Moral and ethical side of using of AI in modern life

Ilia Dunshin

Haaga-Helia University of Applied Sciences

Business Information Technology

Bachelor's Thesis

2022

Abstract

Author(s)
Ilia Dunshin

Degree
Bachelor of Business Information Technology

Report/Thesis Title
Moral and ethical side of using of AI in modern life

Number of pages and appendix pages
33 + 5

This thesis explores AI development topic in terms of its regulations by different countries and people's opinions on it. During work not only scientific questions were raised, but even philosophical ones.

The work begins with an introduction where an explanation about reasons standing behind such topic choice is, as well as research about how and when AI industry has started its existence. Moreover, objectives for thesis and methods, which were used to achieve them, are also in introduction part. Last thing contained in introduction part is keywords definitions, where main keywords, that might be came across during reading of this thesis, are explained.

The theoretical framework focuses on current state of AI. Research about how AI is regulated worldwide was done and results were deeply analysed and compared to each other. Besides, research about people's opinions on AI industry at whole and on AI regulations, in particular, was done as well. As a result of this research work, a set of rules to be considered, while working in AI industry was created and examples of different ethical dilemmas were presented and analysed. Last part of the theoretical framework is devoted to ECCOLA method, which was created by Finnish students as a guidance for check if an AI product corresponds all existing moral and ethical regulations.

As an empirical part, a survey about AI development is created, where are 14 questions about different aspects of AI and its regulations. Respondents were both IT professionals and usual people. As IT professionals FCAI (Finnish Center of Artificial Intelligence) members and Haaga-Helia UAS students on IT programs were contacted. In total, there are 24 respondents.

Conclusions and discussion points were gathered and presented on the last stage. In the end knowledge gained from research made in theoretical and practical parts, is reflected, and analysed. Besides, there is a discussion about possible improvements and different approaches, which could be used while working on this thesis.

Key words
Artificial intelligence, regulations, development

Table of Contents

1 Introduction

This is a research-type bachelor's thesis, in which all aspects of moral and ethical side of working with AI are formulated and considered. This is done by using of different methods of finding the information and working with it.

This topic is very relevant nowadays, because the AI field is constantly growing and developing, which creates a strong need of ability to correspond to this development. As a result, the topic is not only relevant in today's world, but has a lot of useful components for the future studying.

1.1 Background

For the last decades AI as a topic is arising. It all started in the mid-20th century by Alan Mathison Turing. He was the first man, who described an abstract machine, which now can be counted as a descendant of a modern computers and AI machines. In 1947, during one of the Turing's lectures in London, he was speaking about computer intelligence. He stated that he wanted to create a machine, that can learn from its own experience. This goal is reached nowadays, and we have Machine Learning as its result.

In 1950 Alan Turing introduced a test, which is used for checking computer intelligence existence, and is currently known as the Turing test. It consists of people who are holding conversation between each other. The main idea of a test is to replace one of them in such a way, that the main interrogator would not be able to distinguish any differences. Turing has predicted that computers would easily pass his test in 2000, but even now just a few succeeded. Moreover, they did well thankfully to clever ways of fooling judges rather than to their overwhelming memory and computer power.

The next step in development of AI was a program written by Christopher Strachey in 1951. It was the first computer checkers. By the summer of 1952 his program was upgraded to such a level, that it could play a complete checkers game at a reasonable speed. (Copeland s.a.)

A lot of time has passed since then, but the interest to AI topic has only become more. Nowadays, this topic is controversial and has different, sometimes opposite point of views, starting from strong approving of all the AI development processes and ending with fear of total domination of AI in the future.

But how can we define AI nowadays? Basically, AI is a software, which has a mechanism to learn. It can have different types of learning, supervised by a human or unsupervised. Also, AI in general can be divided into type groups, but all these aspects will be discussed later in this thesis.

1.2 Objectives

This thesis has several objectives. The main outcome is a set of rules, which should be considered, while working with AI, which represents basic human rights.
Other targets are:

- Finding and analyzing of the information about current state of moral and ethical regulations in AI usage.
- Making predictions about the future development process of AI.
- Gathering an overall information about how and in which areas AI is being used nowadays.

There will not be concrete guidelines of how different AI applications should be used in this thesis. Also, there will be not exact timeline of history of AI development, only a few main points were discussed in the introduction.
The main aim of this thesis is to be objective. A lot of different points of views on AI development will be looked at, but all of them will be considered with scientific background and without any personal preferences.

1.3 Methods

In this thesis qualitative research method was used. The process of researching was constant and iterative. This required returning to the same resources from time to time to gain an information for gathering it in a better way later.
This approach was chosen as the number of resources which are available and useful for this topic is enormous.
Not only English language resources were used during research process, but Russian as well.
After the main part of research had been done and all the needed information was gathered, the analysis stage was started. During this stage, the main aim was to structure the information in a clarified, understandable, and easy-to-read way. But the research process did not stop on that stage yet. As was said above, it was a constant process, and when the work was coming to different parts of thesis, it was already known where to get the needed information thankfully to iterative research approach. Speaking more of analysis stage, most objectives of current thesis were made exactly during it.
The last stage is finalizing. During that, the analysis of a survey about AI development, was done. Besides, conclusions about all the work done before were also made during that stage.
First two stages were divided with pretty much the same amounts of time (around 20 days), but the last one required less (10 days).
The main resource, that was used to look for relevant information, is internet.

1.4 Key concepts

AI (Artificial Intelligence) – a software, which has a mechanism to learn.

FCAI (Finnish Center of Artificial Intelligence) – a community of experts in artificial intelligence in Finland, initiated by Aalto University, University of Helsinki, and VTT Technical Research Centre of Finland.

ML (Machine Learning) – a type of AI, that allows applications to self-develop. (Burns 2021)

Petaflop/s-day (pfs-day) – Unit of measure in Machine Learning. Consists of performing 10^{15} neural net operations per second for one day. (OpenAI 2019)

Supervised Learning – a type of ML, which is defined by its use of labelled datasets, which are being inputted into the model. (IBM Cloud Education 2020)

Unsupervised Learning – a type of ML, which is defined by its use of unlabelled datasets without the need of human intervention. (IBM Cloud Education 2020)

2 Current state of Artificial Intelligence

Nowadays, AI is being used in many industries, such as Finance, Education, Agriculture, Entertainment, Healthcare, and others.
As one of the main features of AI is autonomy, it is a special criteria, when we're speaking about AI. We can define these levels by looking at Society of Automotive Engineers (SAE) (Krishnakumar 2022) categorization:

- Level 0: No automation
- Level 1: Light automation (for example, cruise control)
- Level 2: Some automation, but human attention is constantly required
- Level 3: Self-driving ability, but required human intervention in severe conditions
- Level 4: Highly autonomous
- Level 5: Full automation

We can look at this categorization at more global sight and divide it onto 2 groups: First 3 levels (0-2) require human intervention constantly, and others 3 (3-5) can be named as 'self-driving' in overall. (Krishnakumar 2022)
Also, AI development will completely change business organizations' model. Currently, there are many job titles, such customer support, marketing specialist, administrative assistant. All these positions will be replaced by AI versions in the future completely. Other duties such as graphic design assistance can be placed on freelancers' shoulders. As a result, businesses organization will look completely different from an old option.
One more feature of AI today is its exponential development. The case is that we were at its pivot point in 2012. On the graph below (Figure 1) we can see how rapid AI growth has been since 2012. At first, everything was developing accordingly to the Moore's Law, which stated, that AI computing power doubled about every 2 years. But since 2012 this power doubled in 3 months

regularly. (Brown, 2020)

