LUDOVIKA UNIVERSITY OF PUBLIC SERVICE

Doctoral School of Military Science

THESIS BOOKLET

Author's Summary of the Doctoral (PhD) Dissertation

Lieutenant Colonel András Tibor FEHÉR:

The human aspects of artificial intelligence in the defense sector

Supervisor:

Colonel Dr. Imre NÉGYESI
associate professor

Budapest, 2025

SUMMARY

PRESENTATION OF THE RESEARCH

THE SCIENTIFIC PROBLEM, RESEARCH APPROACH, AND AREAS OF RESEARCH

Interpretation of the title: the research program

The title summarizes the research program well, so it is worth explaining it as an introduction. The term artificial intelligence (hereinafter referred to as AI) can be understood in the first instance as what the reader already knows about it – clarification of this term is one of the objectives of the research. The term "human aspects" refers to the fact that the topic is related to a number of human sciences, from philosophical ethics to psychology, which gives the research its interdisciplinary nature. It should be clarified why, as a soldier, I do not use the term "national defense" in the title, but refer to a broader area. The answer is that the use of military terms (war, military, combat, etc.) would not adequately cover the approach of the research and would even be misleading, as the research itself will point out. I was looking for a word that goes beyond the traditional military segment and refers to those involved in the defense of the country. I came to the conclusion that the term "defense" is an appropriate compromise and a good choice, even though the study does not intend to examine certain parts of the colloquial "defense sphere" (see the delimitation of the topic).

In summary, although the common set of terms AI, the human aspect, and the defense sphere goes beyond the scope of this dissertation, their intersection refers well to the content of the research. Even so, I can only analyze a few small slices of the area indicated in the title within the present framework.

I completed my research on December 1, 2024. After that, I did not continue with targeted research, but while working on the material, I was forced to incorporate some newer information into the text due to its importance. My goal was not (and could not have been) to produce an up-to-date text in such a rapidly changing field, but I strove to use formulations that would be as timeless as possible.

The fundamental problem of the research

The general problem of the research as a whole can be summarized as follows:

Major technological advances have a significant impact on relationships between people and between human communities (countries), but the expected impact of AI stands out among these due to the novelty of its cognitive characteristics. To examine

this specifically, it is necessary to study the impact of AI on relations between countries (the defense aspect of the problem) and to analyze the impact of AI on the transfer of information between people (the human and terminological aspects of the problem).

This work therefore fits into further impact assessments of AI for defense purposes, and its findings are intended to support such investigations. Although the above problem description already anticipates the main directions of the research, the designated area is still too broad. I can only narrow it down in part through the research objectives and research questions, as they are not intended to include the criteria on which the author bases his investigation or the directions from which the published approaches seek to address the problems (and those from which they do not). Recording the latter characteristics also helps in planning the research, which I will discuss here.

Aspects of the investigation

In my research, I have tried to keep the following perspectives in mind.

1. **Interdisciplinary approach**. I do not expect technical (mathematical, engineering, or programming) findings, so here I combine research in the fields of philosophy, computer science, pedagogy, and psychology with other scientific fields, more specifically with military science, sociology, and certain parts of the history of science.
2. Addressing **defense aspects** in as many sub-research areas as possible.
3. **Clarification of terminology.** Thorough explanation of important or lesser-known key concepts in each field of research (including their various interpretations) and clarification of any terminological anomalies that arise.
4. **Filling gaps**. Presentation of as many subfields as possible on which no publications have appeared in Hungarian, or for which there are few sources even globally.
5. **Practicality in the details**. Not only the planned scientific findings, but also the smaller partial results should be useful in themselves (e.g., own concepts, models, divisions, unique approaches, insights, ideas, suggestions, etc.).
6. **Focus on the background**. The study should focus not only on the phenomena generated by AI, but also on their possible background.
7. **Prevention**. Based on the above, this is also a key consideration given the timeliness of the topic.
8. **Existential approach**. I try to approach the topics in such a way that the reader can apply as many of the findings of the studies as possible in their own life.

Based on research perspectives 2 and 3, the terminological and defensive research attitude permeates the thesis in general, beyond the specific objectives.

Delimitation of the topic (which the research avoids)

This interdisciplinary topic would be most effectively approached from a multidisciplinary perspective, but due to the limitations of human capabilities, true interdisciplinary knowledge is not feasible. In the introduction, I listed the sciences on which I intend to rely most heavily. Below, I exclude those sciences and research areas that are obvious connections to the topic, which I will only touch upon as necessary.

1. Legal aspects – it should be noted that I approach ethics from a philosophical point of view and do not use the term in the more common legal sense. Although I deal with regulatory issues, I only examine the information security and protection aspects of these.
2. As described in the introduction, I do not focus on the aspects of defense related to disaster management, law enforcement, secret services, water management, etc.
3. Some of the examined sections touch upon economics, but I do not analyze them as an economist.
4. For the sake of completeness, I will only touch on medical and biological sciences at a basic level; it is not my intention to conduct a thorough analysis of brain research, biological systems, or AI implants built into humans.
5. I plan to examine the cognitive battlefield in depth at a later date; here, I will only touch upon it.

OBJECTIVES

I examined two problem areas (P1-P2) and assigned research objectives (O1-O2) and hypotheses (H1-H2) to them

- P1: Terminological problems related to the topic
- P2: The problem of the paradigm shift in defense

O1: To identify as many factors as possible without which the definition of AI is ambiguous, or which, when included, make the concept confusing, and to make suggestions for resolving the problems identified (new concepts, divisions, other terms, etc.).

O2: To analyze the relationship between global change and new trends in power enforcement with AI, and to determine the factors of AI that are important from a defense perspective.

HYPOTHESES

I formulated the hypotheses taking into account their expected uses.

H1: Important aspects are missing from the current definitions of Artificial Intelligence, and are burdened by misunderstandings.

- Planned uses: Based on the identified shortcomings, a new, more usable definition of AI that will perhaps remain valid for the next ten years. This will make it easier to deal with regulations and aspects of AI that are not taken into account in the current definitions for defense planning. A better definition will help to raise awareness of the essence of AI and develop the mindset necessary for its use, which is also crucial for education, and such training is of strategic importance.

H2: AI further shifts the emphasis of force projection towards soft operations.

- Planned uses: The findings will primarily be useful in disciplinary and strategic planning, but the terminology, models, and insights developed along the way are also necessary from a military science perspective.

RESEARCH METHODOLOGY AND STRUCTURE OF THE STUDY

I determined the specific questions of the study based on the objectives of the research. I will explain this together with the rationale for the structure of the material, and then I will present the research methods that are tailored to answering the research questions.

