ON THE ROAD TO HUMAN-SCALE ARTIFICIAL INTELLIGENCE, AS SEEN BY OUR EXPERTS

AZ EMBERI LÉPTÉKŰ MESTERSÉGES INTELLIGENCIA FELÉ VEZETŐ ÚTON, AHOGY AZT SZAKÉRTŐINK LÁTJÁK[1,2]

HEITLERNÉ LEHOCZKY, Mária[3] – KOLLÁR, Csaba[4]

Abstract

In our study, we will first present some of the definitions of AI, noting that there is no single definition that is universally accepted, but it is the many definitions that highlight the many readings and uses of AI. In a separate section, we discuss AI on a human scale, which could also be a synonym for ethical AI, since we are convinced that only an ethical approach and its international acceptance by governments, industry and society can guarantee that the human focus will be maintained in the development of technology. In the empirical part of our study, we present the results of two expert studies conducted in the Artificial Intelligence Workshop at Óbuda University, and after presenting and comparatively analysing the opinions of the two groups of experts, we conclude with a short summary.

Keywords

artificial intelligence, ethical AI, foresight, expert interviewing

Absztrakt

Tanulmányunkban először a mesterséges intelligenciával kapcsolatos definíciók közül mutatunk be néhányat azzal a megjegyzéssel, hogy egységes, mindenki által elfogadott definíció nincs, de pont a sokféle definíció világít rá arra, hogy a mesterséges intelligenciának mennyi olvasata és mennyi felhasználási területe van. Külön részben foglalkozunk az emberi léptékű mesterséges intelligenciával, amelyik alapvetően az etikus mesterséges intelligencia szinonimája is lehet, hiszen meggyőződésünk, hogy csak az etikus megközelítés és annak nemzetközi kormányzati, gazdasági és társadalmi elfogadottsága a garancia arra, hogy a technika fejlődésében a humán fókusz megmaradjon. Tanulmányunk empirikus részében az Óbudai Egyetemen működő Mesterséges Intelligencia Műhelyben végzett két szakértői kutatásunk eredményét ismertetjük, majd a két szakértői csoport véleményének bemutatása és komparatív elemzése után egy rövid összefoglalással zárjuk írásművünket.

Kulcsszavak

mesterséges intelligencia, etikus MI, jövőkutatás, szakértői megkérdezés

[1] The present study is based on a detailed research report published in the journal Safety and Security Sciences Review, vol 4, No 1 and 2.

[2] This study was written by humans in Hungarian, translated into English by artificial intelligence (DeepL and Google Translate), and checked for correctness by a human.

[3] lehoczky.maria@uni-obuda.hu | ORCID: 0000-0003-0588-715X | PhD student, Óbuda University Doctoral School for Safety and Security Sciences | member, Óbuda University Bánki Donát Faculty of Mechanical and Safety Engineering Artificial Intelligence Workshop | doktorandusz, Óbudai Egyetem Biztoságtudományi Doktori Iskola | tag, Óbudai Egyetem Bánki Donát Gépész és Biztonságtechnikai Mérnöki Kar Mesterséges Intelligencia Műhely

[4] kollar.csaba@uni-obuda.hu | ORCID: 0000-0002-0981-2385 | senior research fellow and leader, Óbuda University Bánki Donát Faculty of Mechanical and Safety Engineering Artificial Intelligence Workshop | tudományos főmunkatárs és vezető, Óbudai Egyetem Bánki Donát Gépész és Biztonságtechnikai Mérnöki Kar Mesterséges Intelligencia Műhely

DEFINITIONS OF ARTIFICIAL INTELLIGENCE

It is difficult to clearly define AI due to diversity of Artificial Intelligence (AI) problems, solutions, distinction of what AI contains and what not. The shortest and simplest definition is: "AI is not biological intelligence" [1]. AI is one of the most life-changing scientific and technological development of the century. There is no universally accepted definition of AI, it is an umbrella term. AI is a science and computational technology that is inspired by the way how people use their nervous system (acquire, store, manipulate, transmit information, learn, reason) and bodies to learn, reason, take actions.

Encyclopedia Britannica defines [2] as "the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings. The term is frequently applied to the project of developing systems endowed with the intellectual processes characteristic of human beings, such as the ability to reason, discover meaning, generalize, or learn from past experience". The Cambridge Academic Content Dictionary [3] compresses in a shorter version: "the use of computer programs that have some of the qualities of the human mind, such as the ability to understand language, recognize pictures, and learn from experience".

Artificial Intelligence (AI), a term coined by John McCarthy in 1955, was defined [4] in 2007 as "the science and engineering of making intelligent machines especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable".

After 1955, several definitions of AI have emerged. In their popular book Russel and Norvig [5] developed a new taxonomy. They examined the definitions of AI according to two dimensions. On one hand, they separate the human-focused conception of the empirical sciences from the rational-focused conception of mathematics and engineering. On the other hand, thought processes, inference, are separated from behavior. Based on the combination of the two aspects, four groups can be formed.

Human approach

(1) systems that think like humans (e.g., cognitive architectures and neural networks);
Thinking Humanly: there is no comprehensive theory of mind yet, but the ultimate goal for the system is to function in a manner similar to human thinking. The interdisciplinary field of cognitive science brings together computer models from AI and experimental techniques from psychology to construct precise and testable theories of the human mind. If the program's input–output behavior matches corresponding human behavior, that is evidence that some of the program's mechanisms could also be operating in humans.