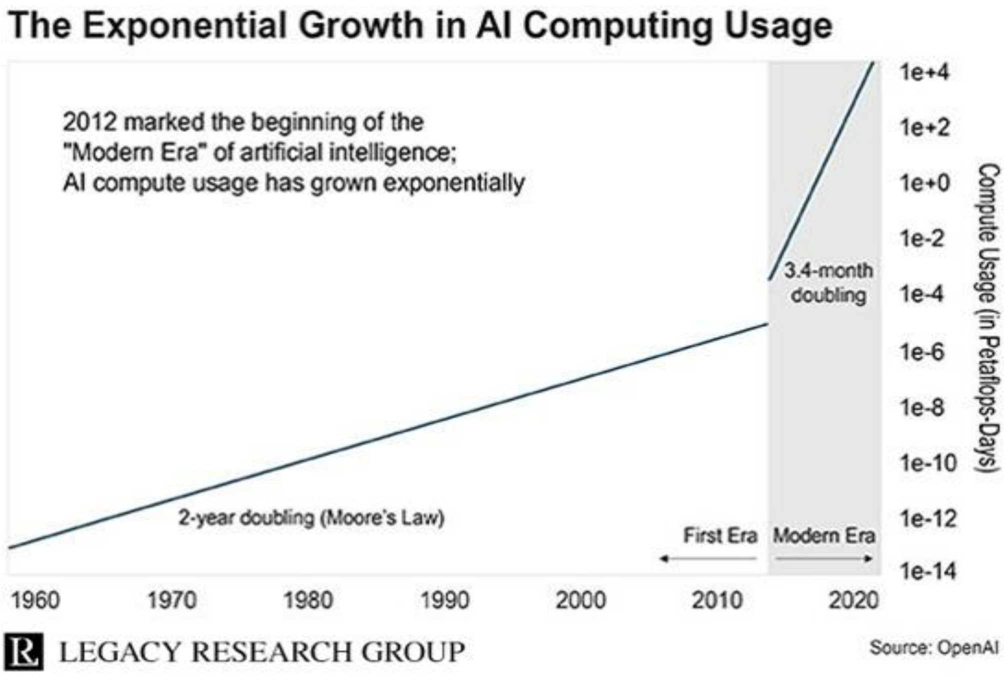


Figure 1. The Exponential Growth in AI Computing Usage (Brown 2020)

2.1 People's views on AI

People's views on AI nowadays are completely different from each other (they can be even opposite). On the graph below we can see the results of survey held by Pew Research Center.

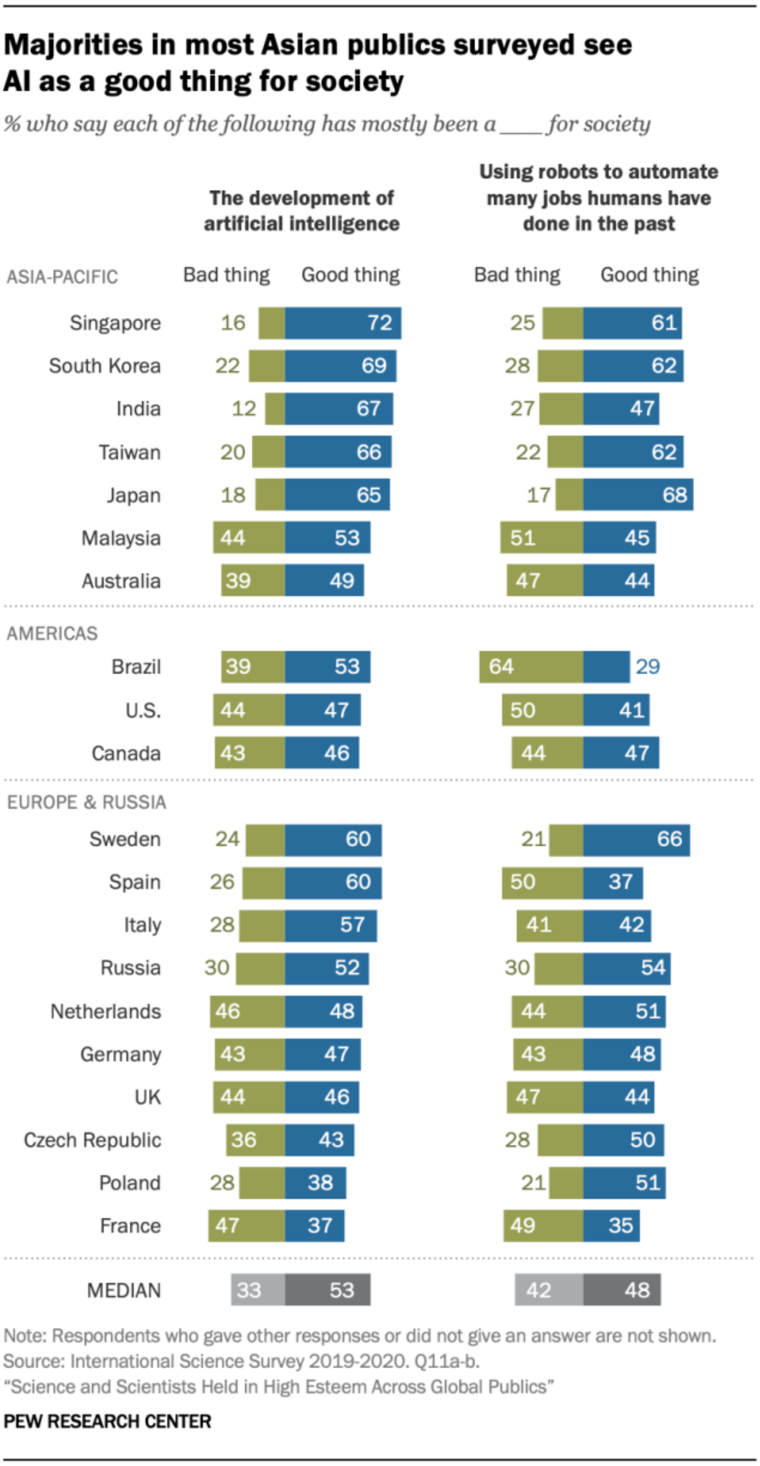


Figure 2. Percentage of supporters and detractors of AI development worldwide (Johnson, Tyson 2020)

According to this survey (Figure 2), in Asian countries most people are considering AI development as a benefit for the world, while in Europe the percentage of people with the same view is much lower. Nonetheless, more people support AI development and see using robots to replace humans on some jobs as a good thing. (Johnson, Tyson 2020)

2.2 Advantages and Disadvantages of AI development

Of course, the reason for the fact that AI development is such a controversial topic is that it has lots of pros and cons.

- **Advantages**. Speaking of benefits, we can start with human error reduction. When a human performs an action, the overall performance depends not only on human's skills and abilities, but also on emotional conditions. Also, people usually do not have an option of doing a job without being distracted by external circumstances and even so, there is always a chance of a mistake. The only thing a human can do is to reduce this chance, but it is impossible to get rid of it completely. In that case, AI development is what can help people a lot. Robots do not have feelings and emotions; therefore, this kind of error will never happen.
  One more benefit is taking risks instead of human beings. As science is developing all the time, it can sometimes lead to new dangerous unknown situations for people. In the last century it was flying to a space, for example. Soon it can be a settlement of different new planets or exploring the deepest parts of oceans on Earth. It will be good if we are able to send robots there firstly to check if it can be dangerous for us and to prepare for that beforehand. Not only exploration of new locations is making a need of safety to people, but also different disasters that are occurring on our planet all the time. All of them require people to eliminate dangerous consequences but these people are taking risks. Replacing them with robots would save lots of lives.
  One more benefit is robots' availability, they don't have any needs that people do, they do not require sleep or food. Therefore, they can perform their duties 24/7 without any breaks. This can improve the quality of immediately required services such as services required in hospitals.
  Also, AI is able to work with information much faster than a human. It has an ability to process a lot of information in just a second and as a result it can come up with the best decision possible. People sometimes can't come up with the best solution possible, because it is hard for us to take in consideration all the information and our conclusion can be not the optimal one. (Kumar 2019)

- **Disadvantages**. Of course, on the other hand there are some disadvantages. First one is high cost of robots' creation. As said before, AI is developing constantly, which means, that there is a constant necessity of different updates, which requires lots of money.
  The other disadvantage, which is spoken about mostly by AI development opposers, is

increasing of unemployment. If we replace some people with robots, it may be hard for some of those people to find a new job, or at least it requires a lot of time.

One more disadvantage of AI development is its high cost in terms of energy. Energy consumption is one of the most important problems in today's world, especially in Europe. And AI development can harm the environment seriously due to its high energy needs.