Research questions and the logical structure of the dissertation

When planning the research, it became clear that the planned findings would need to be supported in several ways. Thus, the logical arrangement of the material is a particular challenge, as the studies are closely interlinked and each depends on the others. (To help the reader understand the many internal connections, I made the internal references easier to follow by using specific subchapter numbers.) Finally, among the many logical ways of organizing the exposition, the following order seemed to be the version that required the least amount of foreshadowing, while also meeting the expectation of chapter proportionality. I have structured the material so that the first four chapters address both objectives, while Chapter V supplements, summarizes, and systematizes the P1 problem area, and Chapter VI brings together the threads of thought with regard to P2.

At the very beginning of the study, I had to introduce the subject I am investigating, so Chapter I reviews the issue of intelligence. In Chapter II, I examine machine learning as the basis of intelligence. This is a somewhat technical chapter, which also outlines AI-related techniques, but I have tried to avoid excessive technical depth in view of the target audience. Learning and intelligence consequence in a certain level of freedom. Freedom, individual and community autonomy are closely linked to defense challenges, and in our time, this is accompanied by the potential autonomy of machines. Therefore, in Chapter III, I summarize the issue of autonomy and my thoughts on it. After examining three basic concepts related to AI, Chapter IV critically reviews the security challenges arising from neural networks, building on these concepts. However, it also examines the possibilities for the development and spread of AI. In Chapter V, I wish to go beyond the planned achievement of Objective O1: as part of the terminology issue, it was logical to discuss the expected change in the term "informatics" in light of AI, which I use in the final analysis. The sub-studies in Chapter VI, based on all the material available up to that point, specifically discuss defense topics in order to achieve goal O2. Based on the structure presented, I defined the research questions that outline the specific program of the chapters:

Q1: The issue of intelligence: *How are artificial and human intelligence interpreted in the world, how are they defined, and are there any important aspects missing from this?*

Q2: The issue of technology and learning: *How do machine learning implementations and other related technologies work, and are there any factors among them that should also be taken into account in the concept of AI?*

Q3: The issue of autonomy: *How can the machine and human implementation of autonomy be captured, what are their main characteristics and main differences?*

Q4: Challenges and development: *What challenges does neural AI pose, how can its errors be handled, and what factors can influence the development or spread of the technology?*

Q5: The issue of terminology: *In what cases is the use of the term AI incomplete or confusing, how can improvements to AI terminology be proposed, and does this affect the term "informatics" in a data management paradigm that differs from the classical one?*

Q6: Defense issues: *How (using which models, concepts, comparisons, and analyses) can we grasp the innovations that are expected in the defense segment as a result of AI?*

Selected research methods and formal solutions

Based on the above questions, previously identified aspects and delimitations, and taking into account that different methods are justified in different phases of the research, I conducted the research using the following methods:

It was worthwhile to rely on deductive methods and an analytical approach in the basic investigations. I used document and source analysis methods when exploring the research findings of each issue or reviewing official documents. Etymology also played a role in the terminological studies, as the older meanings of borrowed words reveal a lot about them. I used correlation analysis methods to systematize the information revealed in the analyses: for example, I compared or contrasted the material obtained, in other cases I divided the subject of the study into levels or parts for better analysis, or I made them transparent in tables – these divisions can also be used beyond their current use. However, it was also necessary to synthesize the systematized but analytical material; without the use of inductive methods, it would have been difficult to achieve findings. I often arrived at conclusions discursively, as a logical connection between the information uncovered and my own thoughts, while at other times I drew conclusions from the results of organizing the information or from comparisons. In some cases, it was worthwhile to prove the proposed conclusion by applying complete induction or by refuting the opposite of the statement. In addition to proofs, these logical constructs were also well suited for syntheses based on predictions or proposals. For ease of understanding and to check the train of thought, I also drew mind maps, two of which I created digitally. I used a descriptive method to communicate the resulting trains of thought, but in the case of phenomena and concepts, I used retrospective analysis, and for the examination of processes, I used an exploratory method. Where necessary, I illustrated the explanations with diagrams. For the sake of logical fluency, I also used an "invisible method": by converting each train of thought into limited-time projected presentations, I was often able to improve the structure, wording, and conciseness.

During my work, I tested several AI-based text processing systems, which I used to search for quality literature (instead of Google Scholar) or to quickly describe unfamiliar concepts (instead of Wikipedia). I couldn't trust AI to write paragraphs, as the result required more follow-up work than if I had written them myself (except for the literature review). The machine text suppressed the characteristics of my own approach, making the style of the text impersonal.

A BRIEF SUMMARY OF THE STUDY

Chapter I: Human and machine intelligence

Although most people have some idea of what AI is, it was important to clarify what is meant by it before examining it in this study. So, after providing a brief explanation of the internal connections between the first three chapters and their order of discussion (I.1.), it was necessary to begin the dissertation with an analysis of K1 question[1] , thereby laying the foundation for further research. The answer to K1 in the narrower sense was provided by the brief history of the concept and the section on official definitions. However, a broader answer could only be addressed in connection with further research: Therefore, in light of the other research questions, I attempted to clarify the concepts of artificial and intelligent agents, interpret the concept of intelligence etymologically, describe the "levels" of AI, and present the types of human intelligence (1.2 and 1.3). This was followed by an example to illustrate that the term intelligence cannot be characterized solely by cleverness and IQ: to this end, I presented how emotional intelligence can be imitated by the results of affective computing. This also highlighted the fact that machines only mimic humans at the level of phenomena, i.e., they do not have emotions, but are only capable of copying their external expression (I.4.). On the one hand, the above provided an appropriate basis for interpreting the other chapters, and on the other hand, it enabled me to take steps not only towards goal O1, but also towards O2.

Chapter II: The mechanization of learning and AI-related technologies

Misconceptions about AI are most often rooted in a fundamental lack of understanding or misunderstanding of the technology. In order to avoid this, it was essential to provide at least a concise answer to question K2[2] in order to achieve goal O1, a better understanding of AI. Within this, it was necessary to explain the key to today's AI systems, the "inside of AI," i.e., deep learning. This abstract outline of deep learning models also presented everything that needs to be known to avoid fundamental misunderstandings, i.e., to correctly examine topic P1. (II.2.). In addition, it was necessary to present the technologies and principles related to AI, i.e., the "outside of AI." However, I not only presented the modern technologies that are closely related

[1] K1: The question of intelligence: *How are artificial and human intelligence interpreted in the world, how are they defined, and are there any important aspects missing from this?*

[2] K2: The issue of technology and learning: *How do machine learning implementations and other related technologies work, and are there any factors among them that should also be taken into account in the concept of AI?*

to the true effectiveness of AI, but I also named this common set (AIRT[3]) (II.1.). Finally, I described aspects (principles, models) that are primarily related to the topic of P2 protection. Here, I outlined some important directions for AI hardware support and presented solutions coming from biology to computer science, primarily inspired by the lifestyle and behavior of swarm-living animal species, but I also mentioned language models, fuzzy logic, and neural databases (II.3.). I would like to emphasize that I have already successfully used the content of this chapter in my teaching, and in the future, I would like to develop this section into teaching material in a revised form.

