



D12.3.1: Insight into Europeans' informed and considered views on Artificial Intelligence - Results from EuropeSay on AI

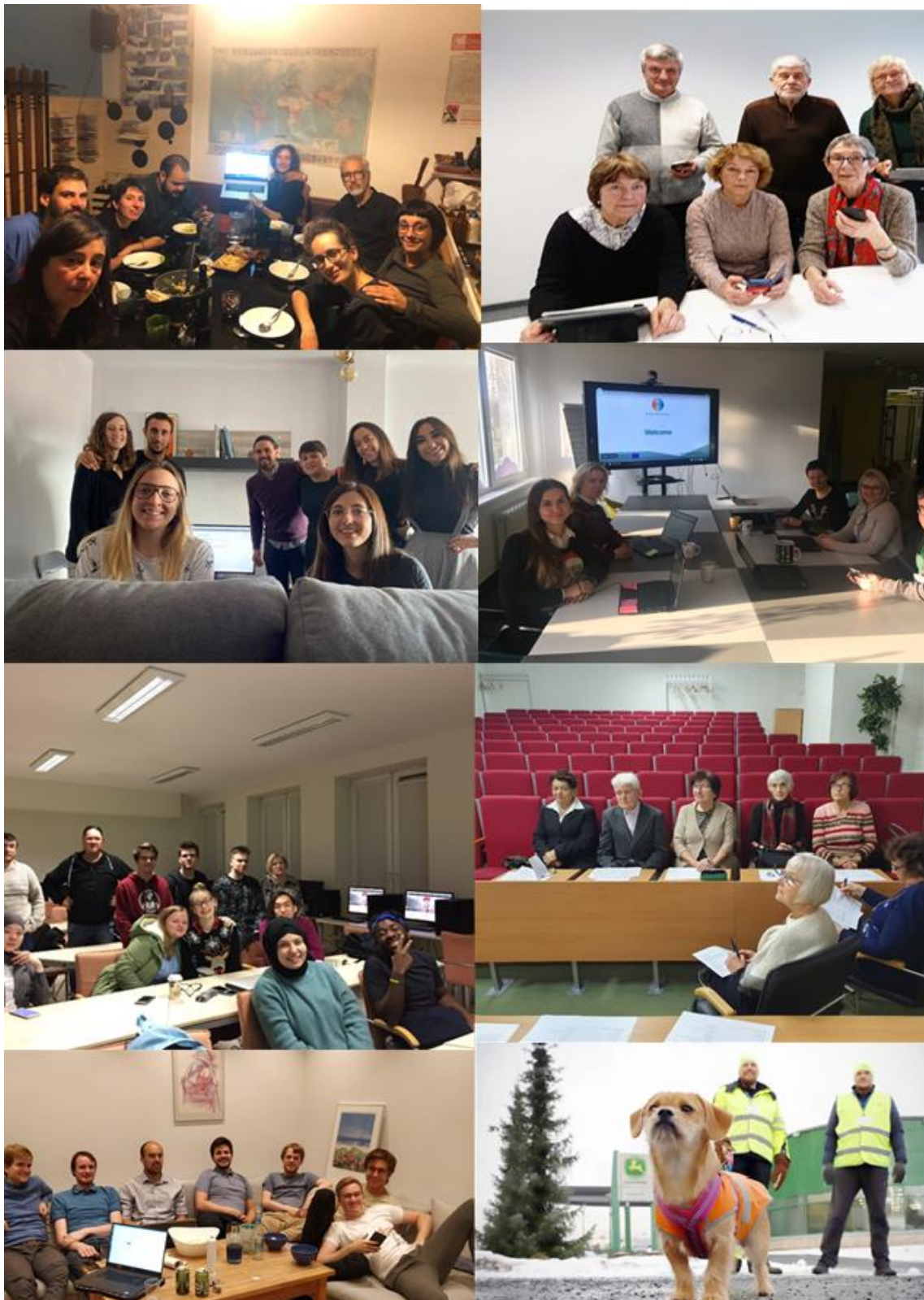


Figure 1: Pictures taken by participants and uploaded as part of the consultation

Project Number:	785907	Project Title:	Human Brain Project SGA2
Document Title:	D12.3.1: <i>Insight into Europeans' informed and considered views on Artificial Intelligence - Results from EuropeSay on AI</i>		
Document Filename:	HBP_SGA2_Basic_Deliverable_190208.docx <i>(to be updated by PCO)</i>		
Deliverable Number:	SGA2 D12.3.1 (D77.1, D113)		
Deliverable Type:	Report		
Work Packages:	12.3.1		
Key Result(s):	KR12.3		
Dissemination Level:	PU		
Planned Delivery Date:	SGA2 M23 / 28 FEB 2020		
Actual Delivery Date:	SGAx M? / DD MMM YYYY <i>(to be updated by PCO)</i>		
Author(s):	Lise BITSCH, Nicklas Bang BÅDUM, Nathan CAMPION & Marie Louise JØRGENSEN (FT)		
Compiled by:	Nicklas Bang BÅDUM & Nathan CAMPION (FT)		
Contributor(s):	Lise BITSCH, Marie Louise JØRGENSEN (FT)		
SciTechCoord Review:	firstname FAMILYNAME, SHORT PARTNER NAME (P#) <i>(to be updated by PCO)</i>		
Editorial Review:	firstname FAMILYNAME, SHORT PARTNER NAME (P#) <i>(to be updated by PCO)</i>		
Description in GA:	Report on citizen engagement activities across Europe, including recommendations from citizens to the Human Brain Project.		
Abstract:	<p><i>This report presents the results of the first European-wide citizen engagement on artificial intelligence, EuropeSay on AI, which ran from September 2019 to January 2020. The consultation focused on getting insights into what European citizens think about artificial intelligence, its potential benefits and risks, as well as some of the applications that it can be put to.</i></p> <p><i>The consultation found that participants were generally positive to AI, and applications in medical and health research were generally seen as acceptable, whereas support for use by public authorities was more contingent. However there were also widespread concerns, e.g. about its potential application for political purposes and for making accurate inferences. In continuation of this, there was also concerns about the lack of control over personal data. In addition, it was found that black-boxing of AI was not acceptable, and that AI systems should be explainable and transparent. Further, the consultation found that the vast majority of participants believed that AI should be subject to regulation, and means to consumer empowerment should be implemented.</i></p>		
Keywords:	<i>Artificial intelligence, AI, Citizen Engagement, Public Engagement, Responsible Research and Innovation, medical research, health research, Regulation of AI</i>		
Target Users/Readers:	Policymakers, researchers, scientific community, AI researchers and developers, funders, students, computer scientists.		

Table of Contents

1. Introduction	4
1.1 Why focus on Artificial Intelligence?	4
1.2 The relevance of the citizen perspective	5
1.3 Focus of EuropeSay on AI	5
2. What do Europeans think of AI's possible applications?	6
2.1 General Support for AI, but...	6
2.2 No Black Boxing of AI	8
2.3 Using AI Services - Transparency and Control	9
2.3.1 Call for Transparency & Control	11
2.4 Citizens perspectives on use of AI in health and medical research	12
2.5 Public Authorities' use of AI	14
2.6 Create Trustworthy AI and Trust in AI Through Regulation and Consumer Empowerment	15
2.7 Conclusions	17
3. Methodology	19
3.1 EuropeSay	19
3.2 AI 360 COPENHAGEN	21
4. Annex 1: References	22
5. Bibliografi	22
6. Annex 2: Participants Demographic Data	23

Table of Tables

<i>Table 1 – Which of the following information about you, would it concern you that an artificial intelligence can deduce based on seemingly unrelated data??</i>	Fejl! Bogmærke er ikke defineret.
<i>Table 2 – Do you think AI predictions should be</i>	15
<i>Table 3 – What is the best way to ensure that AI</i>	17
<i>Table 4 – Number of meetings and participants by country</i>	Fejl! Bogmærke er ikke defineret.

Table of Figures

<i>Figure 1: Pictures taken by participants and uploaded as part of the consultation</i>	1
<i>Figure 2 – Do you already have an opinion about Artificial Intelligence?</i>	6
<i>Figure 3- What is your main concern regarding AI?</i>	7
<i>Figure 4 - Do you trust decisions based on AI?</i>	7
<i>Figure 5 - Do you do any of the following?</i>	10
<i>Figure 6 - Do you feel that you have control of the data available on you?</i>	11
<i>Figure 7 - If yourself were offered the chance of a mental health screening, which could identify signatures of various mental and psychological disease, under what conditions would you agree to it?</i>	13
<i>Figure 8 - Should there be some kind of deanonymizing mechanism so that it is possible to connect the data with the person behind them if necessary?</i>	13
<i>Figure 9 - Do you think public authorities (social workers, public administrators and similar) should be allowed to use AI and predictive profiling and cross-share information to identify people in risk of</i>	14
<i>Figure 10 - Do you think algorithms should be regulated, and, if so, how?</i>	16

1. Introduction

This report presents the results of the European wide citizen consultation on artificial intelligence (AI), EuropeSay on AI. Recent years has seen significant progress in research and development of AI, and thus also proliferation of applications across varied sectors. For this reason, it has also stirred considerable debate among policymakers, researchers and in mainstream media.