Last one is a lack of creativity. Robots can only perform tasks which they were intended to. That means, that if something unpredictable happens, robots wouldn't be able to act accordingly, and an immediate human's interference would be required. (Eliaçik 2022)

3 Ethical regulations of AI

AI regulations question does not have enough attention yet. Even though many people are speaking about its as about necessity, the situation with this aspect is far from excellent. Though, AI is constantly developing, and its usage is becoming more and more massive, we still can't agree on a set of instructions and regulations which should be considered in any case. Moreover, all the instructions vary a lot in different countries, especially if we compare EU and USA. But what almost everyone can agree on is that regulative work should be done, because otherwise people just put their lives under the significant impact of a new industry, nobody is responsible of. Ensuring of accountability and fairness is crucial when we're speaking about any human's life.

3.1 EU regulations

In 2018 the European AI Alliance was created. At the beginning it was started as a forum for discussions, but as the time comes, the Alliance became a huge community consisting of EU citizens, business organizations, different unions, and AI experts. The main aim of this group is to define the main problems with AI and to solve them, as well as use most of the opportunities offered by AI in overall. At the same time, the special high-level expert group was created by European Commission. It consisted of 52 people, and their responsibilities were pretty much the same. They should have been checking the situation with AI constantly and give recommendations on how to cope with current challenges and how to take most of the AI opportunities. These recommendations then were used in creation of draft ethics guidelines, which were eventually created fully in the beginning of 2019. The members of this Expert Group were selected by European Commission itself from around 500 applications.

In February 2020 European Commission has presented a White Paper, in which EU approach regarding AI regulations was. This document was largely discussed by AI experts and by community at whole. As a result of these discussions, a Proposal for a Regulation of AI and a Coordinated Plan Review appeared as the main resources for AI regulation.

By the EU Commission, it is crucial to have regulations on AI, because some AI systems may act in a way, reason for which would not be understandable for people. In that case, it can lead to wrong results and harm us.

Proposed rules set clear requirements for AI systems, proposed a governance structure in Europe, defined special obligations for AI users, and finally created a distribution of AI applications in 4 levels:

- Unacceptable risk
- High risk
- Limited risk

- Minimal risk

As clear from the name, AI systems within an unacceptable risk division are considered as a threat, therefore they were banned completely. (An example of AI system, which is under that division is all forms of social scoring made by governments).
High risk list contains a huge number of systems, but the main examples are: critical infrastructures and public services (for example, transport, or a system, which is used for a review on citizen's credit score to make a decision on his availability to obtain a loan), educational training (for example, systems which can be used for counting students' exam score), employment training (for example, CV-sorting applications), and border control management. These systems can be used unlike ones under the unacceptable risk division, but they will have strict obligations before they can be used. They should have a compulsory logging of activity to ensure results' traceability, a high-qualified datasets, which are being fed to such systems, detailed documentation providing all the necessary information on the system and its purpose of usage, high level of security, accuracy, and robustness, and lastly a clear and correct show of information to the system's user. After such system is created and it corresponds all the requirements mentioned it should be registered in an EU database of AI and only after it can be placed on the market. (Figure 3)

Figure 3. Instructions on a new created high-risk level AI system (European Commission 2022)

Limited risk systems are more common to be came across in the life, but they still do have some obligations. The most important of them is a compulsory aware of a user, that he is interacting with an AI system. A good example of such system would be a chatbot.

The last one division is minimal risk one. In the proposal it is allowed to use AI systems from that division freely. Good to note, that most of all systems that are existing nowadays in the EU are under this category. As an example, we can take AI-enabled videogame. (European Commission 2022)

But not only White Paper and Legal Framework were things published for AI by European Commission. On 28th of September 2022 an Artificial Intelligence Liability Directive was created. Aim of that was ensuring that people harmed by AI will be protected with laws as well as people harmed by other technologies. Liability rules earlier were not adapted for handling compensations for people harmed by AI. Besides, protecting victims of AI-products, the Liability Directive should also have reduced legal uncertainty in AI usage. Also, the Directive 'is expected to generate efficiencies and contribute to the innovation of environmentally friendly technologies'. (European Commission 2022). It was estimated that this new approach will increase overall number of insurances paid annually in the EU to EUR 16.1 million..

The last one document which can be reviewed in terms of European approach to trust in AI as a Proposal for a product liability directive. In overall, defective products were considered in a document since 1985. But AI, of course, was taken into account much later. A compensation for material damage is limited with lower border of EUR 500. In 2020 the European Commission published a new report, which was implied into liability for defective product, where an AI was discussed. In this report all the opportunities and characteristics of AI were viewed and defined. For the conclusion, it was said that AI field does still require a lot of work be done to become a stable and reliable field for every EU citizen.

3.2 USA regulations

Speaking of the USA position towards AI regulations, there is not so much done as done in Europe. Starting from Obama's administration and continuing with Trump's one, USA has always seen AI technology as a strong impactful power, which shouldn't be regulated manually, but should constantly be developed and used in different areas of life, especially in politics. In October 2016 a document named 'Preparing for the Future of Artificial Intelligence' was published by Executive Office of the President of the USA. This paper contains an overall view on AI, as well as view on its impact on people's lives and areas of its usage. There is also a paragraph about AI regulations there as well, but there were not concrete offers for regulations. Instead, there is just a statement, that AI should be developed constantly, and that it's crucial to ensure safety of people before bringing any global innovative AI systems into mass-market. What is also contained in 'Preparing for the Future of Artificial Intelligence' is an acknowledgement that AI is considered as the main progressive force nowadays, and its developing could bring humanity to a completely different new level.

Nevertheless, the regulatory aspect of AI industry is not so one-sided in the US, because it has a federal unit, which allows states to enact different laws separately if they do not contradict the US constitution. As a result, for example, some states have already banned the use of facial recognition technologies firstly (for example, Virginia and New Orleans), but then their governments approved a bill allowing the use of this technology by police in some cases. (Galaski 2021)
As a conclusion, it can be said that while European Commission is trying to create a new regulations approach specifically for AI, USA government has a different point of view, and is trying to pay attention to including AI in already existing regulative system. And mostly USA government is interested in AI development and fast including of innovative systems straight into the work. (Ebers, Navas, 2019)

3.3 China regulations

Speaking of Chinese regulations, on the 1st of March, 2022 Chinese government has published a regulation document in which main points of working with AI were made. According to this document, different algorithms which are being used in AI systems should correspond to moral and ethical standards and be transparent and accountable. In other words, companies from that date on should notify users when AI system is being used to provide their online customer services and also, there should always be an option for a user to kick himself out of AI targeting system (speaking of only online services). (Howell, Chanley 2022)
Notable mention, that the China does not only make attempts to regulate AI but develops its constantly. And AI developing, which is heavily supported there requires much more work being done in regulation's aspect. On 6th of September 2022 Chinese government has published a new regulation document named 'Regulations on Promoting artificial intelligence industry in Shenzhen Economic Zone (the Shenzhen AI Regulation)'. In overall, this document encourages local Chinese governments to develop AI sector massively not only just by plain motivation but by big financial donations made from the national government. According to Shenzhen AI Regulation, AI services in China, which are considered as a 'low-risk' systems can go for testing and eventually in a massive production event without national norms, but only if these systems adhere to international standards. The distributions on the risk levels and official guidelines will be created by Shenzhen government in the future. At the same time, the requirement for the establishment of AI ethics council is also written in the Shenzhen AI Regulation. The main aim of this establishment will also be developing of different standards of AI safety as well as examining of AI impact on different crucial aspects of life, such as employment and data protection. (Wu 2022.)