Chapter III: The Anatomy of Autonomy

The basis of every good decision is that the decision-maker is able to use their freedom appropriately, rather than abusing it. Intelligence can help to judge what is right. Intelligence, in turn, is able to make this judgment through learning. This has long been the case with humans, but with the advent of computer technology, machines have also begun to make decisions in place of humans, and more recently, their learning abilities have enabled them to judge certain things, which creates a sense of freedom. Therefore, after analyzing intelligence and learning—as a consequence of these—it was necessary to examine the content of the concept of autonomy, partly as part of the overview leading to the concept of AI (O1), but even more so for the purpose of defense (O2). Following these two guidelines, I reduced this rather difficult area to a brief overview of the K3 question[4] . In doing so, I took a critical look at the currently prevalent versions of autonomy classifications and defined my own classifications. The basic question focused on the similarities and differences between human and machine autonomy, which can be grasped methodologically by defining the realization of autonomy in both cases in degrees and then comparing these levels.

To this end, after reviewing human autonomy (III.1.), I replaced the key terms of the known classification, thus creating a scheme called the Four-Type Autonomy Classification (4TA) (IV.2.), which also included "extreme autonomy."

Based on the official names of the levels of machine autonomy, it became clear that even experts consider these technologies to be merely "advanced automation," and that it is rather the tabloid press that spreads the word *"autonomy*," e.g., to the levels of self-driving cars. This proved to be a useful finding, especially in answering the P1 terminology problem (IV.3.). In

[3] Artificial Intelligence and Related Technologies

[4] Q3: The issue of autonomy: *How can the machine and human implementation of autonomy be captured, what are their main characteristics and main differences?*

my study, I proposed a more refined, more generally applicable, and multi-level classification than the official one (IV.4.). Then, based on the ideas in the first four subsections, I compared the two types of autonomy, listing the main differences between machine and human autonomy. This will be useful in relation to the P2 defense problem (IV.5.).

Chapter IV: Security challenges of neural AI

From the perspective of both the defense and terminology problems (P2 and P1), it was essential to review how neural AI can be considered novel, i.e., what challenges it poses. I examined only two aspects of the challenges based on the K4 question[5] : one provides a brief overview from the perspective of technology, and the other from the perspective of people's connection to AI. I began the technical overview with a list of challenges, drawing on the information in the first chapters. The overview of the factors presented here also highlighted the security challenges posed by the novelty of AI, for which it was useful to review and classify the types of problems arising from neural networks (IV.1.). The analysis of AI errors is also of paramount importance to the defense problem (P2), so—continuing the ideas raised in the previous chapter on autonomy—I also examined the poor implementation of machine freedom, i.e., I reviewed neural bias (bias, prejudice) (IV.2.). Based on this, I compared the errors of the two types of technology and analyzed the extent to which the presented distortion of AI is novel compared to traditional computer program errors (IV.3.). The overview also required the discussion of two human topics: first, I presented the pitfalls that have already caused setbacks in the development of AI as historical lessons (IV.4.), then, in order to examine AI's learnability in greater depth, I highlighted in a separate subchapter the important aspect that points to why people find it difficult to follow innovations that involve such a major shift in perspective (IV.5.).

Chapter V: Rethinking the concept of artificial intelligence

Based on the research in the first four chapters, I have reached the conclusion of my examination of the terminological problem (P1). In Q5,[6] I received some answers to the questions I posed, but there were still a few sub-research questions left to answer in order to provide an

[5] Q4: Challenges and development: *What challenges does neural AI pose, how can its errors be handled, and what factors can influence the development or spread of the technology?*

[6] Q5: The issue of terminology: *In what cases is the use of the term AI incomplete or confusing, how can improvements to AI terminology be proposed, and does this affect the term "informatics" in a data management paradigm that differs from the classical one?*

acceptable answer to the question. In H1, it had to be proven that *important aspects are missing from the current definitions of AI*. [7]In O1, I divided this deficiency into two parts: on the one hand, without the inclusion of certain content, the concept can be misunderstood because key terms *are* omitted from the definitions; on the other hand, the concept is used inconsistently and remains vague because *it can include* mutually exclusive characteristics, so a clearer definition of the term AI would also be necessary. In the previous chapters, I have already identified the terms that are missing from the definitions. In this study, I have further expanded their scope and pointed out the reasons for the insufficient definition of the concept.

At the beginning of the study, I focused on the contradictory meanings of the term rather than the aspects missing from AI, as I reviewed how many types of non-neural (deterministic) systems the term AI is used for. The technical merging of classical and neural systems is very forward-looking, but it can lead to serious misunderstandings (even security issues) when the use of the term AI blurs the line between whether or not it refers to a neural system. To review this, I systematized the cases in which the term AI is not applied to neural systems (V.1.). I distinguished three cases: I examined the incorrect use of the term by researchers, users, and market participants. I analyzed separately the manageability of the three types of problems when developing a new AI concept. In these studies, I also outlined a "terminological degradation model." I consider this to be a generally applicable principle, which can be used to predict even future inaccuracies in terminology. As an example, I applied the model to the (as yet not widely used) term AGI, and thus arrived at a result that has since been partially confirmed by reality (V.2.).

I then approach AI from the perspective of information and related sciences, as this reveals the broad interdisciplinary context that actually lies behind the technology. I could have titled this "An examination of changes in the meaning of the term 'informatics' in light of AI." This sub-research provided an important addition to goal O1, more precisely to the new concept of AI (avoiding the term "informatics"), but it also helped me discuss the defense problem (P2). In addition to explaining a number of terms related to the topic, I outlined the phenomenon of convergence of the sciences (under the influence of AI), which also had implications for IT training. (V.3.).

[7] O1: Identify as many factors as possible without which the definition of AI is ambiguous or which, when included, make the concept confusing, and make suggestions for resolving the problems identified (new concepts, divisions, other terms, etc.).