(2) systems that act like humans (e.g., pass the Turing)
Acting Humanly: Turing [6] proposed a test called "The Imitation Game": (Turing-test), which is an operational definition of intelligence. A computer passes the test if a human interrogator, after posing some written questions, cannot tell whether the written responses

come from a person or from a computer. The computer would require to possess following capabilities:

- natural language processing to communicate successfully in a human language;
- knowledge representation to store what it knows or hears;
- automated reasoning to answer questions and to draw new conclusions;
- machine learning to adapt to new circumstances and to detect and extrapolate patterns
- computer vision and speech recognition to perceive the world;
- robotics to manipulate objects and move about.

Ideal approach

(3) systems that think rationally (e.g., logic solvers, inference, and optimization);
Thinking Rationally: a system is rational if it does the "right thing", given what it knows, based on irrefutable reasoning process. The logicist tradition within artificial intelligence hopes to build on such programs to create intelligent systems.

(4) systems that act rationally (e.g., intelligent software agents and embodied robots that achieve goals via perception, planning, reasoning, learning, communicating, decision-making, and acting).
Acting Rationally: computer agents are expected to operate autonomously, perceive their environment, persist over a prolonged time period, adapt to change, create and pursue goals. A rational agent is one that acts so as to achieve the best outcome or, when there is uncertainty, the best expected outcome. All the skills needed for the Turing Test also allow an agent to act rationally. important fact: perfect rationality – always taking the exactly optimal action – is not feasible in complex environments, because the computing needs are too high. The history of AI is dominated by study and construction of the rational agent approach and Russel and Norvig [5] define as the study of agents that receive percepts from the environment and perform actions.

The definition, given by Nilson [7], includes a broad interpretation of the concept of intelligence, not only humans, but animals and some machines are intelligent to variant degrees: „AI is that activity devoted to making machines intelligent, and intelligence is that quality that enables an entity to function appropriately and with foresight in its environment."

The definition of AI also differs in documents issued by international organizations.

- European Comission [8]: "AI is a collection of technologies that combine data, algorithms and computing power".
- OECD [9] : "An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments." "AI systems are designed to operate with varying levels of autonomy."
- UNESCO [10]: AI is an ensemble of advanced ICTs that enable "machines capable of imitating certain functionalities of human intelligence, including such features

as perception, learning, reasoning, problem solving, language interaction, and even producing creative work"

- Council of Europe [11]: "Artificial intelligence (AI) systems are software (and possibly also hardware) systems designed by humans that, given a complex goal, act in the physical or digital dimension by perceiving their environment through data acquisition, interpreting the collected structured or unstructured data, reasoning on the knowledge, or processing the information, derived from this data and deciding the best action(s) to take to achieve the given goal. AI systems can either use symbolic rules or learn a numeric model, and they can also adapt their behavior by analysing how the environment is affected by their previous actions."

In general terms, AI refers to a broad field of science encompassing not only computer science, if we construe AI as studying how information is acquired, processed, stored, used, etc. in intelligent animals and machines then it obviously overlaps with several older disciplines [12]:

- Philosophy
- Mathematics and statistics
- Economics
- Neuroscience
- Psychology
- Biology and medical science
- Linguistics
- Computer sciences
- Technical sciences
- Safety and security sciences

Nowadays AI trained and focused to perform specific tasks (playing strategic games, language translation, self-driving vehicles, and image recognition, trip planning etc.). This development level of AI called [13] Artificial Narrow Intelligence (ANI) or Weak AI [14]. The next level is General AI (Artificial General Intelligence, or AGI) [13] or strong AI [14] refers to a future, theoretical form of AI system that exhibits apparently intelligent behavior at least as advanced as a person across the full range of cognitive tasks. Artificial Super Intelligence (ASI) [13] – also known as superintelligence – would surpass the intelligence and ability of the human brain.

STEPS TOWARDS THE REALIZATION OF HUMAN-CENTERED MI

AI will radically transform the world, we can already encounter such worrying phenomena, e.g. ethical concerns that may project a dystopian vision. Humanity is facing an existential challenge whose awareness and active struggle can bring about positive change, where cooperation between machines and people results in an utopian world. Tilesch and Hatamleh [15] are urging the development of a new paradigm in which humanity will define its vision, the institutional systems of AI. It is important to keep in mind that it does not serve the interests of a narrow stratum (the profit-oriented, amoral, manipulative use of MI) but places the public good above individual interests. Individual awareness of AI-related

changes is considered necessary. This includes making digital citizenship an integral part of everyday life, encouraging social dialogue about AI. It is considered essential to maintain authentic and quality media and to restore social trust. Three steps [15] are identified for the future implementation of human-faced AI:

1. planning: developing a globally accepted humanistic and actionable vision, harmonizing divergent interests into a normative and regulatory framework. (ethics, credibility, reliability) Regulation based on continuous feedback through impartial, fact-based supervisory institutions.
2. development, dissemination: creation of institutions dealing with research and educational activities in order to solve systemic problems with a multidisciplinary approach, in the form of open research, in accordance with ethical standards. They ensure the global dissemination of AI knowledge through their educational activities.
3. the transformation of humanity's image of itself, the core of a new ideology is conscience and awareness.