3.4 UK regulations

There is also work done in terms of AI regulations in UK. Published on 18th of July 2022 Data Protection and Digital Information Bill is one of the main documents about AI regulations published by UK government. The main content is 6 core principles, that should be taken into account when the work with AI systems is performed. They are:

- Ensuring that AI is used safely
- Ensuring that AI system is secure technically
- Ensuring that AI system is transparent, and its work results are explainable
- Ensuring fairness in AI systems working
- A requirement to identify a legal person to be responsible for AI working
- Clarifying routes for redressing or contestability

Besides, the UK government also has published a national AI Action Plan. According to this plan, the government is also considering constant huge investing for the long term needs of the AI ecosystem as a beneficial and important thing. Besides, the aim of government is also to transit national economy to an AI-enabled one and ensure that AI is producing benefits for each sector and region. Of course, it all couldn’t work without proper regulations, so according to the AI Action Plan, governing AI effectively is also one of the most important tasks in future developing of AI sector. (Collins 2022)

4 Public opinion on AI regulations

Most people consider AI as a technology, that should be managed carefully. On the graph below the results of the survey made by Eurobarometer in 2017 can be seen (Figure 4). It is obvious that most people do agree on the need of strict AI regulations and only small percentage are against it all over the world. Most inhabitants of Sweden, the Netherlands and Greece do agree with the need of careful management of robots and AI systems at whole. At the same time, the least percentage of the supports in Europe can be found in Malta, Italy, and Poland. Even though, in these countries the percentage of strong supporters is below 50%, the percentage of strong opposers remains low in every European country. The overall strong support in Europe is a little bit above 50%, and the answer “tend to agree” was chosen by 32% of respondents. (Zhang, Dafoe 2019) Good to note, that in Asian countries (such as Japan and Malaysia) the overall usage of AI systems is higher now, and people know a bit more about this technology field. As a result, the percentage of AI regulations supporters there is even higher than one in other parts of the world.

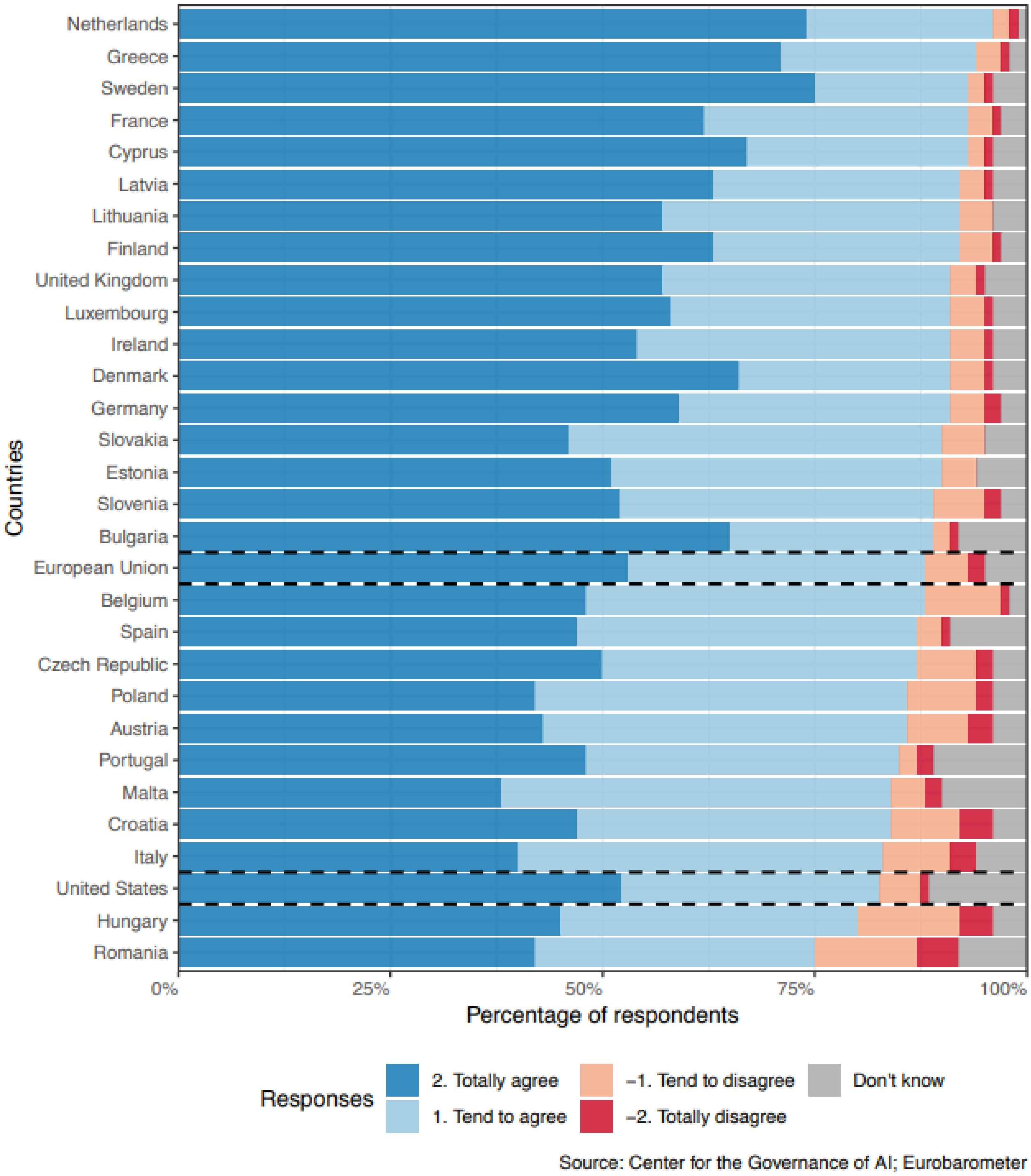


Figure 4. Agreement with statement that robots and AI require careful management (Zhang, Dafoe 2019)

5 Set of rules to consider while working in AI industry

Since the start of AI industry, people were trying to find a way to regulate AI systems. In 1942 Isaac Asimov proposed his view on a moral regulation of AI, known as ‘Three Laws of Robotics’. These three laws are:

1. A robot may not injure a human being or, through inaction, lead a human being to come to harm.
2. A robot must obey the orders, which were given by human beings, except where such orders would conflict with the First Law.
3. A robot must protect it existence, except such protection would not conflict with the First or Second Laws.

These laws still can be applied to AI, but as they are abstract and some points cannot work in a proper way, this set can’t be relied on as a strong option. (Isbert, Koering 2020)
Currently, there is more broaden options for a work in AI industry to be regulated more properly. One of them is Marcel Isbert’s (he is a member of the European Artificial Intelligence) proposal, which was published on the official website of European Commission. His set of rules is:

1. AI shouldn’t harm any individuals or group of those.
2. The usage of AI must not allow a person to be discriminated against by the use of an AI system.
3. The AI system should be correctable by AI experts.
4. The data underlying in AI system should be adaptable and revised in every situation.
5. An AI system requires data quality and constant control by AI experts.
6. The data should be checked by the AI experts before importing into the AI system every time.
7. An approach for the supervised training of the AI system should be clean and transparent to ensure high data quality.
8. The AI system should not be used for discriminating against or exclude individual customers and group of those.
9. The AI system should be used for supporting decisions, and should not be an autonomous system.
10. The AI experts should put their own stamp on the data analyzed with the AI system to ensure transparency and traceability.
11. The AI system should be used for data-based decisions. (Isbert, Koering 2020)

This set of rules should be used by companies when they are using AI systems in their work.

6 Examples of ethical dilemmas in AI

1. Autonomous Car

While driving a car, there can always appear a situation, where a moral judgement should be done, for example a quick, dangerous turn to side if a person is standing in front of the car to save this person. But AI system shouldn't be programmed in such way at least with modern level of technology development. Therefore, there is an important moral question, should we already use autonomous cars even though they can't make good in a moral way decisions. And there already was an incident involving autonomous car in 2018, when an Uber self-driving car hit a pedestrian, who then died in a hospital. (Dilmegani 2022)

2. AI in Art

In 2016 there was a case, when AI has designed a painting named 'the Next Rembrandt'. It was done due to AI analysis of 346 paintings and as a result an amazing painting was created. And this one is not the only example of AI integration into world of art. In 2019, Huawei announced, that their AI system has been able to complete the Schubert's Symphony №8, which was unfinished by a composer. These cases are arising the question about moral ethics of using AI in art, and who should be then counted as an author of a new-way-made masterpieces? It is highly probable, that due to AI development even fundamental questions such as 'What is creativity?' will have new answers. Anyway, all of that require new definitions of culture and art and new approaches towards plagiarism and piracy.