In the following subchapter, I felt that three summaries were indispensable: First, I published a lexicon-like list of different AI classifications, which can be a useful collection for anyone interested in the topic, and I also presented a proposed classification matrix, which can be used to place and visualize the knowledge of a given system on a palette based on the combined consideration of many such classifications. I then summarized the formulations that were missing from or confused the term AI, as revealed in the examinations of the previous chapters. I also briefly touched on why it is impossible to formulate a perfect definition for such a problematic phenomenon as AI. Although it was not part of the relevant hypothesis, and I do not consider it a scientific result, after so much research, it was appropriate to add my own suggestions to the already significant number of AI definitions, showing that it is possible to put the outlined summaries into sentences (V.4.). This concluded the verification of H1 and the related research.

Chapter VI: The new era of power enforcement and AI

In the final part of the research, I was able to complete the ideas of previous research and explore the defensive aspects of AI from a number of perspectives. To this end, in response to question K6[8] , I outlined the four topics in the introduction, which, when explored, fulfilled the O2 objective[9] . Since O2 covers a much broader spectrum than H2 hypotheses[10] , the subchapters were structured accordingly:

1. In the first subchapter, I presented in general terms that the directions in which power is exercised are changing even without AI, but that AI capabilities fit well with these changes and even reinforce them (VI.1.).
2. I then shifted the focus to examining how AI and cyberspace technologies reinforce each other in this regard (VI.2.).
3. I then analyzed a specific aspect of the first topic, the possibilities for abuse in the digital space, but from a much broader perspective than is commonly taken (VI.3.).
4. Finally, I presented two consequences of the hypothesis: the impact of AI on military informatics and its educational aspects (VI.4.).

[8] Q6: Defense issue: *How (using which models, concepts, comparisons, analyses) can the innovations that are expected in the defense segment as a result of AI be captured?*

[9] O2: Analyze the relationship between global change and new trends in power projection with AI to determine the factors of AI that are important from a defense perspective.

[10] H2: AI further shifts the emphasis of power projection towards soft operations

The supporting arguments for the subconclusions necessary to prove H2 are presented within these sections. I planned to prove the hypothesis from four angles:

(1) From the perspective of paradigm shifts in the enforcement of power, which is the subject of this chapter.
(2) From the perspective of autonomy research and AI biases. To this end, it was necessary to conclude the lines of thought previously raised, which arose in the research on autonomy and then in its continuation on AI challenges (in Chapters III-IV).
(3) From the perspective of the interaction between AI and society. This has been the subject of my previous research, and only the most important parts of this research, those that provide relevant arguments, have been included in this chapter.
(4) I also drew arguments from the analysis of the defensive use of non-intellectual AI, i.e., affective computing (machine emotions) presented in the first chapter.

This thesis, particularly this chapter, contributes to the broad examination of the geopolitical-digital paradigm shift, which is one of the most significant challenges in military science today.

CONCLUSIONS AND FINDINGS

SUMMARISED CONCLUSIONS

To prove the hypotheses, I primarily used the numbered subconclusions[11] found at the end of each chapter. I also marked the relevant chapter numbers of the dissertation for the most important ideas. I believed that the logic of the proofs would be easier to follow if I used visual representations to help overview the relationships between the various concepts. Therefore, I have also depicted the essence of the series of conclusions in a mind map,[12] which I present in a space-saving manner in the explanation (Figures I and II[13]), but which can also be found enlarged in the Appendices to the Thesis Booklet (Figures III and IV). The source chapters are indicated by different colors.

[11] When numbering these, e.g., S4.1 indicates the first subconclusion drawn from the fourth chapter.

[12] I will not write out the quoted subconclusions in detail here, only their essence, and for the sake of clarity, I will only give the numbers of some of them in the figures.

[13] The Roman numerals used in the figures are intended to clarify the difference from the figure numbering used in the dissertation.

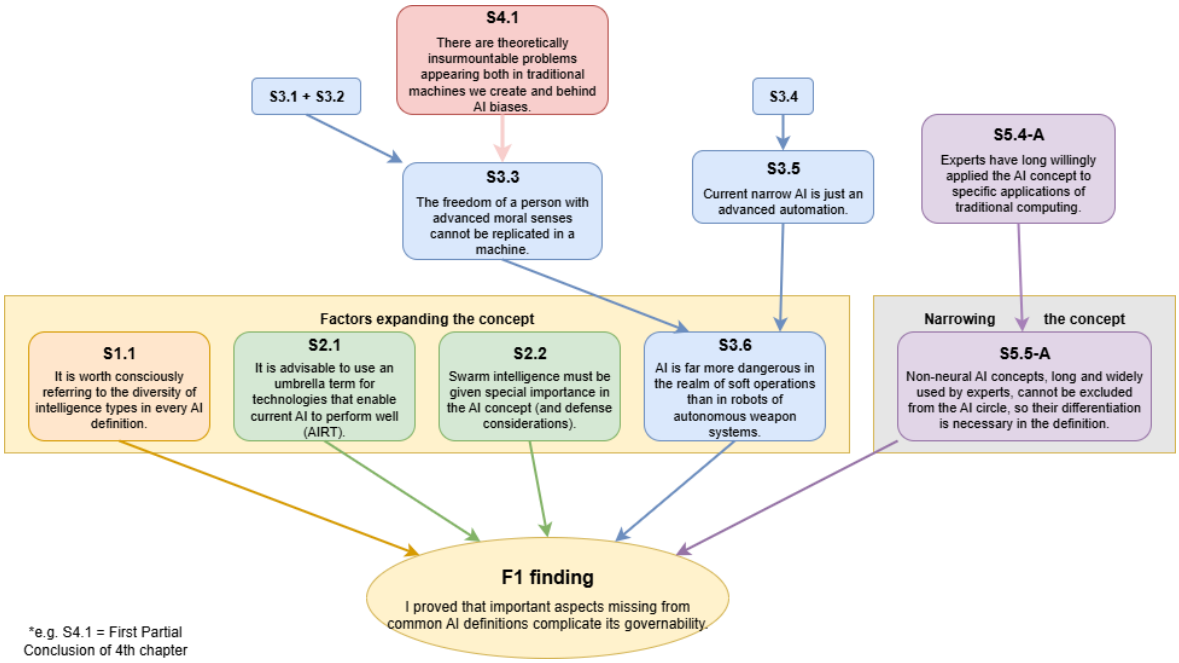
The series of conclusions leading to the F1 findings