EuropeSay on AI was the first major European-wide citizen engagement on the topic of AI. It ran from September 2019 to January 2020 across 13 European countries, and it engaged 928 participants across 156 small-scale meetings. The consultation sought to get an understanding of what opinions and values there are among Europeans concerning AI, what uses they condone and which ones they are critical of or outright against, and whether AI should be regulated and how. The guiding question in the consultation was “What uses can we as a society justify to each other, and how do foster these, while hampering unwanted uses?”.

The consultation made use of a novel methodology, GlobalSay, which produces informed and considered quantitative outputs. Participants set up their own meeting with friends or family, where they are guided by an online platform which presents them with information videos, discussion questions and collects individual responses. An unlimited number of meetings can take place anytime and anywhere over a set time period.

The consultation generates a well-grounded picture of the quantitative tendencies in the European population, both at the national and transnational level¹. In engaging citizens in deliberations about the societal and ethical impact of AI, HBP aims to democratically qualify the debate by including the concerns and opinions of the wider public and thus supplement the viewpoints of those who are already well-represented, such as researchers, policy-makers, engineers, businesses and interest organisations. The output will be used to broaden the knowledge base for policy making at both EU and national level with knowledge.

EuropeSay on AI was organized and facilitated by the Danish Board of Technology (DBT) as part of the Human Brain Project’s public engagement activities. The consultation builds on the results of the multidisciplinary expert workshop, AI360 I COPENHAGEN, which was likewise organized by the DBT. Both activities are part of a wider work on the societal and ethical implications of AI, being conducted in the Ethics & Society subproject, and which will a concerted opinion on AI, *Trust and Transparency in Artificial Intelligence* (Bitsch E. R., 2020).

1.1 Why focus on Artificial Intelligence?

AI has in the past few years risen to a prominent position on the agenda of policymakers, researchers, NGOs, large corporations and start-ups. It is variously hailed as the next technological revolution, and as the end of mankind. Working groups, expert committees, conferences, research and development projects and strategies are popping up left, right and centre. In the EU the development is receiving attention much attention with the establishment of the AI High-Level Working Group and the recent publication of the Commission *White Paper On Artificial Intelligence - A European approach to excellence and trust* (European Commission, 2020).

All of this is testament to the fact that AI holds vast potentials as an enabling technology, i.e. “Equipment and/or methodology that, alone or in combination with associated technologies, provides the means to generate giant leaps in performance and capabilities of the user” (BusinessDictionary, 2020). What makes AI particularly prominent, is its almost universal potential for application. It, or predecessors of it, is already being applied in as diverse sectors as law, medical research and diagnostics, engineering, building and equipment maintenance, transport, finance, insurance, public service provision, and it holds vast potential for further application in all of these

¹ The methodology does not provide statistically representative output, but rather well-grounded insights into the an informed and considered public opinion about AI.

and many more fields. The hope is that AI can increase efficiency, improve services, create growth and release resources that can be redistributed.

1.2 The relevance of the citizen perspective

While the vast positive potentials that AI promises are hard to dispute, there has also been an increased realization of the potential harms and adverse consequences that could be caused by AI, intentional or not, as well as the widespread changes in social, economic, legal and other conditions that will come from increased societal uptake of AI. Thus, recent years has seen an increased focus on an AI development which harnesses the positive potentials, while curtailing the negative consequences. This has also been driven by the risk of tech-lash, i.e. the risk that the popular opinion will turn so decidedly against a technology, that development and uptake will be curtailed, and, in extreme cases, will have to be shelved, at least temporarily, despite its promising potentials. This was, for instance, the case with GMO in Europe in the 1990s. The most prominent response to this risk has been to call for creating trust in AI, by fostering trustworthy AI (European Commission, 2020) (EC High-Level Expert Group on Artificial Intelligence, 2019)

The need to consult with citizens about these questions rests on a fundamental democratic tenet. Something which will profoundly affect the lives of all citizens should not only be discussed by experts, stakeholders and policymakers, but needs to have a broad public debate. A foundational aspect of democratic governance is that the citizens should be involved in deciding what technology can be considered to be societally beneficial and how these technologies should be applied. Further, if trust is to be created, it is necessary to know what the actual concerns are, what uses are seen as desirable and acceptable, and where the line should be drawn for what constitutes unethical, nefarious or simply unacceptable use. Having this dialogue is essential to ensure that the full potentials of the technology can be wielded for beneficial purposes while at the same time addressing and curtailing the negative potentials inherent in the technology. So much the more so, because not having this discussion could lead to tech-lash against AI, which could ultimately lead to a severe curtailing of the positive potentials of the technology.

Furthermore, it is urgently important to have these discussions now, while the society wide application is still in its infancy and the paths that it develops along can still be defined. Technological development and application do not develop deterministically, and now is the chance to steer the development in a societally beneficial and desirable direction.

Apprehension about asking citizens has often been based on a perception of citizens as having insufficient knowledge about AI or simply not being competent enough to provide useful output. But while not all citizens have technical expertise, this is only one point of view. Citizens have in-depth knowledge about how all parts of society functions from their professional and private lives. Thus, the combined knowledge of citizens is vast and in no means inferior to that of experts. It is important to stress that the engineer, the social scientist, and the doctor are all citizens. In addition, in contemporary societies citizens are historically well educated and, not just capable, but also willing to be engaged. There is plenty evidence that citizens are very capable of engaging in discussions on complex topics in a nuanced and enthusiastic way.

1.3 Focus of EuropeSay on AI

Since EuropeSay on AI was the first citizen consultation on AI to span Europe, the aim was to get a general view of what the participants thought of AI. In addition, the consultation was carried out under the auspice of the Human Brain Project, so extra attention was paid to potential application within medical and health research.

The consultation was divided into four thematic segments with each their focus, but also with cross-cutting aspects. The first section focused on participants' general knowledge and attitude to AI and its applications. The second and third focused on various potential applications and ways of applying

AI, while the fourth and last focused on whether and to what extent AI should be controlled and regulated.

As some of the focus areas of the consultation cut across the sections, the analysis will not stringently follow the sequence of the consultation.

2. What do Europeans think of AI’s possible applications?

2.1 General Support for AI, but...

One of the stated purposes of the consultation was to get an insight into what Europeans think about AI and given that EuropeSay on AI is the first European-wide consultation, it was a unique opportunity of getting an understanding of what people think about AI in general. Thus, the initial part of the consultation aimed at gauging the participants’ general understanding and opinion of AI and to investigate what general concerns they have about it.

First off, the participants were asked to state the extent to which they were familiar with AI prior to the consultation, and it was clear that AI as a topic has reached the mainstream. Only 5% knew nothing, while 51% knew some and 14% a lot about AI, while 28% knew a little. Following this, participants were asked if they had an opinion about AI prior to taking part in EuropeSay on AI. Of the participants that had an opinion, the vast majority were positively inclined. However, it is also important to note that 40% either did not have an opinion (22%) or did not know (18%). It is also worth noting that not everyone was keen on AI, as 18% had a negative opinion of AI. Interestingly, it appeared that, the more participants reported to know about AI, the more they tended to be positive towards it. This indicates that information and education of the wider population about AI could be important means of creating and maintaining public support for research, development and application of AI. As mentioned above, and as is clear from figure 1, there was a very large proportion of participants who did not yet have an opinion. If this holds true, it means that there is great potential for positively influence the popular opinion of AI, while it is not yet decided, which is a considerably easier task than turning around a negative opinion. It also means that an important task moving forward will be to provide the general population with reliable and balanced information, and to continue consulting the population.