3. AI Usage in the Court of Law

This is probably the most important ethical question to people, as this one has a direct impact on their lives. Often, a decision in a court cannot depend only on laws. One of the most important factors in decision definition is a motivation of a human, when it's considered, it can either soften or make a decision heavier and therefore will change human's life completely. Besides, as was said above, even now it is crucial to ensure AI transparency, but how can we track it when we're speaking about decision in a court in every exact case? (UNESCO s.a.)

4. Unemployment

This is counted as one of the main fears about AI development in the future. According to McKinsey estimates, AI systems could replace up to 30% of the world's human labor by 2030. Of course, this is an important ethical question, and many people are worried about it. At the same time. The people density is increasing all over the world constantly, which means, that, on the contrary, much more jobs will be required. (Dilmegani 2022)

7 ECCOLA Method

As set of requirements and methods for creating of an AI system does not still exist, there was a need in new approaches towards AI regulations. This need led to appearing of ECCOLA method. Good to note, that ECCOLA method does not provide a view on AI regulations instead of existing ones but takes them into consideration. The sense of creating of ECCOLA method was to decrease gap between research and practice in AI regulations area. ECCOLA method consists of 21 cards with some statements about a topic laying down within AI regulations area. There are 8 main topics, and 21 cards are distributed between them. These topics were chosen from other different ethical guidelines about AI.

ECCOLA card themes.

Card themes (8)	Card number	Card amount (total 21)
Analyze	#0	1
Transparency	#1–6	6
Safety & Security	#7–9	3
Fairness	#10–11	2
Data	#12–13	2
Agency & Oversight	#14–15	2
Wellbeing	#16–17	2
Accountability	#18–20	3

Figure 5. ECCOLA card themes. (Vakkuri, Kemell, Jantunen, Halme, Abrahamsson 2021)

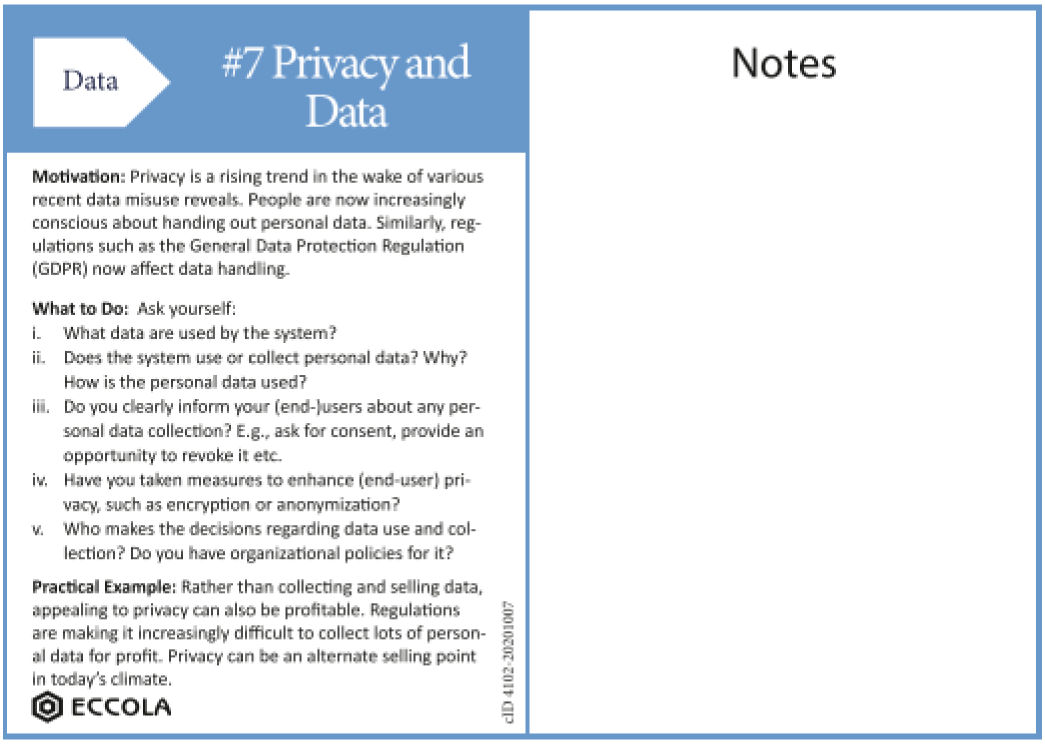
All Cards consist of three parts:

1. Motivation (here it should be said, why this topic is important)
2. What to do (to tackle this issue)
3. Practical example

Cards themselves are:

0. Stakeholder Analysis (understand who and how could the AI system affect)
1. Types of Transparency (understand who you are being transparent towards and what you are being transparent about, or who are being transparent towards you and about what)
2. Explainability (without understanding the reasons behind the actions of an AI system an act of a regulative guidelines is violated)
3. Communication (clear stating about the fact what data does AI machine collect and how it’s being done)
4. Documenting Trade-offs (whenever you make a decision, it is important to highlight what were the alternatives and what good points do they have)
5. Traceability (always document all the actions starting with code and ending with project actions)

6. System Reliability (working with errors and making actions to reduce amount of them)
7. Privacy and Data (clear informing users if a system is using their personal data and encrypt it)
8. Data Quality (data should be good in terms of quality and integrity)
9. Access to Data (important to know what exactly type of data is used, by whom, why and how)
10. Human Agency (users should know exactly what they want from AI system and how to achieve that)
11. Human Oversight (AI systems should support human's decisions but not make decisions for people)
12. System Security (all cybersecurity actions)
13. System Safety (what king of risk does the system involve)
14. Accessibility (ensuring that AI system usage does not discriminate people in any way and equal access is provided)
15. Stakeholder Participation (users should have a basic understanding of the AI system content)
16. Environmental Impact (ensuring that AI system is sustainable)
17. Societal Effects (ensuring that system won't affect negatively on those who didn't even use it)
18. Auditability (it is good to have mechanisms for audit before it would be needed)
19. Ability to Redress (users should know how they can be compensated if something would go in a wrong way)
20. Minimizing Negative Impacts (reducing risks of any incidents)



Data

#7 Privacy and Data

Motivation: Privacy is a rising trend in the wake of various recent data misuse reveals. People are now increasingly conscious about handing out personal data. Similarly, regulations such as the General Data Protection Regulation (GDPR) now affect data handling.

What to Do: Ask yourself:

i. What data are used by the system?
ii. Does the system use or collect personal data? Why? How is the personal data used?
iii. Do you clearly inform your (end-)users about any personal data collection? E.g., ask for consent, provide an opportunity to revoke it etc.
iv. Have you taken measures to enhance (end-user) privacy, such as encryption or anonymization?
v. Who makes the decisions regarding data use and collection? Do you have organizational policies for it?

Practical Example: Rather than collecting and selling data, appealing to privacy can also be profitable. Regulations are making it increasingly difficult to collect lots of personal data for profit. Privacy can be an alternate selling point in today's climate.

ECCOLA

cID 4102-20201007

Notes

Figure 6. An example of card used in ECCOLA method. (Vakkuri, Kemell, Jantunen, Halme, Abrahamsson 2021)

To use ECCOLA in practice, one should use sprint method. After each sprint an iterative process, which consists of analysing and evaluation, should be started. Of course, there is no need to iterative check on every card, but you can take only cards, which would change due to the action. (Vakkuri, Kemell, Jantunen, Halme, Abrahamsson 2021)

8 Survey about AI Regulations analysis

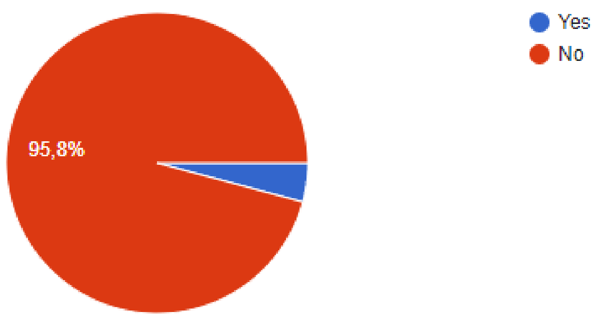
As a practical part of a thesis was decided to create a survey, the results of which would be included in the final analysis and would show people's opinions on AI regulations and AI technology in overall. Respondents consist of people with different background (both professionals in AI industry and people who do not interact with AI at all). The only similarity between all of them is their place of living (Finland).