It had to be proven that *important aspects are missing from the current definitions of AI* (H1). To verify this, I aimed to analyze the overly narrow and overly broad uses of the term AI in O1. I uncovered examples of both problems that need to be remedied. In my proposals, I recommend specific terms and sometimes mention alternative solutions, but these always refer to a particular line of thought. In other words, I do not recommend the use of these terms, but rather see the need to reflect certain characteristics or principles in order to arrive at a better definition of AI. (I also suggested avoiding certain terms in the definition, but the research did not aim to review these, so I will not discuss them here.) I have summarized the words to be represented and their justifications in the two tables included in the appendices to the thesis (see Appendix, Tables I and II). These summaries were aimed at formulating new AI concepts, but below I present a summary of the arguments and ideas as proof of H1. In my research, I identified five terms (factors) without which the definitions of AI are ambiguous, thus confirming the assumption (see Figure I). Four of these are aspects *missing* from the concept of AI:

- I consider the lack of mention of **the many types of intelligence** (R.1.1.) to be the most serious conceptual error. This reinforces the old philosophical-anthropological misconception that intelligence is synonymous with cleverness. Science has already corrected this early error, and the development of numerous other types of intelligence is underway, so omitting this term is technologically misleading (and discriminatory (I.2., I.3.).
- **The technology behind AI** (S2.1.) should be included in a clear definition because there is a very significant difference in usability between a mobile phone's facial recognition AI agent and AI that runs on powerful cloud servers and draws its knowledge from a huge database fed by millions of IoT devices. For the latter technology circle, I have proposed the acronym AIRT[14] (II.1.).
- **Some reference to the imitation of biological cooperation** (S2.2.) would be important in order for the concept to adequately cover the significantly different directions of development. However, modeling the processes taking place in living organisms or the cooperation developed between individuals of certain species differs significantly from the functioning of central AI (e.g., AIRT), as it also allows for distributed (non-central) intelligent systems. I suggested including swarm intelligence in the definition because

[14] Artificial Intelligence and Related Technologies

it is better known, more widespread, and more concise (see II.3.3. – I rejected the much more appropriate term bionics because it is unfamiliar and could be misunderstood).



I. Figure: Mind map illustrating the proof of H1 (own creation – enlarged, see Figure III)

- **Automatics** (S3.6.). I consider its use important in order to avoid the term autonomy and to reduce fear of AI. According to S3.5, current thin AI systems are still only capable of achieving a higher degree of automatics, even hypothetically[15] . The latter statement is based on the international J3016 standard (III.3.2., III.4.1.) and S3.4. (see below), but its theoretical basis is S3.3. According to the latter, such machines are in principle incapable of autonomous decision-making similar to that of humans. I based this theoretical impossibility in part on S3.1, where I used analysis to demonstrate that fundamental differences can be identified between human and machine autonomy[16] (III.5.). However, S3.3 is also supported by the later S4.1, as according to this, there are *theoretically* insurmountable problems (IV.2-3.), in other words, copying humans into machines has not only been "practically unsuccessful" with these systems. Returning to the aforementioned S3.4., this statement points to the same difference in a different way,[17] when it

[15] This refers to the possibility of "fifth-level self-driving cars," which are not yet on the market.

[16] This statement is not evident due to the prevalence of the human-machine concept.

[17] The fact that machines are not actually capable of human-like autonomy is demonstrated throughout Chapter III. Machines designed for this purpose would have to be orders of magnitude more complex for such a thing to even be possible, since humans are orders of magnitude more complex than them.

states that the levels of autonomy in machines should be linked *to the level of built-in AI*[18] (III.4.1.), whereas in the case of humans, the levels of autonomy can be explained on *ethical* grounds. (III.1-2.)

The fifth factor to be considered did not identify a shortcoming, but focused on what should be *eliminated* from the concept of AI. This correction seeks to rectify the contradictory (confusing) use of the term AI with the following recommendation:

- **The neural AI** (S5.5-A) proposal suggests using a prefix to distinguish AI based on neural networks, which operate as black boxes, from traditional systems that are called, believed to be, or presented as AI for various reasons (even though they operate with pseudo-learning or are incapable of learning, see V.1.). Traditional codes can also be marked with a prefix and called "deterministic AI agents" or "non-neural AI." It would be important to use this primarily in the wording of regulations.

In addition to these, I have suggested four further terms for more precise wording, where possible (based on S2.4., S4.3. and S5.6.). These are: *synergy*, *cognition*, *mapping*, and *computing* (instead of informatics!). Since these are not specifically missing from the term, but merely clarify it, they do not need to be proven, so their explanation can be omitted here. (The rationale for their use and the subchapters where I discussed their usefulness can be found in Tables I and II of the Appendix.

The series of conclusions leading to the F2 findings

It had to be proven that *AI further intensifies and supports the trend whereby the emphasis in the use of force continues to shift towards soft operations* (H2). To verify this, I aimed O2 at analyzing new trends in the use of force and their relationship to AI in order to determine the aspects of AI that are important from a defense perspective. I proved H2 from four directions: (1) from the perspective of paradigm shifts in the use of force, (2) from the perspective of autonomy research and AI biases, (3) from the perspective of the interaction between AI and society, and (4) from the perspective of non-intellectual AI (machine emotion imitation).

(1) The arguments in the first line of thought are based on three conclusions. S6.1 states that a paradigm shift has actually been taking place in power enforcement for some time, even without AI, and I have outlined its main features (VI.1.1.). However, based on S6.2-A, the AI capability system fits this trend very well; it is highly applicable on such fronts, and its

[18] In order to ensure that the various systems can be classified from a security perspective.

emergence further catalyzes the processes that have already begun (VI.1.1-3.). It almost follows from the above, but I have also analyzed separately (VI.2.1-2) how AI multiplies the possibilities of cyberspace, while also significantly increasing its dangers.