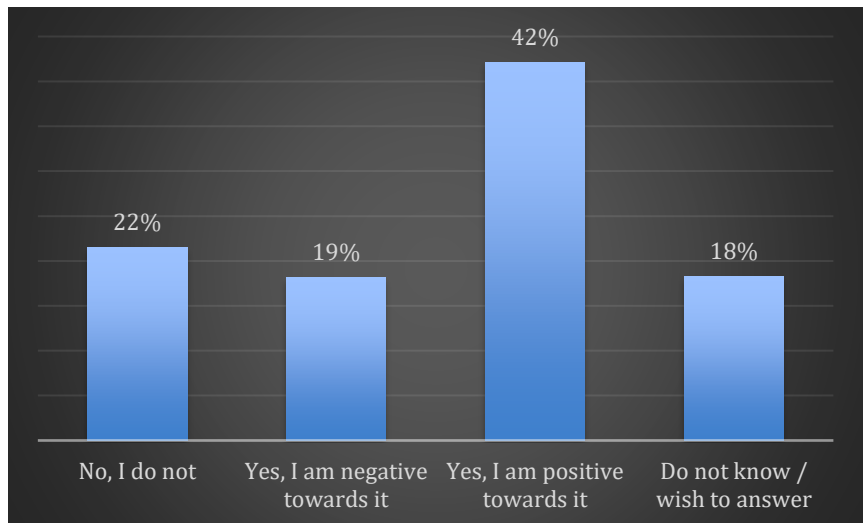


Figure 2 - Do you already have an opinion about Artificial Intelligence?

This was made further clear from the question asking participants directly what their main concerns were. As can be seen in **Figure 3 below**, only 1% did not have any concerns, whereas the concern chosen by most participants was the risk that AI technology potentially being used for social manipulation which was chosen by 51% of the participants. Of the following three most picked areas of concerns we find possible data breach at 42%, sale of personal data at 37% and loss of privacy at 31%. It is interesting to note that the option chosen the least by participants was that they were concerned that we may not know the full reasoning behind AI analysis. However, this should most likely be seen in light

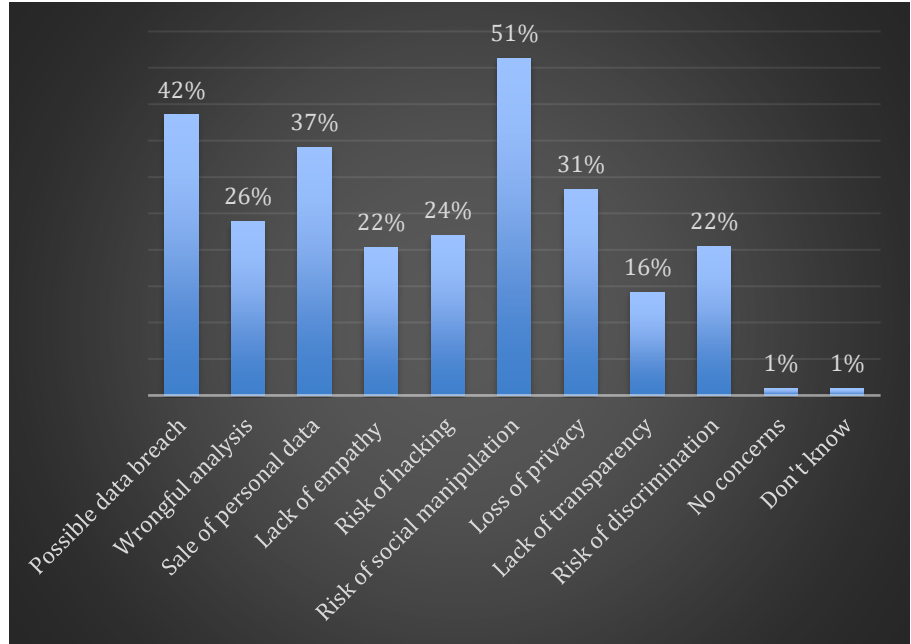


Figure 3- What is your main concern regarding AI?

of the fact that a range of questions prior to this had focused on variations of transparency and explainability, and to these questions, these two aspects were generally considered important, as will be shown below. So it is evident that, while participants do express support for the use of AI, it is not a careless support.

It is also clear that artificial intelligence was not a be-all and end-all solution to the participants. When asked if they thought all decisions can be based on AI, 52% disagreed while 37% agreed, and when asked later in the consultation whether they trust AI based decisions on social matters, this was further supported. While 54% trusted AI, the majority of them (34% of the total) trust it to some degree but are concerned about the lack of human judgement. 42% did not trust AI based decisions, either because the process and reasoning is not clear (13%) or because they do not believe machines can understand the social complexity (29%). Therefore, while there is support for using AI, the

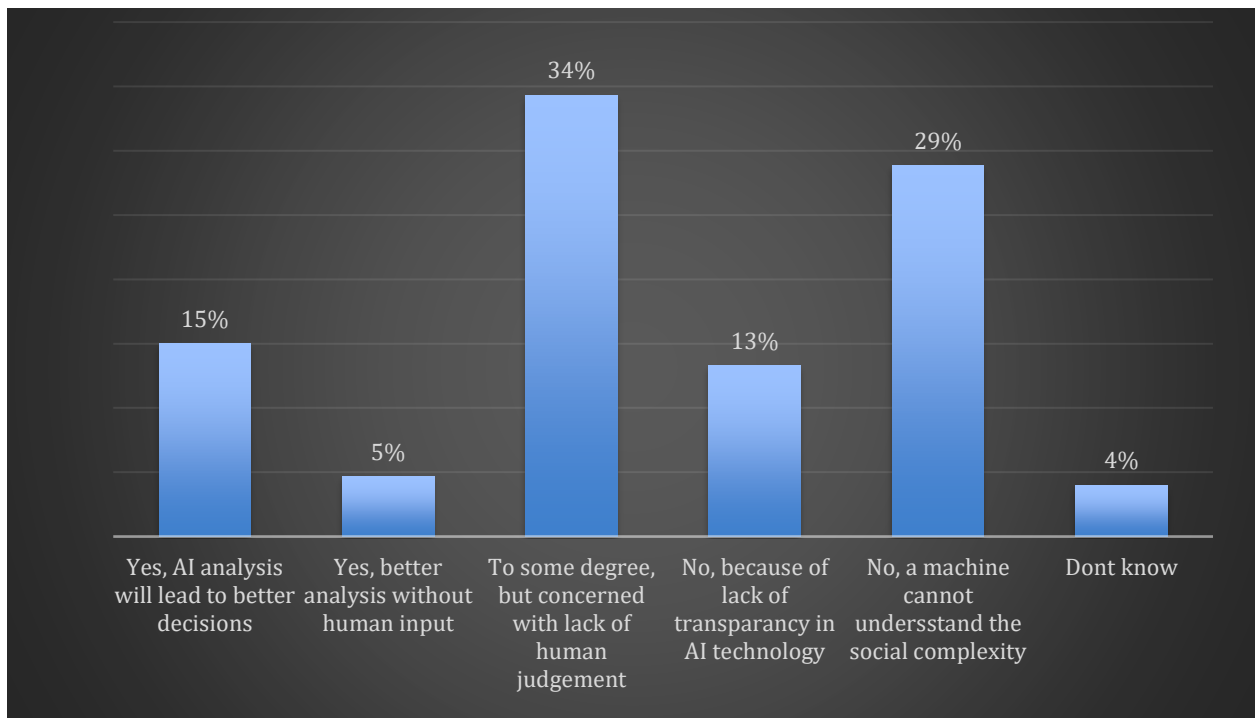


Figure 4 - Do you trust decisions based on AI?

participants did not see it as something which could be a stand-alone decision maker. It indicates, rather, that they saw it as a tool that can be integrated in a decision-making process as a way of improving the foundation for making the decision, but it is still humans making the final call. This is indicative of the participants not accepting that AI systems get to function as black boxes. This will be dealt with further in the next section.

This initial analysis shows that while the participants of EuropeSay on AI were generally positively inclined to AI and the use hereof, they also had significant concerns about its potential applications and abilities.

2.2 No Black Boxing of AI

A prominent discussion surrounding AI is the issue of black-boxing (e.g. (EC High-Level Expert Group on Artificial Intelligence, 2019) (Blanco-Justicia, to appear)), i.e. that some AI systems work in a way where deciphering how the system works, how it weighs different data points and reaches its final conclusion or suggestion, is practically not possible. An important question is, if a decision is made which profoundly affects someone's life, is it important to be able to explain why that decision is made? This goes for provision of public services like social allowances, for use in medical diagnostics or in self-driving cars. In each case, it is likely that AI systems can be highly accurate and effective. But AI can also make mistakes, and sometimes in unpredictable ways and for a variety of reasons. And sometimes black-box systems can run for a long time before anyone realizes that it has systematically been making mistakes. To uncover what participants thought of this issue, a range of questions looked into how important they thought transparency and explainability, was to them.