We are currently in the planning stage. The OECD Recommendation [16] makes it clear that the role of artificial intelligence can be the key to shaping the future positively, promoting people's welfare as well as subjective well-being, contributing to economic development and the achievement of sustainable goals. All this is accompanied by profound social changes. Five principles have been laid down for the present and the near future can promote an AI-powered crisis management that is trustworthy and respects human-centered and democratic values:

1. Inclusive growth, sustainable development and well-being: AI should help the population global prosper by promoting inclusive growth, sustainable development and prosperity
2. Human-centered values and fairness: AI systems must be designed to comply with legal requirements, human rights and democratic values. They should consider adequate safeguards (such as the possibility of human intervention) for a fair and just society.
3. Transparency and explainability: AI systems need to be transparent, with information disclosed responsibly so that people can understand and challenge MI-based decisions
4. Robustness, security and safety: AI systems must operate in a robust, safe and secure manner throughout their lifetime, and potential hazards must be continuously assessed and managed.
5. Accountability: Organizations or individuals developing, installing and operating AI systems must remain accountable for the proper operation of the systems, in accordance with the above guidelines.

The document also makes recommendations to governments:

1. Investing in AI research and development: Promote public and private investment in R&D by encouraging innovation in reliable AI systems. AI systems must respect privacy as well as data protection and should be free of inappropriate biases.

2. Fostering a digital ecosystem for AI: In order to have a trustworthy AI, governments need to support digital technologies, infrastructures, mechanism of knowledge sharing about AI, i.e. the AI ecosystem.
3. Shaping an enabling policy environment for AI: A regulatory environment needs to be created that paves the way for the deployment and operation of reliable MI systems.
4. Building human capacity and preparing for labor market transformation: People need to be equipped with artificial intelligence skills and all support should be given to workers to ensure a fair transition.
5. International co-operation for trustworthy AI: Cross-border and market-sector co-operation is needed to promote responsible care for reliable MI technologies by governments.

DESCRIPTION OF THE RESEARCH METHOD

Focus group interviewing is a qualitative research method and can be found in sociology, and more broadly in the social sciences, as well as in marketing, advertising, market research, public opinion research and psychology. In the development of the methodology used in this research, and in the processing and evaluation of the results as textual content, the social science foundations were laid by the writings of Cseh-Szombathy and Ferge [17], Earl [18], Krippendorff [19], Horváth and Mitev [20], Gordon and Langmaid [21], Langer [22]. Although our research was not marketing-oriented, we drew on the works of Malhotra [23], Scipione [24], Kollár [25] Vicsek [26] in organizing the focus group and conducting the focus group discussions. In the selection of the members of the two groups (senior and junior), as will be described in detail later, we were considerably more rigorous than the hascholarly selection methods and procedures in order to ensure that the members of the senior group met the requirements of expert selection [27] in all respects, while the junior experts were selected less rigorously, but with the aim of ensuring that the members of the group were found to have a verifiably and factually deeper more thorough knowledge of AI and robotics than the average person of interest, due to their technical and IT knowledge.

The general characteristics of the two groups of experts and the methodological criteria we use are described below. The members of the groups have a variety of self-constructed perceptions of reality, which they share with each other and which they shape and form in themselves through the questions they ask and their interactions with each other. Measurement, quantification of data, is not important except for the very simple demographic group description. Our sampling aims to get a deeper understanding of the experts' views and opinions, and to this end we created an atmosphere both online and in the physical world where our experts could talk informally with each other and with us. The questions we asked in the opening and tuning-in sections also served to create a relaxed, friendly atmosphere. Our analysis was based on "contextualized descriptions and understandings of phenomena" [20: 35 p.].

Following Malhotra [23: 206 p.], we carried out our research according to the following design and implementation steps:

1. Setting objectives and defining the research problem: Artificial Intelligence has been at the center of discourse in almost all fields of science over the last 10-15 years. The results of numerous questionnaire surveys and experiments on artificial intelligence and robotics can be found in scientific journals, and books by theoretical authors, either individually or co-authored, provide a wealth of knowledge for those interested in the subject. However, we felt that there was a lack of recent qualitative research conducted in Hungary, in which experts with a much richer and more profound knowledge than the average share their views on artificial intelligence with each other and, through our study, with the readers, their views on the future of AI, their own role and that of society, what challenges it will face in ten years' time, and whether they see AI as a risk or an opportunity.
2. To define the objectives and methodology of the qualitative research: to explore the current and future opportunities and risks of AI, using a rigorous focus group survey of senior and junior experts, and to identify similarities and differences between the views and positions of the two groups. The comparison of senior and junior opinions also gave us the opportunity to examine the perception of AI between generations (see Part 2 of our paper).
3. Identifying the questions to be answered by the focus groups: for details, see the subsection "Questions and areas covered by each expert group".
4. The screening questionnaire: the screening questionnaire in this study was used to screen the experts [27], i.e. to determine the expectations of the group members who were selected for the senior and junior groups.
5. Construction of the interview guide for the moderators: the interview guide for each group was constructed based on (1) the opening, (2) the tuning, (3) the main questions, and (4) the questions of the interview guide. The opening round and the tuning-in, as described above, contributed to creating a relaxed atmosphere, while the main questions addressed the participants' perceptions of the emerging fields of AI, its current and future potentials and limitations, and its impact on them as experts and on the fields they represent/are familiar with.
6. Conduct of the focus group discussions: see Table 1.
7. Data analysis: the data were analysed separately for the two expert groups and are presented in the subsections "Results of the research" for each expert group in our study.
8. Summary of results: our research results are reported separately for each expert group.
9. Comparative comparison and summary of results: we concluded our study by comparing the results obtained for the two groups.