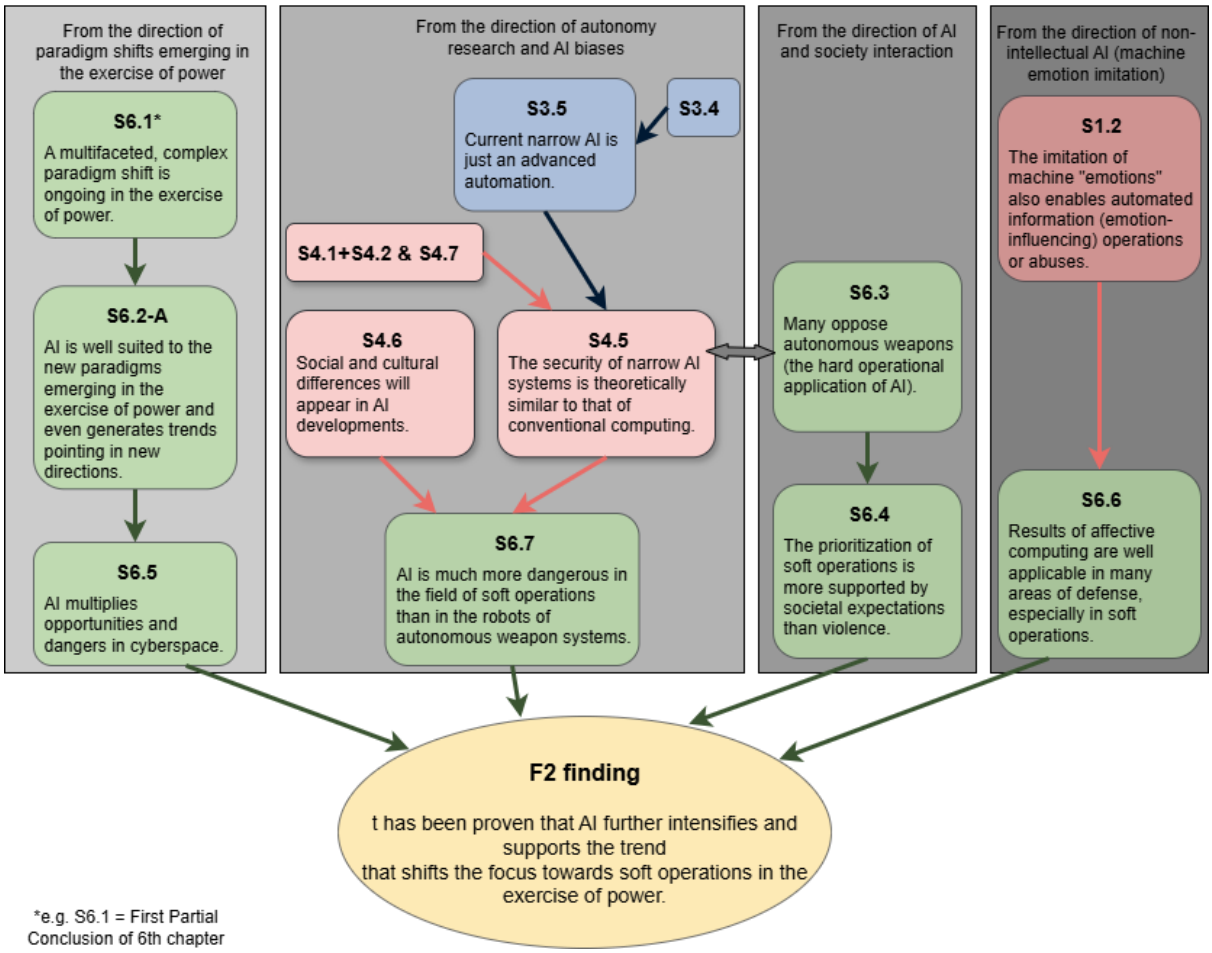
(2) The second line of argument is complex. Based on several interrelated sub-studies, the argument presented in S6.7. states that, currently and in the near future, AI is much more dangerous in the field of soft operations than in autonomous weapon systems (as compared to remote-controlled robots or explosive, chemical or biological weapons) (VI.2.4.). I arrived at this conclusion from two lines of thought. One is a social perspective (S4.6.), in which I point out that cultural and social differences will manifest themselves in AI developments (IV.2.2.(3)). These differences may also manifest themselves in cyberattacks with different ethics, the dangers of which are almost impossible to predict. In the other line of thought (S4.5.), I showed that the security of thin AI systems is similar in principle to that of traditional computing (in practice, they are very different, IV.2-3.). I arrived at this position on the one hand based on autonomy tests, which revealed that current thin AI is merely advanced auto matics (S3.5.[19]), and on the other hand, this was confirmed by studies of AI biases[20] (IV.2-3).

(3) From the third perspective, I point out the problem in S6.4. by showing that trends in the (mainly Western) world are moving towards "non-violence". Soft operations are best suited to these trends in every respect (VI.1.1.). On the part of AI, this is reinforced by S6.3., which refers to the political and social protests, the much greater attention and fear surrounding the deployment of hard AI, i.e., autonomous weapon systems ("killer robots") – as opposed to the use of soft AI, which is hardly ever heard of (VI.1.5., VI.2.4.). The latter is also refuted by the idea in S4.5. of direction (2), since if both the old large systems and the new neural systems can only be considered safe on the basis of testing and statistical values, then the practical difficulties that currently favor traditional systems are only characteristic of the current years of development and do not make AI robots "deadly" *in principle* (IV.2-

[19] For a summary of the ideas leading to S3.5, see the footnote to finding F1.

[20] S4.5 is referred to from this perspective in S4.2, for example, which states that the transparency of traditional systems, which are growing and becoming increasingly complex, and AI are converging towards each other (IV.4), as well as S4.1, in which I summarized that there are *theoretically* insurmountable problems that appear both in the traditional machines we have created and behind AI distortions (IV.2-3). This is captured differently in S4.7., which states that errors can only be found through testing in both robust traditional and AI systems (IV.2-3.).

(4) The fourth direction complements these arguments by pointing out that the results of affective computing are well applicable in many areas of the defense sphere, but are particularly applicable in soft operations, as pointed out in S6.6. (VI.2.3.). I refer back to the initial studies, where, in addition to the basics of the technology, S1.2 emphasized that not only can machine "intelligence" be used for offensive, surveillance, defensive, or other purposes, but the imitation of machine "emotions" also provides opportunities for automated information (emotion-influencing) operations. (I.4.)



II . Figure: Mind map of the proof of H2 (own creation – enlarged, see Figure IV)

SCIENTIFIC FINDINGS

By proving the hypotheses, the research achieved the following scientific results:

F1: I have proven that important aspects are missing from the widespread definitions of Artificial Intelligence
that make it difficult to regulate.

F2: I have demonstrated that AI further intensifies and supports the trend
the emphasis in enforcement continues to shift toward soft operations.

SUGGESTIONS FOR THE USE OF RESEARCH

The multifaceted analyses of the study and the more than fifty subconclusions allow for numerous conclusions to be drawn, but hopefully the following will suffice to illustrate the practical side of the research. I will describe these uses, which I consider to be the most important, in four groups (U1-4), with the remark that each use could be classified into several groups. In addition to considering the two results in this direction, I also based the compilation on the partial results. Before doing so, however, I will present two proposals that led to more complex conclusions.

A conclusion and a conjecture

Of the two practical consequences below, the first emphasizes a fact that may be obvious to professionals but is not yet widely accepted in practice. The other is a conjecture that has not yet been fully proven, but the compiled argument provides a suitable starting point for further investigation.