Tabel 1

Q 3.3	
This kind of analysis is no different from all others, and should be treated that way	3%
The artificial intelligence should always state what data is used in the analysis and how it is weighed	24%
The results of the analysis should always be controlled by a human	20%
There should be a complaints authority that can review the analysis, provide explanations and overrule the decision	11%
Decisions affecting peoples' lives should not be taken using artificial intelligence based analyses	20%
The artificial intelligence system should be open to independent review	18%
I do not know / wish to answer	5%

Thus, one of the questions that the participants were asked, was whether anything should be done to handle the difficulty of understanding analyses from AI systems and their foundations, which in turn could make it difficult to challenge decisions or detect mistakes. Of the three most picked answers, the most chosen answer, with 24% of the participants picking this option, was that that AI systems should always be transparent and state what data that has been used and how this data has been weighed in their analysis and conclusions. If it is impossible to understand the rationale of an AI's decisions, the risks would be too great, or it would be too difficult to correct mistakes in the algorithms. Being the favoured option of the participants shows

that transparency and explainability is very important to them, even if its barely a quarter of them who picked it. Another 20% of the participants was of the belief that the results of an AI analysis should always be controlled by a human, which indicates that they trust an informed human decision more than a decision solely based on an AI and that AI's should **only advise and inform**. Another fifth of the participants believed that decisions affecting peoples' lives should not be taken using AI based analyses, which points towards distrust in the decision-making capabilities of the technology.

Only 3% thought that this kind of analyses does not differ from any other analysis and should be treated accordingly. In other words, AI is not just the run of the mill analysis, and close to half of

the participants think that it should either be transparent and explainable or only be used as a decision support. At any rate, there is not support for using AI indiscriminately.

Another question which gives us an insight into the participants views on black boxing (**Fejl! Henvisningskilde ikke fundet.**), is the question whether they trusted decisions made based on AI. While 20% answered yes (15% because it would lead to better decisions and 5% because analyses would be better without human judgement involved), most of the participants picked either a “neutral” or no option. A third of the participants, at 34%, stated that they trusted the decision to some extent but were concerned with a lack of human involvement and judgement. The majority of the participants, with 42%, picked an ‘no’ answer, either saying that they didn’t trust AI analysis because the process and reasoning aren’t clear, 13%, or saying that a machine can’t understand the social complexity of human society, 29%.

Based on this, it seems fair to say that there is definitely demand for AI systems that are explainable and transparent. In addition, it seems to strengthen the point made previously, that AI should not be stand-alone decision-makers, but rather function as decision support elements in a carefully designed decision making process, which is also supported by the fact that 26% of participants had wrongful analysis as one of their main concerns about AI.

When asked about what their main concern on AI was, of the answers relevant to this section, 26% of the participants said that wrongful analysis 26% was a main concern and another 16% said that their main concern was that we may not fully understand the reasoning behind AI analysis. These numbers show that, although they are not outliers, that the citizens truly are concerned with the potential issues that come with AI analysis.

However, as shall be demonstrated below, it was also clear that to the participants, it was not sufficient that AI systems are transparent and explainable to those operating them and utilizing the analyses; 81% of the participating also considered it to be either very or somewhat important that the rationale behind an AI’s decision is obvious and easily understood. This will be elaborated more in the next section.

2.3 Using AI Services - Transparency and Control

AI can be used to provide a host of very useful and convenient services, such as search engines, social media, route planning, suggestions for news-items, music suggestions or personalized advertisements, however, these services usually also conduct large scale collection of personal data across services and time. This data is, among other things, used to create very precise profiles of users, regarding anything from their preferred brand of ketchup or what holiday destination they can be influenced to choose, to more personal things such as their relationship status, economic and social standing, mental and physical health and much more. For users of these services it is often difficult, if not impossible, to know exactly what data is collected, by whom, what the data can tell about the data subject and how it used. Nonetheless most people make widespread use of these

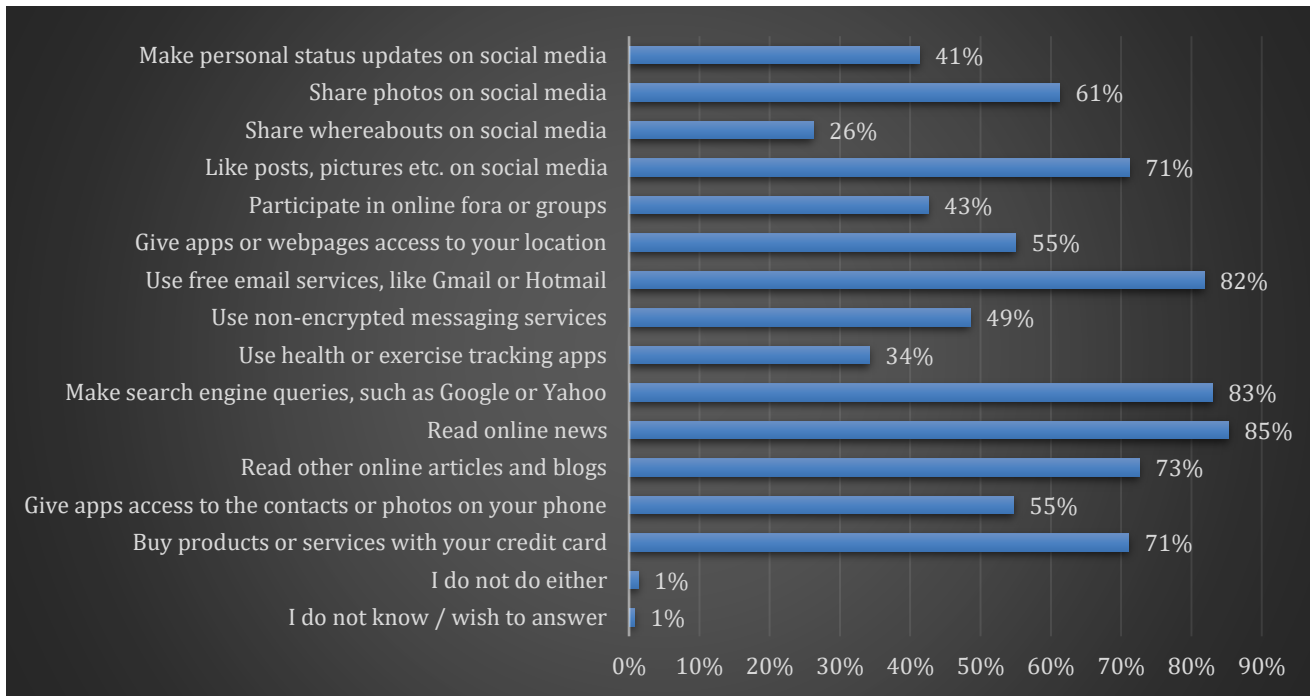


Figure 5 - Do you do any of the following?

services. For this reason, a number of questions focused on the participants’ digital behaviour and their opinion about the potential uses that their data can be put to.

Table 2 - Which of the following information about you, would it concern you that an artificial intelligence can deduce based on seemingly unrelated data?

Mental health	43%	Employment status	12%
Physical health	27%	Career ambitions	12%
Sexual orientation	10%	Mood	15%
Romantic liaisons	8%	Circle of friends and acquaintances	20%
Relationship status	14%	Economic standing	29%
Political observations	12%	None of this concerns me	10%
Philosophical beliefs	35%	I do not know / wish to answer	6%

When asked what services that they make use of, it was clear that almost all of the participants made widespread use of digital and online services. Only 1% reported not using any of the services listed. The participants were asked to select all the uses that applied, and the 904 participants who did report using services, made a total of 7648 selections. This means that each of them on average

reported making use of 8,5 services, which indicates that use of these services are deeply ingrained throughout society. The vast majority read online news, search engines, free email services, make use of social media, make online purchases and read other online content. Most of the other suggest applications also had widespread usage, and even the least used services, openly sharing whereabouts on social media and use of health or exercise trackers, were used by more than a quarter of all participants.

However, it was also clear that the participants were concerned about what the data about their usage of these services would enable AI systems to tell about them. As displayed in Table 1 the participants were asked what information, about themselves, that they would be concerned about an AI being able to determine based on seemingly unrelated data. Only 10% stated that none of the suggested examples concerned them, while the most frequently chosen option was that it could deduce their mental health (43%). Second to that was philosophical beliefs (34%), followed by economic standing (29%) and physical health (27%). It is interesting to note that very few participants reported being concerned about AI being able to work out the romantic liaisons, their sexual orientation or political observations. So, it appears that while the participants make widespread usage of online services, they are at the same time concerned about what information about this usage can be used for.

One of the potential uses of these this data is microtargeting, in which this data is used to make accurate profiles on individuals. These profiles are variously used for commercial and political purposes, where content is tailored to the individual profile and their perceived susceptibility. Using this kind of profiling and micro-targeting was not condoned by the participants. They were asked if it concerned them that microtargeting can be used for political purposes. 75% answered “yes, it should be made illegal”, while 9% did not mind either way. 6% answered “No, I see no problem with this.” This clear signal should also be seen in the light that the most chosen concern regarding AI was that it can be used for social manipulation, as shown above, and it could be an indication that AI based profiling for political use is not seen as a widely condoned potential. It is interesting, though, that when they were subsequently asked if it would make them want to change the way they find information about political issues, knowing that commercial companies register what they look at, the answers were less clear cut. While 48% answered yes, 38% answered no, while 14% did not know or want to answer. So while there is clear concern about profiling being used for political purposes, the majority willing to change how they find information about political issues was substantially smaller.