The general characteristics of the focus group expert discussions are listed in Table 1, based on Gordon and Langmaid [21: pp. 57-58].

Description	Senior Expert Group	Junior Expert Group
Date of the expert discussion	Monday, 21 February 2022, 18:00-20:05	23 February 2022, Wednesday, 10:00-12:15
Location of the expert discussion	Online (Zoom)	Óbuda University, Bánki Donát Faculty of Mechanical and Safety Engineering, Room 306
Duration of the expert interview	125 min	135 min
Demographic characteristics of the participants	1 woman, 7 men min. age: 28 years max. age: 61 years average age: 43 years professionals with tertiary education	3 women, 2 men min. age: 20 years max. age: 28 years average age: 24 years University students (safety engineer)
Number of participants	8 persons	5 persons
Seating arrangements in the room	Participants were able to see each other and the moderators on a monitor, thanks to the interface provided by the Zoom application	On either side of the table, facing each other
Location of respondents and moderators		The two moderators sat on the shorter side of the table
Identity of moderators	Experts in qualitative and quantitative research in online and offline environments are Mária HEITLERNÉ LEHOCZKY (psychologist-economist) and Csaba KOLLÁR (engineer, humanities, PhD in economics).	
Observers, other participants	In order to keep all selected and invited participants active throughout the discussion, we decided to exclude passive participants and observers from the discussion.	
Technical equipment, IT background	Zoom application, not recorded at the request of the participants	Dictaphone, participants contributed to the audio recording

1. Table: General characteristics of focus group interviews with experts [5], own ed.

Introduction of the Senior Expert Group

In selecting senior experts, we followed Kollár's [27] recommendations as follows. The following statements apply to the experts we invited to participate in the online focus group discussion:

- have a higher education qualification
- is a member of at least one professional and scientific organization:
  - Óbuda University Artificial Intelligence Workshop
  - Artificial Intelligence Coalition
  - Communication and Information Science Association Artificial Intelligence Section

    - National Association of Human Resources Professionals
    - Hungarian Military Science Society Electronics, Informatics and Robotics Section
    - Hungarian Academy of Sciences Public Board
- in areas related to artificial intelligence, digital society, human-robot interaction, industry 4.0
    - have at least 3 papers published in peer-reviewed scientific journals
    - at least 3 presentations at scientific and professional conferences

In classic focus group interviews, experts are named in the research reports and the resulting studies. The experts we asked did not contribute to this, as many of them work in a public or governmental sector where they would have had to seek permission from the institution before making a statement, and because internal workplace rules require them to express an opinion that reflects the institution's position. Due to time constraints, our experts did not have the opportunity to ask for permission to express their views by name and job title, and it was considered more important to analyse the individual views expressed by the experts than to analyse the position of the organization. 8 experts participated in the online expert focus group discussion, 1 woman and 7 men. The youngest of our senior experts is 28 years old and the oldest is 61 years old, with an average age of 43. The participants in the online focus group interviews have the following higher education qualifications: computer engineer-economist, engineer-economist, computer engineer (2), economist-psychologist, sociologist, engineer, computer scientist. The average number of diplomas per person is 1.75. 3 people have a doctorate (economics, military engineering, security science) and one participant is an expert of the National Association of Human Sciences and one is an expert of the Hungarian Military Science Association.

Introduction of the junior expert group

The core group was made up of students from the Óbuda University, Bánki Donát Faculty of Mechanical and Safety Engineering (students of mechanical engineering, mechatronic engineering and safety engineering). Within the core group, a smaller group was formed, which includes those who took the course "Artificial Intelligence in Engineering", announced by Csaba Kollár, in the spring semester of 2022. Students who have taken the course were invited by email to participate in a focus group discussion on the challenges, opportunities, risks and future of artificial intelligence. We invited 5 full-time students to participate: Dávid BÁRCZI, Rita Bianka BARNA, BAYARAA, Burtejin, Noémi BENKŐ, Zsolt FÁBIÁN. The junior expert focus group consisted of 3 women and 2 men, the youngest participant was 20 years old, the oldest 28 years old, with an average age of 24 years. All five of our junior experts are all students of security engineering with a bachelor's degree in information security specialization. Their previous training included: logistics administrator, fire safety lecturer, medical college, high school, network administrator, mechanical engineer, mechanical engineering technologist, CNC programmer, systems operator.

CONCLUSIONS AND RECOMMENDATIONS BASED ON THE OPINIONS OF SENIOR EXPERTS

The first experience of the majority of the expert group were based on theoretical, cybernetic models, which they are currently using in the field of AI development and education, and thus have a historical overview of the development of AI and are therefore authoritative in estimating future trends. Just as career choices are often based on childhood experience, their experience of films and reading about AI and robots have also played a key role. They also use fiction as a constant source of inspiration in their work, a relatively rare phenomenon among professions. In the future, there will be a growing demand for highly qualified experts and developers, and building on this interest, films and games could be used to orient children towards different areas of storytelling intelligence as a potential career path.