Conclusion: I have demonstrated that the integration of AIRT and other paradigm-shifting technologies into education is essential for military and defense training.

I believe that the educational insights of this research can contribute useful arguments to increasing the priority of AI education in Hungary and to its acceptance by society and educators. (This is why I am sharing it, even though the conclusion itself is professionally evident.) I have mentioned the field of education in many places due to my personal involvement: I analyzed it on a theoretical level in the studies in Chapters IV, V, and VI, and its direct practical applicability is discussed in Chapter II and Table 3 of Chapter V.[21] Here, I summarize the arguments of the theoretical studies in such a way that they support the importance of the topic from four directions.

1. **Self-taught learning does not work in the case of AI and other new paradigms.** S4.13. (the vertical learning model) states that it is difficult for people to change their mindset. However, it is not possible to make good use of technological paradigm shifts (often simply using them) without this, i.e., merely by acquiring additional information. Of course, this also requires teachers who are capable of conveying this new mindset. However, it must also be acknowledged that when vertical learning is required instead of horizontal learning,

[21] This serves as the basis for the first teaching material and can provide a foundation for the second topics.

it cannot be achieved quickly at the course level, as it requires much more mental energy and inspiration than is available there. In other words, students can only achieve an adequate level of understanding and application (or the beginnings thereof) of new paradigms through training courses lasting several months or, in some cases, several years, which also teach a new approach (IV.5.1.).

2. **Due to the convergence of sciences, all graduates need a basic knowledge** of **AI** in order to use and shape their own goal systems.[22] According to S5.2, all sciences converge towards AI (V.3.4), therefore the basics of neural networks and practical demonstrations of their multifaceted use should be incorporated into all higher education (V.3.4), as pointed out in S5.10. However, S5.9 supports this from a different perspective, stating that technical developments have reached a level where they must wait for models to be developed in the humanities and other sciences (e.g., defense) in order to make significant progress (V.3.1.). Therefore, the safe spread and development of AI requires the involvement of representatives from all sciences.
3. **If there is no technical breakthrough in the near future, then the stagnation must be exploited, and if there is, then citizens must be prepared for it.** S4.14. states that "a long summer of AI is expected," meaning that AI will spread and generate significant revenues, but in my opinion, disruptive changes affecting users cannot really be expected for some time (IV.4.). This situation can be exploited, according to S4.15, if the population is consciously inspired to learn about the current use and paradigms of AI. (This is not unnecessary even if a new AI paradigm does emerge, as it is almost certain that it will inherit many of the characteristics of the current technology in many respects) (IV.5.3.).
4. **General education in AI fundamentals is also important in the military and defense sectors for other reasons.** Based on the arguments presented in F2 (due to the paradigm shift in defense and military affairs), it is extremely important for all soldiers, but especially for information and communications technology and IT officer training, to learn about AI at the deepest level possible (VI.4.2.).

[22] Or even as a contribution to AI models from the perspective of one's own field of science.

Conjecture: The current AI models and the much-criticized "autonomous weapon systems" implemented on the basis of the guidelines do not raise any fundamentally new ethical problems.

In this conjecture, I formulated an insight in the field of philosophical ethics based on the ideas of the partial studies already used in the two results, and I reorganized certain subconclusions related to autonomy and machine errors. The practical significance of this hypothesis is pointed out in S4.10, which states that the general opinion regarding the insecurity of AI autonomy will continue to undermine trust in AI in the future, even when there is statistically little difference between the safety of traditional and AI systems (IV.3.). I use two approaches to prove this:

1. First, I gathered arguments that refute exaggerated approaches to AI capabilities.
    - Based on S4.9-A, current thin AI can only act contrary to human intent due to the will of its creators or their serial, large-scale negligence, and is not capable of becoming autonomous "just like that" or even awakening to consciousness (IV.3.).
    - According to S6.7-B, traditional systems are much more dangerous in the event of a detailed malfunction, as they continue to operate partially, whereas a trained neural network either simply does not work if it physically loses some of its neurons, or is able to repair itself or shut down dangerous capabilities.
        - In other words, the "danger of runaway technologies" exists even without AI, according to S6.7-A, and in fact, errors can be corrected more easily in thin AIs than in traditionally coded automatons (VI.2.4.).
2. Another set of arguments points out that traditional systems are not so much more reliable that it would justify a certain degree of "demonization" of AI.
    - According to S4.2, the complexity of traditional systems and the transparency of AI are converging (IV.4.), meaning that from a safety perspective, the difference between them is decreasing, as S4.5 also states (although from a practical perspective, traditional systems are still significantly better).
        - In this context, a "convergence of security" is already taking place, since the essence of the practical problems stated in S4.7 is that security, or more precisely, a statistically determined degree of security, can only be achieved through testing in both types of systems (IV.2-3.).

- The practical advantage of traditional systems is relativized by the fact that, according to S4.1, there are *theoretically* insurmountable problems that arise in both traditional machines and behind AI distortions (IV.2-3.). This also points to the fact that the two types of systems face fundamentally similar obstacles.

In summary, traditional computing and AI are not in sharp opposition. The above approach to the new paradigm (AI) makes it clear that the new era of mechanization "only" requires the rewriting of rules, but does not create an uncontrollable, autonomous evolutionary step that rebels against its creator.

Possible uses of the analyses of the P1 (conceptual) problem area

U1.1: Based on item F1, the shortcomings identified in the study can be used in AI

- For regulators: the wording helps to reduce the number of loopholes by reducing unregulated areas.
- In defense planning: it makes aspects of AI that are not taken into account by current definitions more manageable
- In education: it helps to raise awareness of the essence of AI and to develop the mindset necessary for its use. This is also crucial for military training. Education is very important in all areas as part of the defense system.

U1.2: Based on the shortcomings and inconsistencies identified in the study, I have proposed a new definition of AI in a more comprehensive and a more concise version, and I authorize their use with reference to the source (V.4.). This concept will hopefully be usable for a few years in the areas listed in U1.1.

U1.3: The collection of numerous possibilities for dividing and systematizing AI, as well as the classification matrix (V.4.1.), can be used

- in the development of regulations;
- in education;
- in further refining the concept.

U1.4: I have shown that the content of the concept of "informatics" can also change due to AI, and can even be significantly transformed (V.3.). Beyond its use in defense, this insight may be important in university courses related to AI technology, as it can elevate the current technology-centric teaching of AI to broader theoretical and practical perspectives.

Possible uses of the analyses of the P2 (defense) problem area

U2.1: The F2 finding provides decision-makers and experts with insights that can be used in the areas of defense strategy, regulation, economics, and education.