2.3.1 Call for Transparency & Control

Still, from subsequent questions it was confirmed that this was something participants thought about. Participants were asked how much they think about what the data from their online behaviour can tell about them and what it is used for. 32% reported often thinking about the data they leave behind and refraining from using certain products and services, while 15% reported that they think about it a lot and that they are doing what they can to minimize the data they give away. On the other hand, there were also quite a few who thought less about it and took no action. 16% think about it but do not act, while 19% think about it sometimes without acting. Only 7% do not think about it at all. So, while it is not top of mind for all participants, almost all of them think about this, and more than half take steps to curtail collection of their personal data.

But even if it was something that most participants thought about, and almost half of them take steps to limit their digital footprints it was also clear that they did not feel that they had control over their digital identity. When asked, 69% of the participants responded that they have either no or only very little control of their data. That control of data appears to be an important issue, is further underlined by the results reported previously, that breach of and sale of personal data are the second (42%) and third (37%) most frequently chosen concerns that participants had about AI. That this is the case, is strengthened even further by the fact that, when asked if they think there should be restrictions on what kind of data companies can sell and to whom,

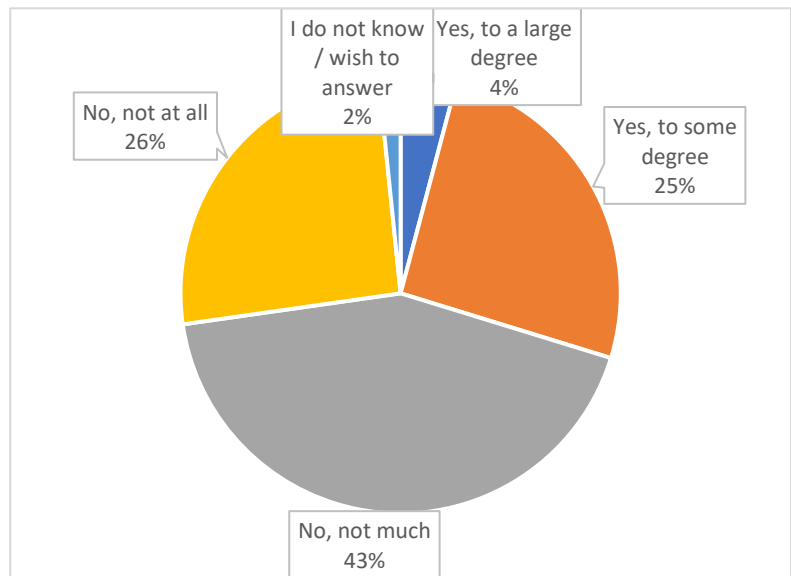


Figure 6 - Do you feel that you have control of the data available on you?

89% answered yes, while only 5% did not think so. Which speaks to a fundamental issue. Modern social relations are to a large extent contingent on online presence on social media and the like and increasing parts of modern work-life takes place online, just like being an informed citizen keeping up with international news and politics today means reading news online. This makes these services hard to circumvent for most people. So, the fact that participants still make uses of these services should not be seen as an accept of loss of privacy and control. Rather, there appears to be a conflict for most participants between wanting to use these services and at the same time protecting their privacy. There is little doubt that the participants are concerned about what their data footprints

can say about them, and at the same time, they do not feel that they have control over these footprints, even if they take measures to limit the data they give away. And this leaves the question, what should be done, when it is clear the people do value control and privacy, but at the same time wants to, and some need to, make use of services that they know will compromise the control they have of their own data and their privacy. One possible solution is to implement regulation. What the participants thought of this will be covered in section 2.6.

2.4 Citizens perspectives on use of AI in health and medical research

Parts of the consultation focused on used of AI in medical and health research and for diagnostics. This is because it is one of the application areas where the ethical dilemmas are clearest and with the most tangibly questions of balancing ethical considerations with potentials to save lives, but also where the consequences of mistakes are most directly a question of life and death. The questions generally focused on the use of AI systems to extrapolate future risks, of illnesses or otherwise and what to do when researchers identify individuals predisposed to various risks.

As briefly presented above, among the participants who can be said to have an opinion about AI, the majority were positively towards it. This also became clear from their support of a number of the applications that AI currently is used for and potentially could be used for. When it came to the use of AI technology in the sphere of medicine and health, the participants must be to be said generally positive. This can be seen in the results of several of the questions on the topic of health and medical

Tabel 2 - Do you think researchers should be obliged to react if they attain knowledge about people that are predisposed to or at risk of certain illnesses?

Yes, if the risk is potentially life-threatening	44,58%
Yes, if the risk can negatively change their physical living conditions	10,72%
Yes, if the risk can negatively change their mental living conditions	4,97%
Yes, if the risk can negatively affect their social skills	1,35%
No	26,75%
I do not know / wish to answer	11,63%

research, which will be reviewed below.

The question regarding if researchers should be obliged to react, if they discovered that an individual was predisposed to certain risks in the results of large dataset analysis. By far the majority of the participating citizens were positive in this regard. A combined 62% of them thought that researchers should react and alert them if something was discovered. Of these, 45% agreed to be contacted if the illness was potentially life-threatening to them, another 11% would agree if it was their physical health that was at risk, 5% would agree if their mental health was at risk and 1% would agree if it could put their social skills at risk. Of those opposing, 27% of the participants answered simply no, they thought

that researchers should under no circumstances contact them, if an illness was discovered. The use of predictive diagnostics is a contentious topic, but it appears that for the majority of participants it is considered to be acceptable. However, the 27% of participants who thought otherwise is a fairly large number. As has been discussed elsewhere/above/below in this project, the participants seem to be very concerned in regard to the subject of privacy.

In continuation of the above, it can be seen that in regard to health concerns, the participating citizens had a fairly positive attitude towards the use of AI technology for preventive profiling. Additionally, as can be seen in Figure 9 in the flowing section on page 15, that the approval rate for the use of AI for health-related subjects was 28% for mental health, 22% for physical health and 21% for life-style diseases. These are not numbers which much can be concluded on alone, though, in combination with the prior mentioned and the results presented in the graphs below regarding health

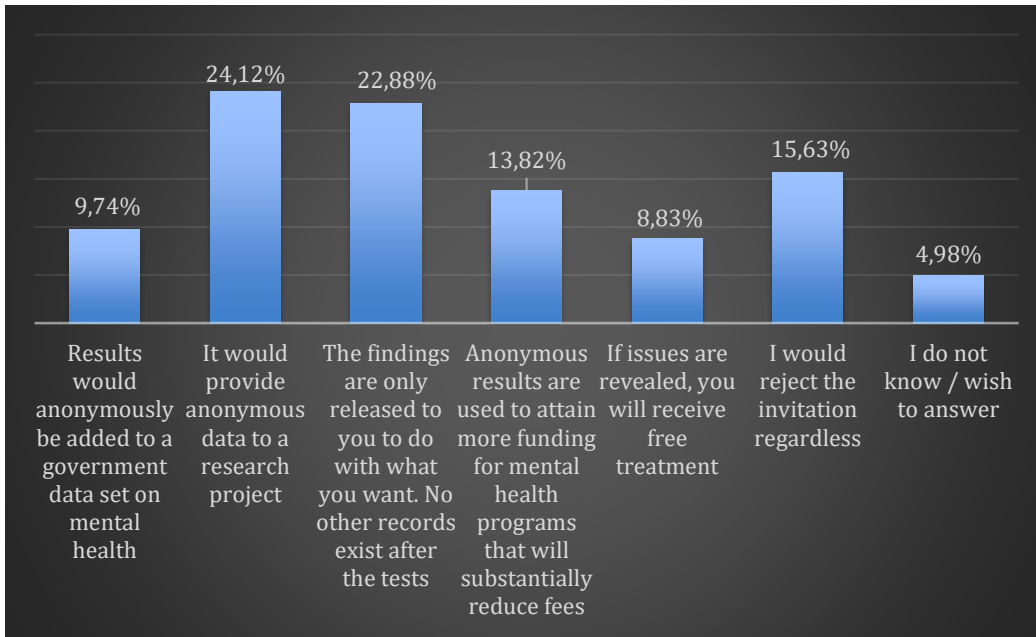


Figure 7 - If yourself were offered the chance of a mental health screening, which could identify signatures of various mental and psychological disease, under what conditions would you agree to it?

and medical research, it is possible to state that the participating citizens are positive towards the use of AI technology in health and medical research.