For representation of professional opinion participants of FCAI (Finnish Center for Artificial Intelligence) were contacted and asked to answer the survey. Besides, students of Haaga-Helia on different programs (mainly Business Information Technology) have participated in a survey as well. There are 24 answers in total. Only 1 person have worked in AI Regulation area before. Besides, 50% of respondents do have some connections with AI industry (Business Information Technology program students, IT experts) and another 50% do not.

Have you ever worked in AI regulations area?

24 ответа

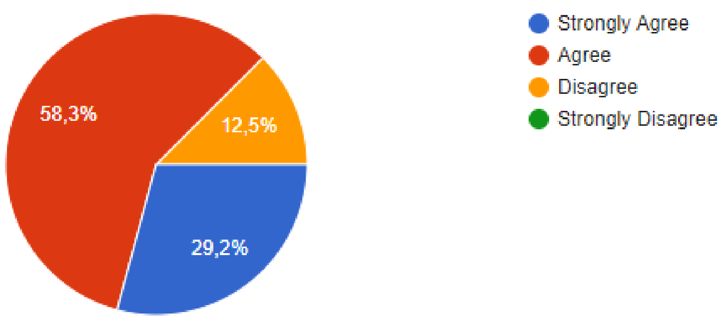
Копировать



Using robots to automate many jobs, humans have done in the past, is a good idea

24 ответа

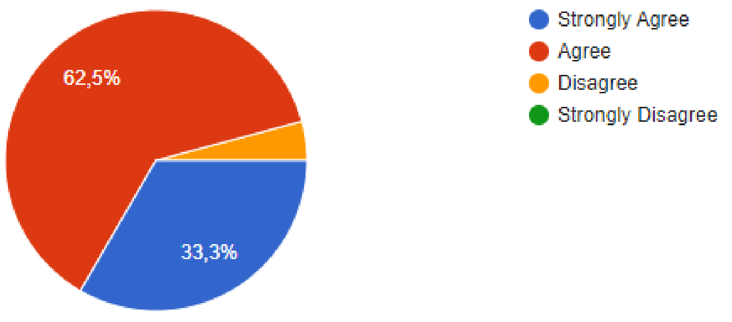
Копировать



Nowadays usual people do not have enough knowledge about AI technology in overall

Копировать

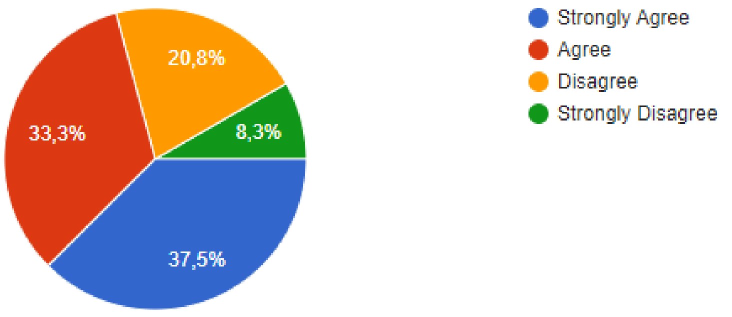
24 ответа



We should create a worldwide organization responsible for AI regulation.

Копировать

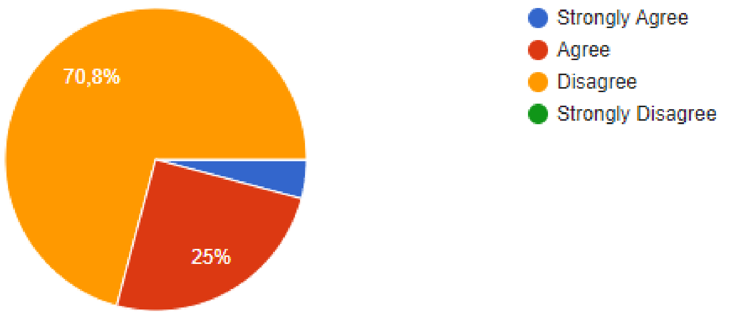
24 ответа



Amount of work being done in AI regulations area in Europe is enough

Копировать

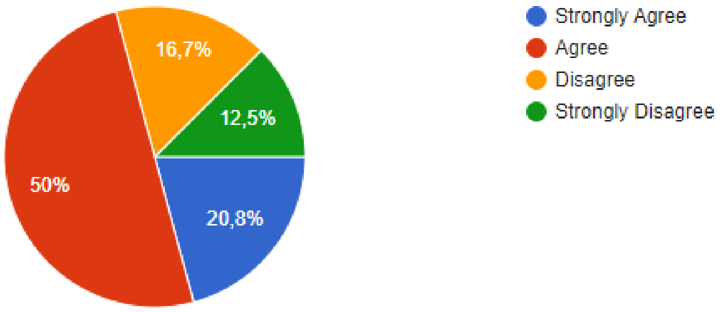
24 ответа



If an AI system has created a painting, it still can be considered as a piece of art

Копировать

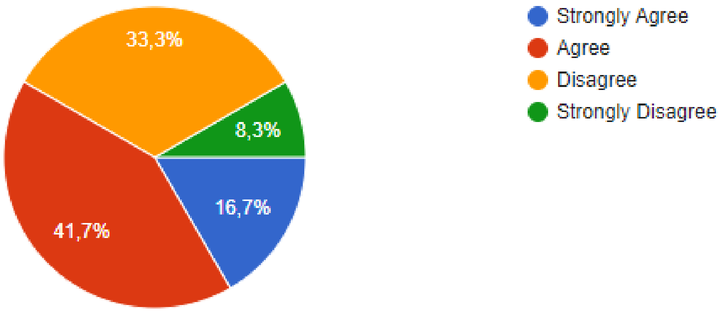
24 ответа



A creator of an AI system is the one who is responsible for decisions made by this system

Копировать

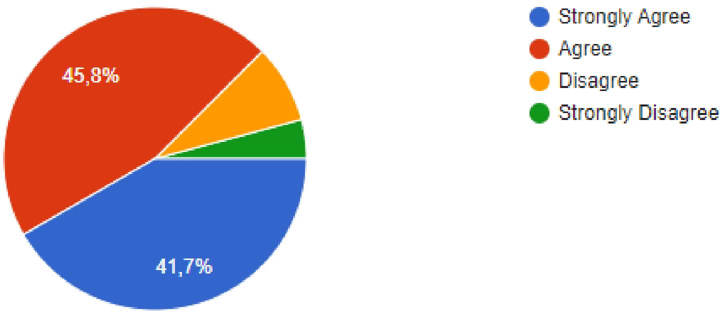
24 ответа



It is unlikely, that AI systems will take the dominance over people in the near future

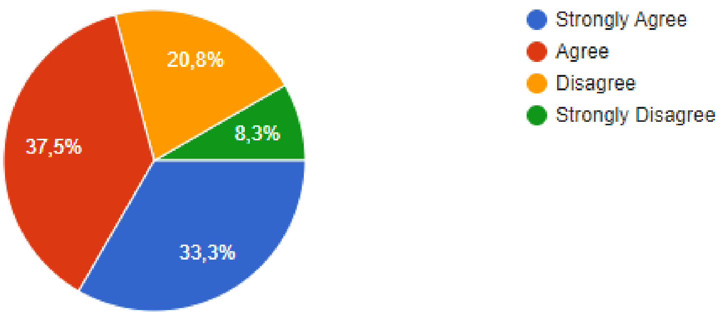
Копировать

24 ответа



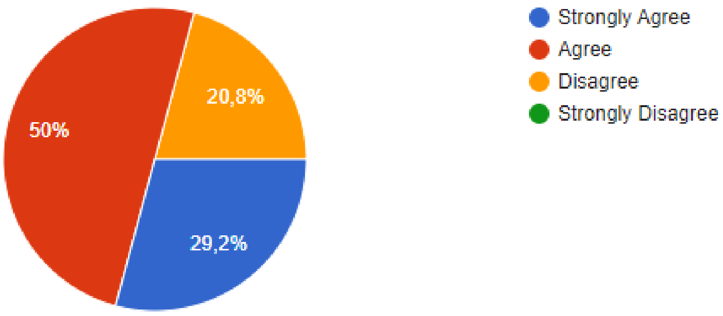
There are some AI systems, which should never be used by people and their usage should be marked as 'unacceptable risk'