There is no consensus among experts on the definition of artificial intelligence, nor is there a consistent literature, but it is a generic term. The distinction between robots and artificial intelligence has led to two possible scenarios: humanoid/animoid robots competing with humans in a dystopian future, taking over and subjugating humans, but the risk of increased tensions as social inequalities increase was not mentioned. The optimistic view is that robots and artificial intelligence will bring about positive changes in the quality of human life, because human society will be able to use its coping potential constructively.

The difficulty of defining AI is also due to the wide range of its current applications, which experts have listed, while implicitly touching on the issue of social trust: people may become suspicious of technical backgrounds they do not know, fearing misuse and therefore rejecting their use because of the bi-sensitivity deficit. Experts have also raised the impact of the attention economy, the contradiction that "user-friendly" ("attention-grabbing", convenient, fun, requiring little cognitive effort) applications are not necessarily safe, reliable, serving the interests of making a sure profit. In contrast, the return on investment of professional systems is not always guaranteed, and investors are often sceptical about the economics of AI [28].

The complexity of AI makes it challenging for professionals to carry out risk analysis on a horizontal basis, to assess potential threats (mis-programming, external attacks, etc.) and to estimate the extent of the damage. This raises the issue of addressing secondary damage in AI risk assessment, i.e. systematizing errors resulting from human actions and exploring ways to address them, such as algorithm biases, limitations of human cognitive capacity [28].

The techno-pessimistic view of the experts is that the dominance of competitive motives will increase in the foreseeable development, i.e. the acquisition of economic, power and security superiority, which may eclipse ethical aspects, human values and democratic rights, and that people may find their happiness in the meta-world rather than in the physical world. Although a number of international regulations have been put in place to implement AI in a human-centered way (e.g. 193 countries have signed the Ethical AI Agreement in 2021) [29], experts argue that these will not prevent socially dangerous developments, but are not sufficient on their own and that more broad-based action is needed. In addition to civilian applications, they do not consider the unrestricted use of artificial intelligence in defense and national security areas to be permissible, nor do they consider it permissible to override the relevant existing internationally valid rules and agreements on

ethical standards and the destruction of the enemy. Experts consider a broad 'socialisation' of artificial intelligence, education, awareness-raising and education to be key to countering these threats, and to maintaining the social control that will provide the basis for an optimistic vision of the future.

All participants agreed that the transmission of universal human values is essential in the development and education of artificial intelligence and that this is a priority in their work. Although Hungary's Artificial Intelligence Strategy 2020-2030 [30] comprehensively describes the guidelines for preparing society and the expected changes in the labor market, experts believe that we are already lagging behind in anticipating these changes, not only in terms of dissemination, but also in terms of concrete plans and their implementation, especially in measures to deal with the mass unemployment that will arise as a result of robotization and automation, the failure of which could escalate into a social crisis.

FORMULATED ON THE BASIS OF THE OPINIONS OF JUNIOR EXPERTS CONCLUSIONS, RECOMMENDATIONS

The junior experts were university students with a background in artificial intelligence, who, due to their age, had a different experience and time perspective from the senior experts. The emotional impact of the recent childhood film and other experiences with AI was vivid in their minds. The junior experts agreed that the unpredictability of the pace and direction of AI development means that the potential futures range widely. It was felt that if social awareness and pro-activity were to be replaced by a lack of interest and fears shaping people's attitudes, this passivity could make the techno-pessimistic scenario a self-fulfilling prophecy. A distopical vision of the future was a recurring element in their manifestations, which could be a real threat to the shape of their lives. In the social perception of AI, there was unanimous agreement that fear predominates in the middle-aged and elderly, both as a result of media representation and as a natural human reaction to the lack of certainty. Young adults, i.e. their own generation, are not aware of AI, they are uninformed, they are not interested in it, they do not appreciate the dangers, they are vulnerable and have a resigned attitude towards AI. Among children at large, the presence of AI is already a natural part of their socialization and they are not threatened by the use of their data, for example, which is less accepted by older people. Differences in attitudes between generations were also a recurring organizing force in our thinking.

The definition of AI and robots did not result in a unified view among junior experts, as theorists use different definitions, rather as an umbrella term. The need for control over AI was stressed as an essential component, e.g. in the form of an emergency stop button providing security, which could be interpreted as a means of actively coping with the threat and leading to the desired optimistic scenario. The dangers of AI included mass unemployment and the 'awakening of self-awareness' that would result from increasing autonomy as technology developed, and the associated aggression against the human race. In both cases, the fabulous intelligence is transformed from being a subordinate to a rival, rather than a servant of humanity, which today reinforces the trust deficit that goes hand in hand with fear.

The junior experts also use a wide range of artificial intelligence applications [31], some of which they themselves use mostly with their smartphones. Their drawbacks include language and geographic accessibility limitations, but reliability issues have become more

important, for example an application that performs computational tasks and generates erroneous results.

The expected qualitative change in interpersonal relationships based on multidimensional experiences as the boundaries between virtual worlds and physical reality blurred was also highlighted, with mixed views, both positive and negative, expressed by participants. The hope of comfort, security and a happy life defined their hopes, but they also saw the danger of the world becoming boring and "too comfortable". In the world they envision for 2032, they expect, among other things, smart cities, homes and vehicles to operate safely, predictably and with humane choices. In the case of major decisions, responsible human decision-makers and bodies using decision-support systems are still considered acceptable.