Another question concerning the views of the participants towards the use of AI in health and medical research, is when the participants were asked if they would agree to a

hypothetical mental health scan, which could identify signatures of various mental and psychological diseases. Of the participants, only 16% of them answered that they would reject the invitation, while the remaining 84% answered, under different conditions, that they would agree to a scan. The three most popular options were; if it would provide anonymous data to a research project at 24%, 23% would agree if the findings only were released to them and with 16% agreeing if they would receive free treatment, if something was revealed. This reinforces the above suggestion that participants support the use of AI for predictive profiling in medical and health research, and it appears there are wider allowances for use in these fields.

However, it should be noted that a desire for privacy regarding the participants information still appear to be something of importance for the participants with 27% declining contact from researchers in the one question and 16% stating that they would reject an invitation in the other.

When asked about whether there should be a deanonymization mechanism for health data, so it is possible to contact otherwise anonymous individuals if necessary, the participants' answers supported this. In the results it was clear, that if data was anonymized, 52% of the participating citizens thought that deanonymization was contingent on explicit consent from the data subject, for the data to be deanonymized under

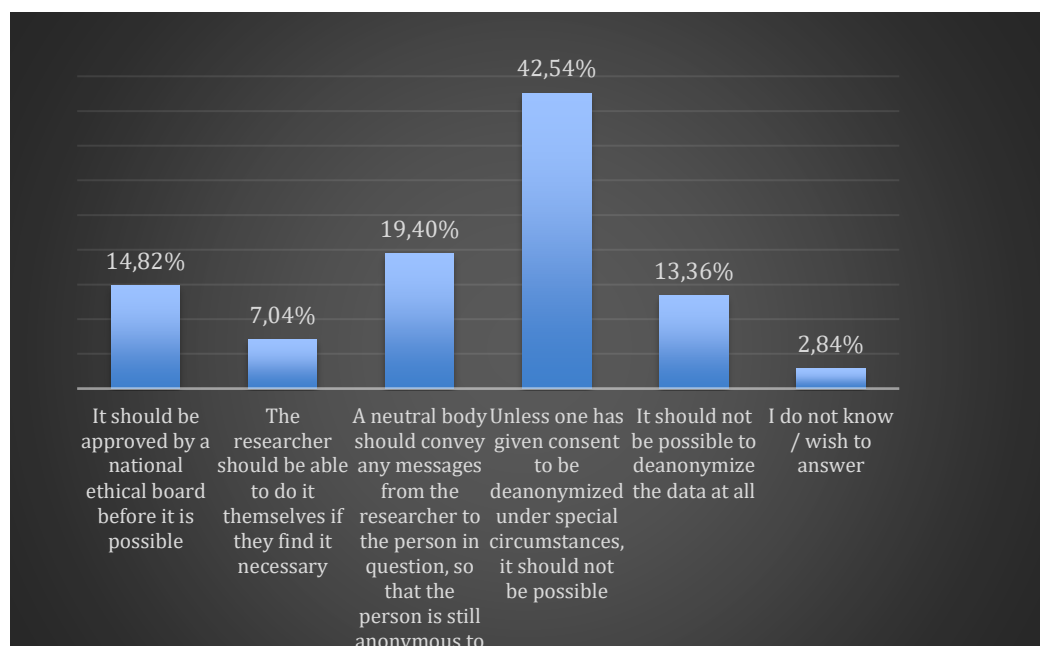


Figure 8 - Should there be some kind of deanonymizing mechanism so that it is possible to connect the data with the person behind them if necessary?

special circumstances. On the other hand, 24% of the participants believed that it would be acceptable to deanonymize the data, if it was a neutral body that was in charge and that would convey messages from researchers to the individual in question and 18% thought that the decision of approving deanonymization of data, should fall to an ‘national ethics board’. Clearly using AI in medical and health research, including for predictive profiling, is contingent on providing adequate privacy for the research subjects. Given that this is provided, there is widespread support for this application.

This seems to underline that the participants generally condone the use of AI and profiling for health and medical research purposes, but that deanonymizing data and contacting research subjects if they are found to be predisposed should only be done in case of serious illness. Furthermore, it was clear that this contact should not be carried out by researchers, but rather by a neutral body, in order to maintain the research subject’s anonymity to the researcher, which is still considered very important. Although the majority ‘are in favour’ of using AI in relation to health and medical research, a large proportion of the participants still consider their privacy to be more important.

2.5 Public Authorities’ use of AI

The technology used in medical and health research is not fundamentally different from that used in some other sectors. Thus, the consultation sought to investigate if the same technology applied in different sectors and for different purposes is equally acceptable. Public authorities are always under pressure to increase efficiency, save resources and at the same time deliver improved services. For these reasons, AI is seen as silver bullet by public authorities in many countries, whether it is the using it for processing applications and payments of social benefits or using AI to identify children currently at risk or predict which individual children might become at risk, or for predictive policing. The intentions are generally laudable, however, these applications also hold substantial risks.

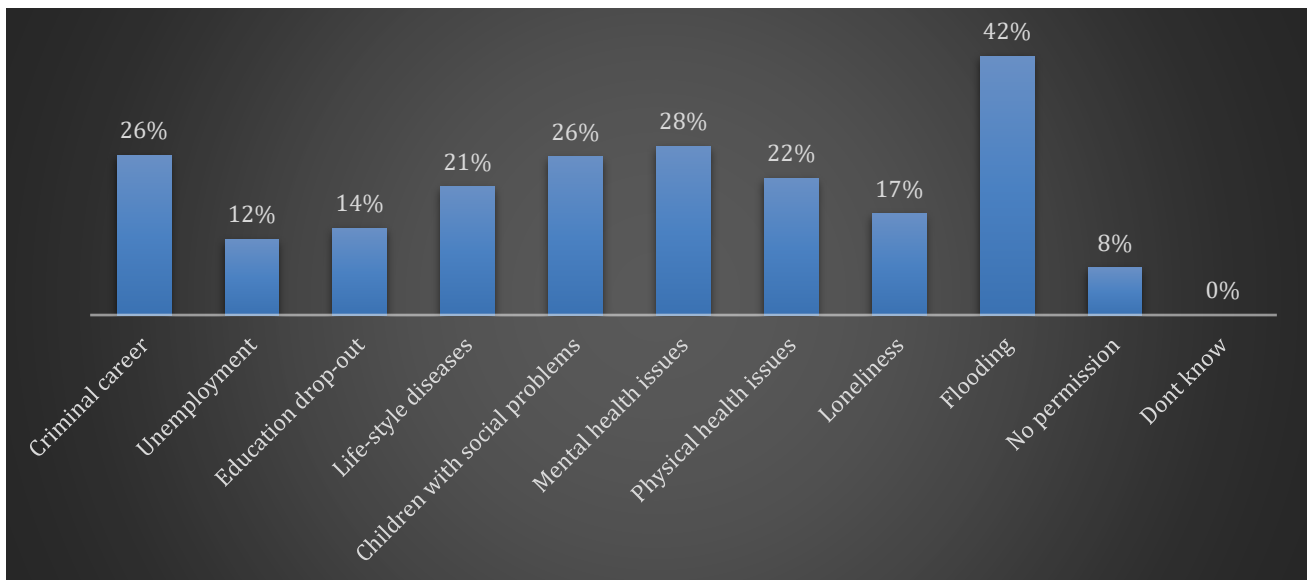


Figure 9 - Do you think public authorities (social workers, public administrators and similar) should be allowed to use AI and predictive profiling and cross-share information to identify people in risk of...

As can be seen in Figure 9, the participants generally thought that public authorities should be allowed to make use of AI; only 8% did not think so. The participants were allowed to choose as many options as they wanted, and the 813 participants who chose at least one made a total of 1846 choices, amounting to 2.3 choices on average per participant out of the 9 choices. So while a clear majority did think public authorities should be allowed to use predictive profiling, it was also clear that it was for most of them, this permission was limited to a few specified uses, and not a carte

blanche to use it across the board. Among the uses that were listed the use that most participants condoned was identifying people in risk of flooding (42%), while, on the other side, it is clear that of the other uses none received enough support that it can be argued that they were widely considered to be acceptable.