24 ответа



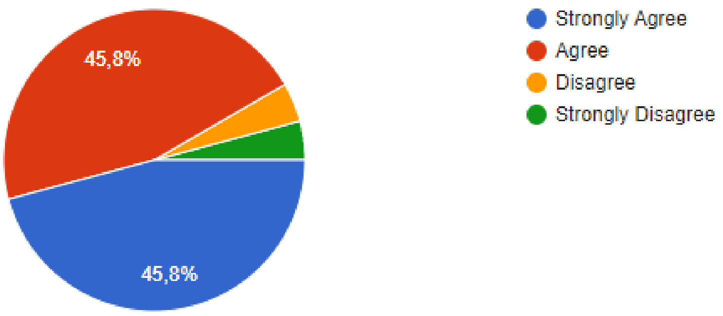
There should be more resources being put in the AI regulations area

24 ответа



We should create a set of rules in AI area, which would be followed worldwide

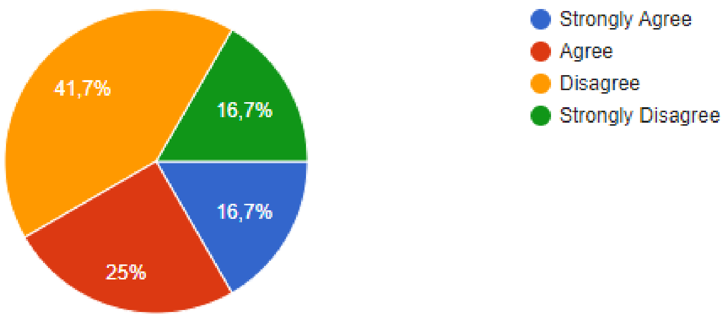
24 ответа



AI development will cause the problem of unemployment in the future

Копировать

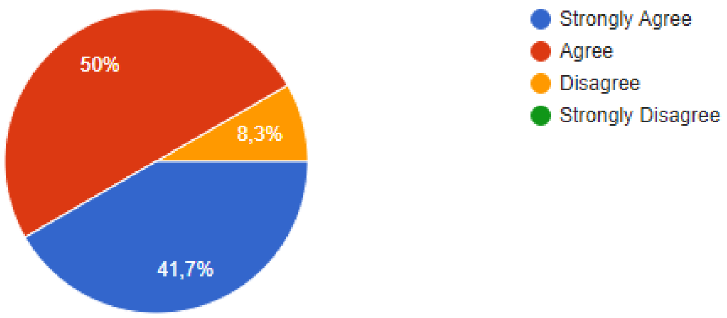
24 ответа



We should create more university programs about AI

Копировать

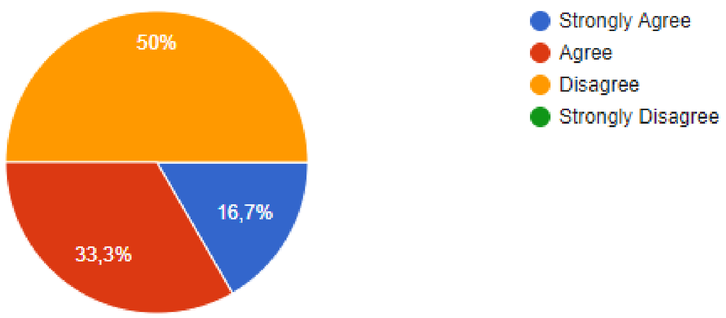
24 ответа



We should use AI systems in vital areas (such as court of law)

Копировать

24 ответа



Most people agree with the need of AI developing and respondents believe, that humanity will benefit a lot from AI development in the future. Moreover, people think, that AI industry will be and already is one of the most important industries to take care of. Besides, 95% of respondents consider that knowledge about AI topic of an average human is not on the sufficient level. 92% of respondents also agree with creating of more university programs about AI industry. Good to note, that 2 respondents, who gave negative opinions on this question were IT experts, so one of the reasons for them to give such an answer, could be satisfaction with the current number of programs. At the same time, this question contains one of the highest percentages of strong agreement option chosen of all questions in the survey.

Speaking of regulative aspect, answers were more balanced. 71% of respondents think, we should create a worldwide organization responsible for AI regulation (and 91% agree with the need of creating of a set of rules for be followed worldwide in AI industry).

In overall, respondents also feel, that the amount of work being done in AI regulations area is not enough. 71% consider increasing the amount of work in this area. It is also notable, that this question does not have strong answers (strong agreement or disagreement) which means, that the topic is pretty difficult, and respondents do not have a solid confidence in their words, as half of them do not really know what work is exactly done in Europe regarding AI technology.

One more controversial question was a question, if a painting created by AI can be considered as a piece of art. 71% agree with this statement. An interesting point is that 29% of respondents who disagreed with the statement mostly were usual people, but not IT experts. A conclusion can be made, that with the development of AI, even the basic things such as culture and art definitions do not stand still and are developing as well.

On the contrary, IT experts mostly do not agree with the statement, that a creator of an AI system is the one who is responsible for decisions made by this system. At the same time 59% of respondents agree with this statement. (Most of them were not IT experts).

What is similar between 2 groups is their opinion on possible AI domination in the future. Only 12% do think, that AI system will likely take the dominance over people soon and all of them are not from IT experts' group. Of course, this worry does have a strong reasoning behind itself, but from the results it can be seen, that for now, this fear is still irrational.

50% of all respondents think, that we shouldn't use AI systems in vital areas (such as court of law) yet. Notable, that this percentage consists of both IT experts and usual people. It is interesting also, that this question is 50/50 balanced with negative and positive answers. I believe that this balance is reached because of a definition of a vital area. It is a good idea to reconstruct this question using concrete examples and ask about them exactly. Using this approach will allow us to understand the situation better.

As a conclusion to the survey at whole, it can be said, that people accept AI industry and consider it as an important part of life in today’s world. They also support investment of different resources in AI development and regulations. Though, AI is still not an essential part of all the aspects of life, it is likely, that it will change soon, and this industry will be the dominant one.

9 Discussion

As during this thesis work a survey was created and used in a final version, thesis is research-typed and contains quantitative research. As a result, it is required to ensure research reliability through the concepts of validity and reliability. According to Dr. Roberta Heale (2015), to ensure validity of a quantitative study, an author should ensure, that survey's research object corresponds its content. Approach in this thesis is validated, as the topic of a survey ('survey about AI technology') is wide and matches its content.

Besides, there is a need to ensure thesis' reliability. According to Dr. Roberta Heale (2015), it can be done by ensuring, that a participant completing the survey might have approximately the same responses each time the test is completed. As was said above, the topic of the survey itself is wide, so answers are strongly divers and often opposite. Respondents were gathered by using personal emails and messages in group chats, which means that the traceability is ensured, and it is easy to find each respondent of 24 in total.

During thesis work, a survey about AI development was created and analyzed. Besides, all theoretical information was gathered and presented as a set of rules to consider while working in AI industry. Moreover, a deep analysis of current AI regulations state in different countries was done. During that analysis, approaches were compared to each other and analyzed at whole.

Speaking of objectives, all of them were successfully reached. A current state of moral and ethical AI regulations worldwide was deeply researched and reflected. Also, an analysis on how and where AI systems are being used was done as well.

As a conclusion to this thesis, it can be said, that more and more people currently see AI industry as one that needs strong regulations and will become the main industry in the world in the near future. To be added, governments of different countries all over the world have already done a huge amount of work regarding AI regulations and nevertheless are planning to increase it.

One of the main difficulties while working on this thesis was gathering feedback for a survey about AI development. Though, this survey was sent to FCAI, the feedback from them consists only of 2 answers. As a result, a decision to expand this survey on to Haaga-Helia students from different programs as well as just Finnish citizens, was made.