In the content of the opinions expressed in the junior group, the psychological needs described by Ryan and Deci in their [32] self-determination theories can be identified from above. According to this theory, three basic needs can be distinguished in humans:

(1) autonomy, i.e. the possibility of free choices and actions, which is threatened by "non-human decision-makers" or "self-aware" master-intelligence competing with the human species;

(2) the need for competence, i.e. a sense of self-efficacy, threatened by the loss of control over artificial intelligence and by "coddling";

(3) the need to connect, i.e. to form relationships based on love and respect, threatened by the multiverse of virtual worlds.

The experts also highlighted the need for human traits in certain applications of AI, such as the need for social and emotional intelligence, or aggression to protect the person in one's care. The latter is particularly important in the case of military AI or the use of robots, e.g. the validity of the rules of military warfare in the case of human-robot or robot-robot combat.

COMPARATIVE ANALYSIS OF THE OPINIONS OF SENIOR AND JUNIOR EXPERTS

Despite their differences in experience and perception, senior and junior experts were similar in many areas.

In both groups, the influence of childhood film and other experiences on career choices can be identified as a background to AI-related careers. Among seniors with a stable career identity, it was complemented by a commitment to their professional work by providing continuous development opportunities and a livelihood. Due to the generational difference between the two groups, juniors approached the issues from a future time perspective, while seniors have a broader view of the past, present and future of AI. Both groups are familiar with a number of AI-based applications, most notably smart homes (domotics), in-smart cities, autonomous vehicles, healthcare and pharmaceutical applications, applications in the fields of banking, military, agriculture, environment, education.

Despite the fact that the focus groups were composed of competent experts, there was no consensus on the definition of AI in any of the groups, which is in line with the fact that there are many different definitions in the literature and therefore it can be considered as a generic term. In both groups, the associations were structured around security and convenience, which the juniors defined in emotional terms (good-bad, scary) and the seniors in

rational terms (technical, economic). In the distinction between robots and artificial intelligence, robots were associated with physical embodiment, with both groups having cyber-physical systems at one pole and humanoid/animoid robots at the other, the latter endowed with social, emotional skills. Both groups agreed that media representations of robots and AI are significantly more dystopian, where the existence of the human race is threatened, and therefore the social perception of AI is dominated by negative emotional content, mainly fear.

The experts in the focus groups elaborated on the risks and challenges associated with AI, which covered the following main areas:

- Social distrust: AI is often presented in contexts that pose a threat, can be abused and thus lead to suspicion and rejection.
- Lack of transparency: due to its complexity, it is a challenge for highly skilled and experienced professionals to analyse the risks of practical applications of AI, e.g. to identify potential threats (mis-programming, mis-education, cyber-attacks), to estimate the resulting damage to financial or other resources.
- Bias, discrimination, which reflects systemic errors in human decision-making, e.g. may result from machine learning based on biased data (religion, gender identity).
- The security of information and data becomes less and less guaranteed due to lack of regulation and transparency, the use of data becomes opaque, privacy and human rights may be violated, e.g. social credit systems.
- Artificial intelligence is used for unfair political and economic power purposes, democratic values and human rights are violated.
- Restructuring of the labor market: on the one hand, the replacement of human labor by AI-based technology may lead to mass unemployment due to lack of appropriate retraining. On the other hand, the growing demand for a workforce with the skills needed to develop artificial intelligence could lead to labor shortages, with the emergence of new professions that are not yet known. Labor market restructuring could lead to increased social inequalities and social fragmentation.

According to junior experts, generations have different approaches to AI. The younger a person is, the more accepting and natural it is for them, while older people are distrustful and find it threatening. The experts say that steps and measures to prevent threats cannot be delayed, that adverse phenomena can be controlled today, that the future can be shaped and that professionals must play an active role in this.

Experts are confident that in 10 years' time, AI will be serving the well-being and comfort of mankind, contributing to an improved quality of life, and operating in a humane and ethical way. To this end, the widest possible range of society must be prepared for the expected impact and use of AI. Both groups believe that the boundary between virtual worlds, meta-experiences and physical reality will blur, that the search for happiness will be shifted to the virtual world, that flesh-and-blood human relationships will be replaced by worlds constructed by avatars, and that sensory experiences will be replaced by sensor technology.

A specific area of human-centeredness in AI is warfare, where aggression and the application of traditional rules of warfare raise a number of questions. The seniors approached the issue from the side of regulation: in this field, the introduction of artificial

humanity is necessary, i.e. the possibility of human intervention should always be a given, the counter-attack should not be significantly greater than the strike. Nor can the goal be the total destruction of the opponent, and compliance with the rules and regulations of international law applicable to war must remain a requirement. Rules of war must also be developed for AI-powered devices. The juniors examined the issue from a technical point of view: the use of robotic soldiers was seen as a risk factor in terms of uncontrollability and the technical sophistication that could determine the outcome of war. On the aggressive behavior of robots and AI, the juniors agreed that it could be acceptable to protect human physical integrity, imagining as an example a family robot protecting a child from a burglar, which would switch into defensive mode by the child's physiological reactions.

