The plasticity of well-being: A training-based framework for the cultivation of human flourishing. *Proceedings of the National Academy of Sciences*, 117(51), 32197-32206, 2020: <https://www.pnas.org/content/117/51/32197>
- [105] <https://centerhealthyminds.org/science/overview>





## Annex: Questionnaire

### Introductory remarks

- The answers to the present questionnaire will feed into a short report on the potential impacts of the digital transformation, and in particular of artificial intelligence, on creativity, culture, education and also more broadly on individuals and societies. The main aim of this report destined to policymakers in the European Commission will be to stimulate a discussion around a number of emerging issues.
- While most of the ongoing discussions on the digital transformation and on artificial intelligence are focussing on the short and medium-term effects, the aim of the report will be to look also into the longer-term and to capture some of the key questions to which we will be confronted and for which a large public debate and a policy response will most certainly be needed.
- This longer-term approach allows us to go beyond the state of the art in technology and beyond the developments that are expected for the two or three years to come. The questionnaire therefore explores also developments which may still sound unrealistic at this stage and which may possibly remain unreachable, but for which some weak signals are emerging and which should therefore already be on the radar of policymakers and part of their reflection.
- AI is already playing a central role in the digital transformation, and this role will without any doubt increase in the years to come. This is why the main focus of the questionnaire is on AI. However, some questions consider digital tools more generally, including social media. Furthermore, some of the questions address very broad social transformations for which AI is only one potential factors among many others, including also other scientific developments such as bio-engineering for example.
- The European Commission has initiated important work in relation to AI over the last few years and some of the issues raised by the questions below were already touched on in the White Paper on AI published last February or in the Ethics Guidelines for Trustworthy AI prepared by the High-Level Expert Group. The angle is however slightly different: these documents cover extensively the challenges that will need to be addressed by the Commission as a matter of priority in the months and in the years to come, while the present questionnaire tries to look already a little bit ahead and therefore to be more exploratory.

- For this questionnaire, AI is considered in a broad sense in line with the definition used in the JRC Flagship report on AI <https://ec.europa.eu/jrc/en/publication/artificial-intelligence-european-perspective> ("AI is a generic term that refers to any machine or algorithm that is capable of observing its environment, learning, and based on the knowledge and experience gained, taking intelligent action or proposing decisions." ). Many different technologies fall under this definition, but it implies some degree of autonomy.
  - The questionnaire is structured around three sections and encompasses a rather long list of questions. You are not expected to answer all of them, but only to select and answer the few that seem most relevant to you. The full list of questions is meant to frame the topic and the scope of the entire reflection.
  - If rather than answering the questions, you prefer to share your thoughts, knowledge and opinion on these issues in a more open and free form, you are most welcome.
  - Please send your contributions through e-mail before the 31<sup>st</sup> of October to [vladimir.sucha@ec.europa.eu](mailto:vladimir.sucha@ec.europa.eu) and [jean-philippe.gammel@ec.europa.eu](mailto:jean-philippe.gammel@ec.europa.eu)
  - Whenever possible and relevant, we would be grateful for references to useful books, articles, research projects or sources of data.
  - The questionnaire will not be used for any quantitative purpose, but only for a qualitative analysis. Therefore, please feel free to answer in a very unrestrained and open way.
  - The questionnaire will be sent to a limited number of experts, not only from the digital and AI field, but also from other scientific disciplines and from the cultural and creative sector.
-

## **I. The digital transformation, cognitive and socio-emotional capacities**

### ***Human intelligence and artificial intelligence***

1. What can AI teach us about human intelligence?
2. To what extent can AI already decrypt and understand human sensations, emotions and sentiments?
3. To what extent can AI already use, influence, or transform human sensations, emotions and sentiments?
4. How similar to human intelligence can AI become? Are there any limits for AI?
5. Will AI for example be able to develop some kind of consciousness?
6. Will AI be able to develop some kind of sensations, emotions and sentiments?
7. What new doors would the acquisition of consciousness, sensations or emotions open for AI? What could be the implications?

### ***Impact of digital tools on human intelligence***

8. How does the interaction with digital tools affect human cognitive and socio-emotional capacities? Does it augment and enhance them? Does it change them? Does it diminish them?
9. What are the biological transformations taking place in our brains because of the interaction with digital tools?
10. How do digital tools affect our decision-making, our biases, our behaviours?
11. How do digital tools impact our mental health?
12. What is the impact of digital tools on loneliness or the perception of loneliness?
13. In all these areas, are children impacted differently than adults?

## II. The digital transformation, AI and creativity

### *New forms of events*

14. Has the Covid crisis accelerated the use of digital tools in the cultural sector?
15. Have some particularly innovative ways to use digital tools emerged?
16. Have innovative new formats of events emerged (including for example hybrid forms of live/online events)?
17. Have the interactions between artists and the public increased at the various stages of the creative process thanks to digital tools?
18. Is the division between cultural content producers and cultural consumers starting to blur?