What results of survey mean in practice, is that the level of understanding of AI technology in overall is pretty low and we should still gain more knowledge about the topic and spread it onto other people not to be unprepared to AI arise soon.

One of the main actions that might be took differently is a practical part of thesis. It might be a good idea to choose more qualitative approach and make an interview with an AI professional during which we would have discussed many interesting things. Besides, there is a room for improvement in the survey about AI development. Some questions were difficult to answer for people who are not interested in AI industry and do not have any information about work being done in this industry in the world. So, it might be a good idea to distribute such questions only over AI experts or reconstruct them fully.

Sources

Brown, J. 2020. Growth in Artificial Intelligence Is Beyond Exponential. Legacy Research Group. URL: https://www.legacyresearch.com/the-daily-cut/growth-in-artificial-intelligence-is-beyond-exponential/. Accessed: 02 October 2022.

Burns, E. 2021. Machine Learning. TechTarget. URL: https://www.techtarget.com/searchenterpriseai/definition/machine-learning-ML. Accessed: 10 October 2022.

Collins, D. 2022. UK sets out proposals for new AI rulebook to unleash innovation and boost public trust in the technology. Gov.UK. URL: https://www.gov.uk/government/news/uk-sets-out-proposals-for-new-ai-rulebook-to-unleash-innovation-and-boost-public-trust-in-the-technology. Accessed: 17 October 2022.

Copeland, B. J. S.a. Alan Turing and the beginning of AI. Britannica. URL: https://www.britannica.com/technology/artificial-intelligence/The-Turing-test. Accessed: 27 September 2022.

Dilmegani, C. 2022. Top 9 ethical dilemmas of AI and how to navigate them. AIMutiple. URL: https://research.aimultiple.com/ai-ethics/. Accessed: 31 October 2022.

Ebers, M., Navas, S. 2019. Regulating AI and Robotics: Ethical and Legal Challenges. URL: https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3392379. Accessed: 15 October 2022.

Eliaçik, E. 2022. Can the cons of AI outweigh its benefits? Dataconomy. URL: https://dataconomy.com/2022/08/the-cons-of-artificial-intelligence/. Accessed: 10 October 2022.

European Commission. 2022. Executive Summary of the Impact Assessment Report. URL: https://ec.europa.eu/info/sites/default/files/1_5_197611_summary_impact_asse_dir_ai_en.pdf. Accessed: 15 October 2022.

European Commission. 2022. Regulatory framework proposal on artificial intelligence. URL: https://digital-strategy.ec.europa.eu/en/policies/regulatory-framework-ai. Accessed: 11 October 2022.

European Commission. 2022. Shaping Europe’s digital future. URL: https://digital-strategy.ec.europa.eu/en/policies/european-approach-artificial-intelligence. Accessed: 11 October 2022.

Galaski, J. 2021. Liberties. AI Regulation: Present Situation and Future Possibilities. URL: https://www.liberties.eu/en/stories/ai-regulation/43740. Accessed: 19 October 2022.

Heale, R., Twycross, A. 2015. Validity and reliability in quantitative studies. CrossMark. URL: https://ebn.bmj.com/content/ebnurs/18/3/66.full.pdf. Accessed: 24 November 2022.

Howell, Chanley T. 2022. AI Regulation: Where do China, the EU, and the U.S. stand today? The National Law Review. URL: https://www.natlawreview.com/article/ai-regulation-where-do-china-eu-and-us-stand-today. Accessed: 17 October 2022.

IBM Cloud Education. 2020. Supervised Learning. URL: https://www.ibm.com/cloud/learn/supervised-learning. Accessed: 10 October 2022.

IBM Cloud Education. 2020. Unsupervised Learning. URL: https://www.ibm.com/cloud/learn/unsupervised-learning. Accessed: 10 October 2022.

Isbert, M., Koering, D. 2020. Rules for dealing with artificial intelligence in companies. European Commission. URL: https://futurium.ec.europa.eu/en/european-ai-alliance/best-practices/rules-dealing-artificial-intelligence-companies?language=en. Accessed: 24 October 2022.

Johnson, C., Tyson, A. 2020. Here's how opinions on the impact of artificial intelligence differ around the world. World Economic Forum. URL: https://www.weforum.org/agenda/2020/12/mixed-views-of-the-impact-of-artificial-intelligence. Accessed: 05 October 2022.

Krishnakumar, M. 2022. The 6 Autonomous Driving Levels Explained. Weights & Biases. URL: https://wandb.ai/mukilan/autonomous-cars/reports/The-6-Levels-of-Autonomous-Cars--VmlldzoyNTcwOTQ1. Accessed: 02 October 2022.

Kumar, S. 2019. Advantages and Disadvantages of Artificial Intelligence. Towards Data Science. URL: https://towardsdatascience.com/advantages-and-disadvantages-of-artificial-intelligence-182a5ef6588c. Accessed: 05 October 2022.

OpenAI. 2019. AI and Compute. URL: https://openai.com/blog/ai-and-compute/#:~:text=A%20petaflop%2Fs%2Dday%20(,to%20kW%2Dhr%20for%20energy. Accessed: 04 October 2022.

UNESCO. S.a. Artificial Intelligence: examples of ethical dilemmas. URL: https://en.unesco.org/artificial-intelligence/ethics/cases. Accessed: 31 October 2022.

Vakkuri, V., Kemell, K-K., Jantunen, M., Halme, E., Abrahamsson, P. 2022. ECCOLA – A method for implementing ethically aligned AI systems. ScienceDirect. URL: https://www.sciencedirect.com/science/article/pii/S0164121221001643. Accessed: 20 November 2022.

Wu, Y. 2022. AI in China: Regulations, Market Opportunities Challenges for Investors. China Briefing. URL: https://www.china-briefing.com/news/ai-in-china-regulatory-updates-investment-opportunities-and-challenges/. Accessed: 17 October, 2022.

Zhang, B. 2019. Public opinion lessons for AI regulation. Brookings. URL: https://www.brookings.edu/research/public-opinion-lessons-for-ai-regulation/. Accessed: 19 October 2022.

Zhang, B., Dafoe, A. 2019. Artificial Intelligence: American Attitudes and Trends. SSRN. URL: https://deliverypdf.ssrn.com/delivery.php?ID=332092117088002108096109116089108121026011068081017086005005103078092114078094110108063055059052058025062074030079023095007092041036013081092090065013123099119120065058087060028095124126116120064100091002094029119107120095024001024108126030127103000099&EXT=pdf&INDEX=TRUE. Accessed: 20 October 2022.

Appendices

Appendix 1. Survey about AI technology regulation

AI Regulations

You need to pick, how do you feel about each statement. Pick the answer, that is the closest to your opinion please.

Kirjaudu Googleen, jotta voit tallentaa edistymisesi. Lue lisää

*Pakollinen

Have you ever worked in AI regulations area? *

- Yes
- No

Using robots to automate many jobs, humans have done in the past, is a good idea *

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

Nowadays usual people do not have enough knowledge about AI technology in overall *

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

We should create a worldwide organization responsible for AI regulation. *

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

Amount of work being done in AI regulations area in Europe is enough *

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

If an AI system has created a painting, it still can be considered as a piece of art *

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

A creator of an AI system is the one who is responsible for decisions made by this system *

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

It is unlikely, that AI systems will take the dominance over people in the near future *

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

There are some AI systems, which should never be used by people and their usage should be marked as 'unacceptable risk' *

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

There should be more resources being put in the AI regulations area *

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

We should create a set of rules in AI area, which would be followed worldwide *

- Strongly Agree
- Agree
- Disagree
- Strongly Disagree

AI development will cause the problem of unemployment in the future *

- ○ Strongly Agree
- ○ Agree
- ○ Disagree
- ○ Strongly Disagree

We should create more university programs about AI *

- ○ Strongly Agree
- ○ Agree
- ○ Disagree
- ○ Strongly Disagree

We should use AI systems in vital areas (such as court of law) *

- ○ Strongly Agree
- ○ Agree
- ○ Disagree
- ○ Strongly Disagree