This result might be explained by two aspects. The fairly widespread support of using AI to identify people in risk of flooding might be because it is not very intrusive or invasive in the privacy of the individual, since the analysis does not pertain to any intrinsic characters of the data subject, and thus does not require personal profiling. This might in turn indicate that while using profiling is seen

Table 3 - Do you think AI predictions should be used to provide special treatment?²

Yes, good citizens should receive preferential treatment	4%
Yes, if you have nothing to hide, there is not problem	8%
Yes, but for other reasons	8%
No, I think it intervenes too much in the personal life of citizens	29%
No, everyone should receive equal treatment	39%
No, but for other reasons	7%
I do not know / wish to answer	5%

as less controversial for medical and research purposes, it is not seen as equally unproblematic when public authorities make use of this. The other aspect is that the discussion about what AI should be allowed to be used for, is still fairly new, despite the fact that AI technology and development has been around for decades. So clear positions on some of the more detailed questions are yet to clearly manifest themselves, though from the above there can be little doubt that predictive profiling by public authorities is seen as acceptable.

No special treatment of citizens based on AI predictions

Another suggested use for predictive profiling by public authorities, was to use it to determine who should have access to special treatment and provisions of public service. As can be seen in Table 2, this was met with widespread resistance. A total of 65% were against this kind of predictive positive discrimination, 39% because they believed that everyone should receive the same treatment, 29% because they believed that it intervened too much in the private life of individuals and 7% for other reasons. 20% found it ok, either because good citizens should receive preferential treatment (4%) or because if you have nothing to hide you have nothing to worry about (8%), and 8% for other reasons. This seems to support the notion above, that while AI and predictive profiling in medical research is generally not very controversial, when it comes to usage by public authorities, the participants are considerably more sceptical, and it would appear that they draw the line at using select citizens for special treatment based on predictions about their future behaviour, needs and potentials.

2.6 Create Trustworthy AI and Trust in AI Through Regulation and Consumer Empowerment

As described above, it is clear that there are some fundamental trade-offs when it comes to AI, as there is with most other technology. The participants wanted to make use of AI services and condoned its spread, but at the same time had a large range of concerns about them, including privacy and black-boxing. The question is how to ensure the right balance between realizing the positive potentials of AI, while curbing the negative ones. The discussion is a classic one between to

² Actual full question text: Do you think Artificial Intelligence predictions should be used to determine who should have access to certain rights and services such as cost and level of health care coverage, prestigious schools, certain kinds of beneficial interest rate loans etc?

sides. One position proposing free unhindered innovation as the best way of ensuring that the potentials of AI can be realized, suggesting that the gains to be realized will far outweigh the relative negative consequences, and the other position which argues that an unbounded development and application of the technology will have unsustainable levels of negative consequences. A branch of this position also sees good regulation as a way of steering technology in a way that creates not just more innovation, but better and more societally beneficial innovation. Currently there is no legislation specifically addressing AI. Instead it is regulated by a patchwork of sector and topic specific regulation. However, a clear recommendation from AI360 | COPENHAGEN was the need for a well thought out regulation of AI. For this reason, the participants were asked a number of questions pertaining to whether and how, AI should be regulated.

To gauge whether and how AI should be regulated they were asked if they thought algorithms should be regulated, and if so, how. There was a very clear general sentiment, as can be seen in table 10³ that AI should indeed be regulated. Only 4% of the participants, did not think so. The participants

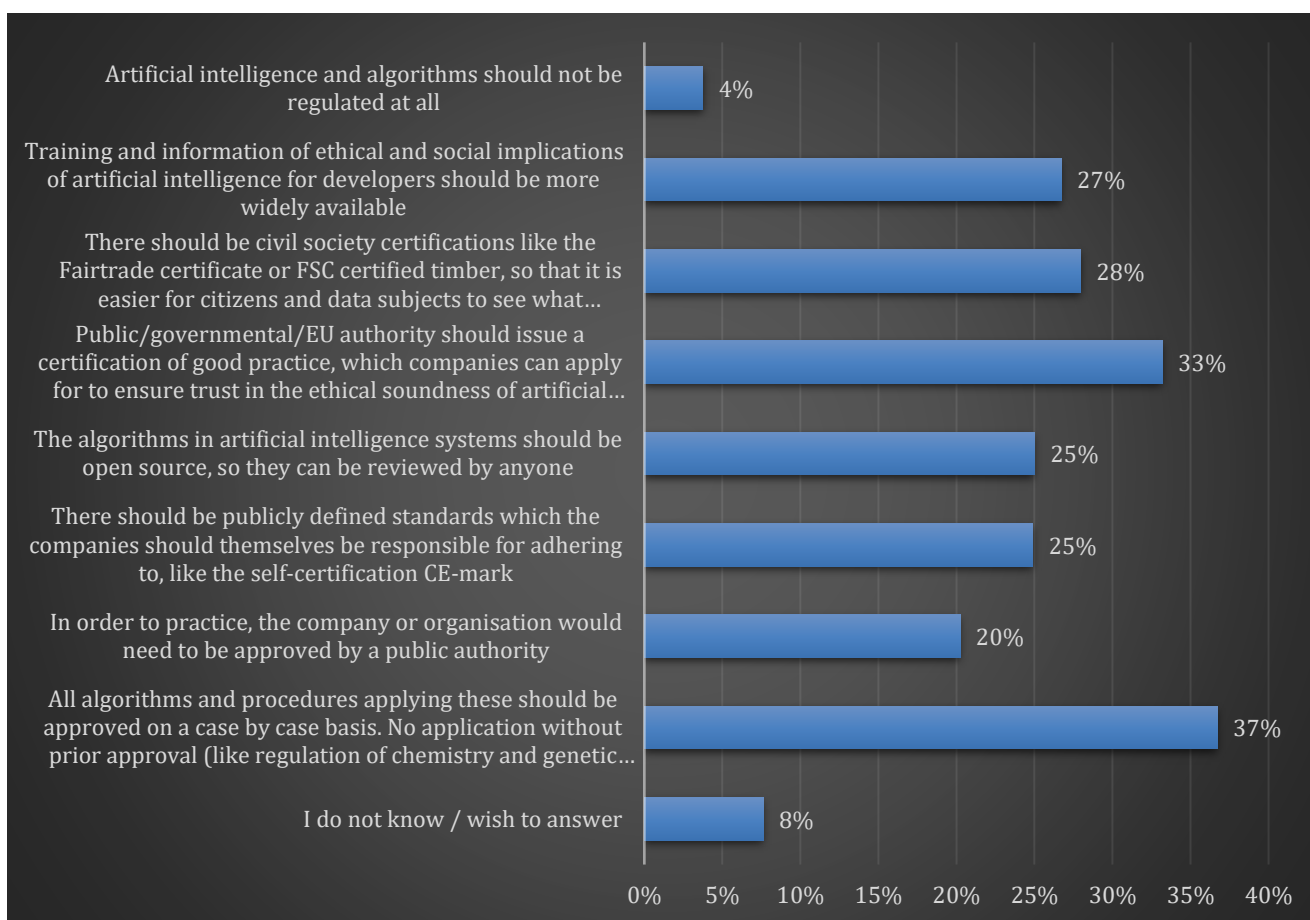


Figure 10 - Do you think algorithms should be regulated, and, if so, how?

were allowed to choose up to three options, and the 780 participants who thought AI should be regulated in some way, made a total of 1713 choices, averaging 2,2 choices per participants, which could indicate that the participants are in favour of applying more than one regulation. Of the available options, the one chosen by most participants was that regulation should happen on a case-by-case basis (37%), closely followed by a public, governmental or EU certification of good practice (33%).

Regulation on a case-by-case basis was a suggestion generated at AI360 | COPENHAGEN, thus the participants were also asked what criteria such an approval should be based on, and two options stood out clearly. 34% found that it should be based on ethical and normative acceptability, while

³ Full question text: “Currently there is no regulation of algorithms, how they work or how they are used in decision making processes. Do you think this is something that needs to be regulated, and, if so, how?” For the full text of the question options, see question 5.1 in Annex 3.

31% found that it should be based on risk of abuse through a multiple impact analysis. Only 3% did not think there should be any regulation, while 7% thought it should be based on functionality, 6% that it should be on bias and discrimination, while 10% found that it should be based on the sensitivity of the application area.

Table 4 - What is the best way to ensure that AI is developed responsibly?

Certification for responsible artificial intelligence, so consumers can see what artificial intelligence based products and services are responsible, and make informed decision	28%
Training of researches in ethics and societal impacts of technology	10%
Multi-disciplinary approach to artificial intelligence, ensuring broad competence spectrum behind the development	15%
Make schooling in IT, coding and knowledge of IT and its impact on social conditions part of engineers' education	6%
Use public research funding strategically, to create research in the technological and social solutions that underpin responsible artificial intelligence	6%
Require social, privacy and data protection impact assessments of all artificial intelligence systems	25%
I do not know / wish to answer	10%

The participants were also asked how best to ensure that the development of AI's happens in a responsible, socially and morally acceptable fashion. As can be seen in Table 3 the option that most participants preferred was a certification for AI, so that consumers can see what AI-based products and services are responsible and make informed decisions based on this. Almost as popular as this, was that a social, privacy and data protection impact assessment of all AI systems is made mandatory. This is supported by the fact that to a clear majority of participants it was very important to understand the rationale behind AI analyses. Seen in connection with the call for increased transparency and control over data described in section 2.3, this can be assumed to indicate that the participants were generally asking for implementation of means to create increased consumer empowerment.