Both groups expressed the importance of creating ethical artificial intelligence, without which their concerns about a dystopian future are perceived as a real threat. Artificial intelligence is a man-made "creature" that can become a partner or rival in coevolution, living in peace with the human race. The junior experts were more likely to express emotional opinions, while the seniors shared concrete experience and knowledge, thus expressing their views in a more factual and rational way.

A further content analysis aspect for future research, beyond the scope of this paper, could be the investigation of the psychological needs of humans in attitudes towards mastery intelligence. The widely known Maslow's hierarchy of 5 levels of needs can serve as a starting point. Providing for the analysis of physiological needs in the field of production and services, by further developing the processes already in use, can make life more and more comfortable, but a persistent sense of comfort can lead to boredom. Safety needs are focused on the ability of artificial intelligence to be detectable, reliable and to minimize risks and hazards, e.g. the "protective robot". Social robots ("lovable robots") already exist to satisfy the need for belonging and love, and artificial intelligence with human empathy, emotional and social intelligence is filling this role. The need to be valued is satisfied if people retain their autonomy and creativity, thereby gaining the recognition of others, e.g. by taking an active role in the development of AI, or in the work of regulatory bodies, and by lifelong learning to meet the challenges of the labor market. Finally, there is ample scope for self-fulfilment, not only in the metaverse, but also through the creative use of more leisure time in a comfortable life.

Both groups found the focus groups useful to share experiences and opinions, to reflect together, to learn new perspectives and to become aware of the wide range of threats.

SUMMARY THOUGHTS, CONCLUSIONS

Humanity is at a critical stage in the development of artificial intelligence, which has the potential to dramatically improve the quality of human life, but also to realize a dystopian future, as experts have explicitly stated. Sinderman [34] and his colleagues have distinguished five attitudes towards AI, reflecting these two directions: trust in AI, AI will be beneficial for the human race, on the contrary, AI is scary, threatening, will lead to the destruction of the human race, will cause unemployment in many fields. In order to reduce the potential threats, regulatory measures by international organizations (e.g. UNESCO [35], OECD [36], [37]) are not sufficient, action at societal level is needed to protect democratic values and human rights, to increase public welfare and well-being, to enhance personal data and information security. To improve social trust, the ethical development of

intelligent systems must ensure efficiency, transparency, professionalism, define responsibilities and minimize the potential for abuse, as declared by a community of 700 experts [38]. Experts, as stated by Tilesch and Hatamleh [39], call for a globally accepted, feasible vision of human-centred AI, based on universal human values, harmonizing divergent interests, and a unified regulatory framework, but the paradigm shift will only be complete if humanity creates a new worldview based on conscience and trust, thus ensuring the survival of the species.

BIBLIOGRAPHY

[1] M. Tegmark: *Life 3.0: Being Human in the Age of Artificial Intelligence*, Knopf; 1st edition, 2017.

[2] Encyclopedia Britannica: Available: https://www.britannica.com/technology/artificial-intelligence

[3] Cambridge Dictionary, Definition of artificial intelligence from the Cambridge Academic Content Dictionary © Cambridge University Press, Available: https://dictionary.cambridge.org/dictionary/english/artificial-intelligence

[4] J. McCarthy,What is artificial intelligence? Revised November 12, 2007, Available: http://jmc.stanford.edu/articles/whatisai.html

[5] S. Russel, P. Norvig, Artificial Intelligence, A Modern Approach (3rd Edition), Prentice Hall, 200., ISBN 0-13-604259-7, Global Edition: ISBN 9781292153964

[6] A. M. Turing Computing Machinery and Intelligence, Mind 49: 433-460, 1950, Avaiable: https://www.csee.umbc.edu/courses/471/papers/turing.pdf

[7] N. J. Nilsson, The Quest for Artificial Intelligence: A History of Ideas and Achievements, Cambridge, UK: Cambridge University Press, 2010, Web Version Avaiable: http://www.cambridge.org/us/0521122937

[8] European Comission, WHITE PAPER On Artificial Intelligence - A European approach to excellence and trust, Brussels, 19.2.2020 COM(2020) 65 final, Avaiable: https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

[9] OECD, Artificial Intelligence in Society (Summary in English) Avaiable: https://read.oecd-ilibrary.org/science-and-technology/artificial-intelligence-in-society/summary/english_9f3159b8-en#page1

[10] UNESCO, Steering AI and advanced ICTs for knowledge societies: a Rights, Openness, Access, and Multi-stakeholder Perspective, UNESCO Series on Internet Freedom, 2019, Avaiable: https://unesdoc.unesco.org/ark:/48223/pf0000372132

[11] Council of Europe: Towards Regulation o AI Systems, Global perspectives on the development of a legal framework on Artificial Intelligence systems based on the Council of Europe's standards on human rights, democracy and the rule of law, Council of Europe Study DGI (2020)16, Avaiable: https://www.coe.int/en/web/artificial-intelligence/-/-toward-regulation-of-ai-systems-

[12] Cs. Kollár: *A mesterséges intelligencia és a kapcsolódó technológiák bemutatása a biztonságtudomány fókuszában* in Kiberbiztonság/Cybersecurity, edited by Z.Rajnai Biztonságtudományi Doktori Iskola, Budapest, 2019, ISBN: 978-963-449-185-9, Available: https://drkollar.hu/blog/2020/01/27/kiberbiztonsag-cybersecurity-ujingyenesen-elerheto-kiadvany/