### *Artists and AI*

19. Does AI boost the creativity of artists and open new perspectives for them?
20. How are artists using data and AI to produce new forms of art?
21. How do artists make sense of the AI revolution that is underway? What new narratives are they producing about AI? Are they starting to “resist” to some of the risks that AI poses? How?

### *Understanding creativity*

22. Can data and AI give us a new and deeper insight into how creative processes work in different artistic areas?
23. More broadly, can they help us to better understand how creativity works?
24. Could this boost the creative and innovative capacity of our societies?

### *Machines as artists*

25. Thanks to AI, machines are also starting to produce works of art by themselves without (or with very limited) human intervention. How creative can machines become?
26. What makes the difference between art produced by humans and art produced by machines? Can machines be “artists”?
27. What would be lost by replacing human intelligence with AI?

### III. The digital transformation, AI and society

#### ***Democracy***

- 28. How will AI affect our identities, our values, our ways of living together, our cultures?
- 29. Is democracy at risk? How can we prevent humans from being nudged, manipulated or used with the help of AI or even by AI?
- 30. Have the deluge of data and algorithms made traditional politics based on political parties and elections obsolete? What new forms could democracy take?
- 31. If AI understands or will soon understand humans and their needs better than humans do themselves and if it is better at decision-making, could we let it decide for us at least in selected areas? What would be the main risks?

#### ***Privacy***

- 32. How far are our societies ready to go in terms of loss of privacy and of transparency?
- 33. Can we still avoid full transparency or will it progressively become the new normal?
- 34. How far are our societies ready to go in terms of surveillance by AI to increase security?

#### ***Individualism***

- 35. How far are our societies ready to go in terms of predictability? What could be the impact of AI on the free will and the autonomy of individuals?
- 36. Is the community more important than the individuals are? Should common good always prevail over individual choices?
- 37. Is AI and in particular machine learning “conservative” because it predicts the future and guides our choices only based on the patterns of the present and of the past? Is there a risk that in the longer term AI could freeze our societies and slow down or even stop evolution?

#### ***Uniformisation***

- 38. What are the main impacts of the recommendation algorithms that “guide” our choices? Do they prevent us from making new and original experiences and from broadening the range of our interests? Are they putting us in “isolation bubbles”? Will they accelerate uniformisation and polarisation?
- 39. More specifically, do they put cultural diversity at risk? Or do they also have positive aspects for artists and consumers of cultural goods?



40. How could we counterbalance the potential negative effects of these algorithms on cultural diversity?
- Would it be sufficient to raise the awareness of cultural consumers? How?
  - Are there other ways to empower consumers?
  - Some researchers and companies have started to introduce “diversification factors” or “serendipity components” in their algorithms. Should this be encouraged? How?
  - Should “public algorithms” be developed and made available to compensate the potential negative effects of the commercial algorithms provided by platforms? How? By whom?
41. Some of the key factors that were at the heart of all “avant-gardes” and that have always allowed humanity to move forward are chance, mistakes, imperfection, difference, outliers, transgression... How can we make sure that humanity will be able to preserve these factors so it can continue to move forward? Could AI become a new “avant-garde”? How?
42. Will there still be a place in the age of AI for “atypical human intelligences” and non-conformists?

### **Well-being**

43. How do digital tools affect our well-being?
44. How could we compensate the potential negative impacts, especially on young people?
45. Is the capacity for well-being something that can be taught in schools? How?
46. Will our conception of well-being evolve with the new opportunities that AI will bring?
47. Could well-being be replaced by some kind of virtual state created by AI in which our sensations and emotions would be controlled, suffering would be diminished, our basic desires full-filled and our uncomfortable desires erased?

### **Immortality**

48. Could “physical immortality” or at least a significant increase of life expectancy become possible? What would that imply for humanity?
49. If physical immortality cannot be reached, some programmes are already aiming at “digital immortality” by uploading the consciousness of human beings on computers. Could this become possible? By when? What would be the consequences?
50. And even if “digital immortality” remains very far away, humans are already leaving huge amounts of digital traces (pictures, texts or other forms of data) which will remain in the digital space after their deaths. Who should inherit the rights for these digital traces? Shouldn’t this be regulated?



### **Finding information about the EU**

#### Online

Information about the European Union in all the official languages of the EU is available on the Europa website at: [https://europa.eu/european-union/index\\_en](https://europa.eu/european-union/index_en)

#### EU publications

You can download or order free and priced EU publications at:  
<https://publications.europa.eu/en/publications>. Multiple copies of free publications may be obtained by contacting Europe Direct or your local information centre (see [https://europa.eu/european-union/contact\\_en](https://europa.eu/european-union/contact_en)).



■ Publications Office  
of the European Union