Furthermore, results also indicates that the participating citizens strongly believe that the responsibility of ensuring that AI remains socially and ethically acceptable, should befall the public sector. When asked who should be responsible, 22% answered it should be the EU, while 21% answered national governments.

Conversely only 5% thought it should be left to the market, while 9% thought that product developers should be responsible. 16% thought that corporations offering AI products and services should be responsible, while 19% found that it should be the researchers developing the technology. That two of the three lowest scores fall to choices within a category that represent the private sector, namely 'The Market' and 'Corporations' could arguably indicate that the general public does not fully trust the private sector with ensuring social and ethical standards, in the future development of AI.

2.7 Conclusions

From the analysis it is clear that while the participants generally condone use of AI, and particularly use for medical and health research, there are also considerable concerns about the technology and its potential applications.

In the first section of the analysis, it was shown that the participants of EuropeSay generally were positively inclined to AI, but at the same time they had general concerns about its potential uses and the lack of human judgement. In continuation of this, it was also clear that they did not consider AI to be a stand-alone decision-maker, but rather believed that it should primarily be used as a decision-support tool.

In the subsequent section, and in continuation of the first section, the results clearly showed that the participants did not want black-box AI. The analysis shows that the participants found it important that AI is explainable and transparent.

In the third section of the analysis, the results showed that the participants made widespread use of AI based services, but at the same time were concerned about what the data they leave behind can tell about them and what it can be used for. At the same time they reported not feeling that they had control over their own data. Thus, there was an interesting conflict between the want to and need for using these services, but at the same time being concerned about the data that they collect about them and what this can be used for.

The fourth section of the analysis showed that there was widespread support for using predictive profiling in medical and health research, and that participants generally wanted risks of physical and mental illnesses discovered in research to be reported to them.

The fifth section found that public authorities should also be allowed to use predictive profiling, but only in limited ways.

The last section of the analysis showed almost all participants were in favour of implementing regulation on AI, as well as means to empower consumers when choosing whether to use AI based products and services.

3. Methodology

3.1 EuropeSay

EuropeSay on AI was a European-wide citizen consultation on the topic of AI making use of the GlobalSay methodology, and it ran from September 2019 to January 2020. The focus of the consultation was guided by the results of AI360 | COPENHAGEN, which was likewise organized by the DBT under the auspice of HBP. A description of AI360 | COPENHAGEN can be found below.

The Danish Board of Technology organized and facilitated the consultation, and coordinated the Danish meetings as well as the partners in 12 other countries, that were responsible for promoting meetings in these. As can be seen in the table below, 157 meetings took place across all countries, where a total of 928 participants were engaged.

Country	Number of meetings	Number of participants
Bulgaria	13	72
Croatia	11	62
Denmark	13	87
Finland	10	58
Germany	12	86
Hungary	10	48
Italy	11	71
Lithuania	11	79
Poland	13	75
Portugal	14	74
Slovakia	17	79
Spain	12	70
United Kingdom	10	67
Total	157	928

The methodology applied is called *GlobalSay*. GlobalSay is a concept for distributed dialogue that is designed to engage citizens in deliberations about select topics. The methodology is inspired by the WWViews-concept which has multisite face-to-face citizen summits, where 100 citizens or more in each place meet face-to-face at a specific time to discuss a topic and subsequently answer a range of quantitative questions about this topic. With EuropeSay, however, the citizens are engaged in micro-meetings of 5-8 participants which are organized by regular citizens volunteering to host the meetings. The meetings can take place where and whenever it is convenient for the participants, and instead of having a human facilitator, as with a citizen summit, the event is facilitated by an online platform: EngageSuite. At the meetings, participants will gather around a computer in the living room of the host, the break room of the office, the local library or where it is convenient. The online platform will guide them through a predefined process during which they will engage in consecutive rounds of deliberation, alternating between presentation of information in the shape of

short texts and video vignettes and deliberation on questions addressed by the videos. At the end of each round they will be asked to provide answers to a range of questions with predefined answering options.

In short, the methodology enables anyone to invite friends, family, colleagues etc. to join them for a face-to-face deliberation, wherever they like and whenever they like, using a digital platform to facilitate and inform the meetings and to collect the results.

GlobalSay in short:

- Self-organized: Anyone can set up a meeting and invite whoever they would like to. No prior knowledge or interest in the topic is necessary to take part.
- Distributed: The meetings can be set up where and when it is convenient for the host. The only requirement is a computer with internet connection.
- Digitally supported: Meetings are supported by an online engagement platform, EngageSuite, which provides the structure for the meetings and facilitates the deliberation, as well as collecting the participants' assessments and opinions. In addition, it enables everyone to participate in their own language.

Thus, the methodology provides the flexibility of online participation while ensuring that participants have had time to reflect over their answers and have had their preconceptions and values challenged in open face-to-face deliberation, thus producing informed and considered responses to the questions posed.

The methodology is designed to run along two tracks. In the one, a partner recruits 10 hosts locally in the country in question. These are chosen based on a range of criteria intended to ensure diversity of the hosts. As there are only 10 of these meetings per country, representativeness is not an option, so instead the intention is that the recruitment should ensure diversity among the hosts. There are, however, no constraints to who the hosts want to invite.

In this track of the consultation, a minimum of 650 citizens from 13 European countries are engaged.

The other track, which is a central part of the methodology, is that anyone who is interested, can organize and host a meeting. The participation process is intended to snowball by viral diffusion of the offer to host a meeting. So beyond the 10 directly recruited hosts, the possibility for taking part in the meetings is open to everyone. The hope is that this will travel and citizens will start organizing meetings on their own. Each local partner is responsible for putting in an effort to this effect.

With this composition of participants, the methodology can and does not make claims to statistical representativeness. Rather what it seeks to achieve is a well-grounded picture of the quantitative tendencies in the European population, both at the national and transnational level. This picture can be very useful for getting a rough understanding of what the European population thinks of AI and how its development should be steered, and thus lay out a first guiding line for policy-makers to follow.

A central aspect of the GlobalSay methodology is having local partners in each country. This is both necessary in order to provide reliable translations of the materials into local languages, but it is also necessary for recruitment and promotion. The local partner in each country is thus responsible for translation of EngageSuite content, promotion material and producing voice-over for information videos, for recruiting hosts and promoting the consultation further to make it snowball. In addition, they provide hosts support with setting up meetings and using the EngageSuite platform. To read more about EuropeSay on AI

3.2 AI 360 I COPENHAGEN

AI360 I COPENHAGEN was designed and organised by the Danish Board of Technology and took place in Copenhagen, Denmark, in the spring of 2019 with the aim to encourage the engagement of stakeholders and experts on the topic of AI in thorough deliberations about the prospects of future developments.

The AI 360 COPENHAGEN workshop, provided an all-around perspective (360 degrees) where various trade-offs related to different choices, developments and implementations may be considered against each other. Furthermore, the aspiration of AI 360 COPENHAGEN was to ‘go one step further’ than most other debates and analyses of artificial intelligence and put an explicit emphasis on concrete solutions to the identified challenges posed by AI technological developments.⁴

⁴ You can read more about AI360 I COPENHAGEN at <http://hbp.tekno.dk/events/ai-360-l-copenhagen/> where you can also find the recommendations report which presents the primary findings of the workshop.

4. Annex 1: References

5. Bibliografi

Bitsch, E. R. (2020). *Trust and Transparency in Artificial Intelligence*.

Bitsch, L. K. (2019). *AI 360 Copenhagen: Report from the workshop*. Hentet fra https://sos-ch-dk-2.exo.io/public-website-production/filer_public/a4/f2/a4f2aabf-6821-4a8d-b082-5070e2797b27/a

Blanco-Justicia, A. D.-F. (to appear). "Machine learning explainability via microaggregation and shallow decision trees.". *Knowledge-Based Systems*.

BusinessDictionary. (31. March 2020). *BusinessDictionary - Enabling Technology*. Hentet fra BusinessDictionary.com: <http://www.businessdictionary.com/definition/enabling-technology.html>

EC High-Level Expert Group on Artificial Intelligence. (2019). *Ethics Guidelines for Trustworthy AI*. Brussels: European Commission.

European Commission. (2. February 2020). *White Paper On Artificial Intelligence - A European approach to excellence and trust. COM(2020) 65 final*. Brussels.

6. Annex 2: Participants Demographic Data