[13] Executive Office of the President National Science and Technology Council Committee on Technology, Preparing for the Future of Artificial Intelligence, October, 2016, Avaiable: https://obamawhitehouse.archives.gov/sites/default/files/whitehouse_files/microsites/ostp/NSTC/preparing_for_the_future_of_ai.pdf
[14] J.R. Searle, Minds, brains, and programs, Behavioral and Brain Sciences, Volume 3, Issue 3, September 1980, pp. 417 - 424 Avaiable: DOI: https://doi.org/10.1017/S0140525X00005756
[15] G.A.Tilesh, o. Hatamleh: Between brains, Taking back our futrein the AI age, GTPublishDrive; Publishdrive ed. edition, 2020.
[16] OECD, Recommendation of the Council on Artificial Intelligence, OECD/LEGAL/0449, Adopted on: 22/05/2019, Avaiable: https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449
[17] Cseh-Szombathy L. és Ferge Zs. (szerk.), *A szociológiai felvétel módszerei*. Budapest: Közgazdasági és Jogi Könyvkiadó, 1971.
[18] B. Earl, *A társadalomtudományi kutatás gyakorlata*. Budapest: Balassi Kiadó, 2000.
[19] K. Krippendorff, *A tartalomelemzés kódszertanának alapjai*. Budapest: Balassi Kiadó, 1995.
[20] Horváth D. és Mitev A., *Alternatív kvalitatív kutatási kézikönyv*. Budapest: Alinea Kiadó, 2015.
[21] W. Gordon és R. Langmaid, *Kvalitatív piackutatás*. Budapest: HVG Kiadó, 1997.
[22] Langer K., *Kvalitatív kutatási technikák*. Gödöllő: Szent István Egyetemi Kiadó, 2009.
[23] K. N. Malhotra, *Marketingkutatás*. Budapest: KJK Kerszöv, 2002.
[24] P. A. Scipione, *A piackutatás gyakorlata*. Budapest: Springer Hungarica, 1984.
[25] Kollár Cs., *Reklám- és reklámszöveg kutatás*. Budapest: Protokollár Tanácsadó Iroda, 2004.
[26] Vicsek, L., *Fókuszcsoport*, Budapest, Osiris kiadó, 2006.
[27] Kollár Cs. A szakértővé válás, illetve a szakértők kiválasztásának és megkérdezésének módszertani kihívásai *VEZETÉSTUDOMÁNY* vol. 49:2 pp. 63-75., 2018, doi: 10.14267/VEZTUD.2018.02.07.
[28] D. G. Harkut, K. Kasat and V. D. Harkut, *Introductory Chapter: Artificial Intelligence - Challenges and Applications*. Artificial Intelligence - Scope and Limitations Edited by Dinesh G. Harkut 2019.: DOI: http://dx.doi.org/10.5772/intechopen.84624 (2022.03.10.)
[29] UN News, *193 countries adopt first-ever global agreement on the Ethics of Artificial Intelligence,* 25 November 2021., https://news.un.org/en/story/2021/11/1106612 (2022.03.10.)
[30] *Magyarország Mesterséges Intelligencia Stratégiája 2020-2030*, Mesterséges Intelligencia Koalíció, Digitális Jólét Program, Innovációs és Technológiai Minisztérium, https://digitalisjoletprogram.hu/hu/kiadvanyaink (2022.03.10.)
[31] Kollár Cs.: *A mesterséges intelligencia és a kapcsolódó technológiák bemutatása a biztonságtudomány fókuszában* in Kiberbiztonság/Cybersecurity, edited by Z.Rajnai Biztonságtudományi Doktori Iskola, Budapest, 2019, ISBN: 978-963-449-185-9.:

https://drkollar.hu/blog/2020/01/27/kiberbiztonsag-cybersecurity-uj-ingyenesen-elerheto-kiadvany/
[32] Ryan, R. M., Deci, E. L., *Self-determination theory: Basic psychological needs in motivation, development, and wellness.* 2017. New York: Guilford Publishing
[33] Maslow, Abraham H. A *Theory of Human Motivation (1943),* Originally Published in Psychological Review, 50, 370-396. http://psychclassics.yorku.ca/Maslow/motivation.htm
[34] Sindermann, C., Sha, P., Zhou, M. et al. *Assessing the Attitude Towards Artificial Intelligence: Introduction of a Short Measure in German, Chinese, and English Language.* Künstl Intell 35, 109–118 (2021). https://doi.org/10.1007/s13218-020-00689-0
[35] UNESCO: *PRELIMINARY STUDY ON THE ETHICS OF ARTIFICIAL INTELLIGENCE (COMEST)*, 2019. https://en.unesco.org/artificial-intelligence/ethics/cases#biasedai
[36] OECD (2019) *Artificial Intelligence in Society.* https://www.oecd.org/publications/artificial-intelligence-in-society-eedfee77-en.htm
[37] OECD, *Recommendation of the Council on Artificial Intelligence*, OECD/LEGAL/0449. 22/05/2019, https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449
[38] IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems 2019. *Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems.* 1st ed.: IEEE. https://ethicsinaction.ieee.org/#series
[39] G.A.Tilesh, o. Hatamleh: *Between brains, Taking back our futrein the AI age*, GTPublishDrive; Publishdrive ed. edition, 2020.