U2.2: Based on the F2 finding, military science needs new models and concepts to address the emphasis on soft operations. Of these, the following have been developed here (VI.):

- Characterization of the composite (hybrid) cold war to describe global processes.
- The theory of virtual power centers and the systematic model of virtual power projection, which can be useful for identifying possible defense vectors and areas of vulnerability.
- The distinction between adapted and direct military informatics, which can clarify the proportions and directions of external procurement and in-house development.
- A much more comprehensive systematization, analysis, and practical demonstration of digital abuses can serve as a useful perspective for legal defenders, in addition to its many military applications.

U2.3: Proposing a new place for military and defense informatics, summarizing the models developed and the problems encountered can also help to overcome the shortage of specialists in the long term (VI.4.).

Educational applications

U3.1: Chapters I and II were primarily developed with the aim of creating notes or textbooks for various courses, but I also plan to include excerpts from the other chapters of the study in more comprehensive versions.

U3.2: I have introduced the concepts of pseudo-learning, pseudo-AI, and pseudo-autonomy (V.1.4.), which can be used in promotional materials and basic education.

U3.3: Based on S4.13, taking into account the separation of horizontal and vertical learning (IV.5.1) could be very helpful when revising the requirements for adult education and continuing education, and when planning such training (e.g., not thinking in terms of courses if the trainees need a change of perspective)..

U3.4: I have shown that the widespread social integration of AI cannot be achieved at a pace commensurate with its rate of development (S4.15.), which may be useful in educational planning (IV.5.2-3.).

Other uses

U4.1: The assumption regarding autonomy—if fully confirmed—can be used effectively, according to S4.10, to dispel fears about AI and increase trust, which could contribute in many ways to solving the world's problems.

U4.2: New, more useful classifications of human and, in particular, machine autonomy (III.) can be used to classify and regulate AI systems (in my opinion, better than the current ones) as well as for information security issues – even in specific cases such as high availability systems (HAS) that can be implemented with a maximum of delta-level, i.e., complex autonomy that can be left to itself (without "machine user error") (S3.9., III.4.3.).

U4.3: With regard to information security, it is also useful to note that the transparency of traditional systems, which are growing in size and complexity, and AI are converging (S4.2., IV.4.), as this may necessitate a change in approach to system security (S4.8.). For this reason, it is necessary to treat the two systems together and eliminate them at the same time in all definitions and regulations.

U4.4: Demonstrating that all sciences converge towards AI (S5.2., V.3.4.), it has become apparent at the system level that knowledge of AI fundamentals is important in all scientific fields (both at the researcher and user levels). This can be put to good use in human resource management expectations in both the public and private sectors, as well as in education, since, according to S5.10, it is necessary to incorporate the fundamentals of participation in AI projects into training programs in higher education everywhere (V.3.4.).

U4.5: The terminology degradation model (S5.3., V.1-2.) aims to contribute to making marketing more ethical and promoting consumer awareness.

U4.7: The multidimensional (horizontal and vertical) description of development (S4.12., IV.4.1.) is a model that can be put to good use in the economic sphere.

U4.6: In the defense sector, it can be used primarily in the field of soft operations to reveal the shortcoming that citizens' hidden emotions are not yet legally actionable (they are not yet personal rights, such as hidden illnesses), even though data on them can be obtained using AI, even for abusive purposes. (S1.3., I.4.)

- The same insight may also be useful for legal protection.

RECOMMENDATIONS

To whet your appetite, I will highlight a few professions here, but this does not mean that the material is of no interest to those not listed here. In fact, I hope that all my readers have found material that is interesting and useful to them. I trust that some of these ideas will have an inspiring effect on their lives.

0. The study did NOT intend to provide technical innovations to colleagues researching AI technical solutions, but it would be great if as many of them as possible read it, as the aspects of their achievements analyzed here could also inspire them.
1. The primary target audience was **professionals** in the **defense sector**, including civilian and military personnel, national security and law enforcement agencies. The study provides useful insights for them in many places, but Chapter VI may be particularly useful. Those who have not dealt much with AI so far can gain an introduction to this world through its presentation of important problems in their profession, and the technological presentation of AI has been specifically formulated for such an audience.
2. A related target audience is those **interested in defense and military topics** who do not work in these fields but who, through the numerous specific approaches, models, and terminology, can gain a deeper insight into the emerging trends in power projection than they would from a single publication or popular article.
3. The other key target audience is **educators**, primarily colleagues working in higher education institutions. They can already make direct use of many of the subchapters and do not have to wait for notes to be compiled from the relevant excerpts. Furthermore, the section on military higher education can also be seen as a discussion starter for defense and military higher education professionals.
4. The approach taken in Chapter II may be useful for **humanities** scholars and anyone with an interest in the humanities, as they are likely to be interested in a general overview of AI, similar to the groups mentioned in point 1, and do not wish to delve into the technologies at an engineering or programming level. However, I also expect them to be interested in the philosophical nature of the study, which can provide a wide range of inspiration.
5. The material can be put to good use by **professionals involved in regulation**, especially IT security researchers and practitioners, but also lawyers. Numerous divisions and approaches have been directed at resolving regulatory issues related to AI, which they can hopefully further develop with the help of their own expertise and apply in the practice of drafting regulations.

6. **Legal advocates** can also make good use of certain parts of the paper, primarily the systematization of digital abuses, for which I have not found such a comprehensive study in Hungarian, or, for example, my view on the inclusion of citizens' secret feelings in personal rights.
7. **Economists** can also benefit from the ideas in this study. I consider my multidimensional development model to be particularly applicable in this field, but they may also find useful insights in the sections on AI regression and in the chapter on protection. The model of terminological degradation can contribute to ethical marketing, while other studies in the thesis can contribute to their forecasts or their own models.
8. Due to the demonstrated convergence of the sciences, it is quite certain that any profession will find interesting and useful parts in some section of the study, so I hope that representatives of all disciplines will be able to make use of the research due to its interdisciplinary (and somewhat multidisciplinary) nature.

APPENDIX

TABLES USED FOR PROOF[23]

I. table: Factors to be included in the new definition of AI (own creation)

	Summary	Term	AI-def.	Education
1.	**Types of intelligence.** I have shown (I.2.) that most current official definitions revolve around reason (keywords: thinking, cognition, teaching, autonomy). Thus, they are overly mind-centric and do not highlight the many other types of intelligence, even though research is being conducted on their machine simulation. I demonstrated this with the presentation of affective computing (I.3.). The new definition must primarily address this shortcoming. (Note: the Hungarian approach avoided this mistake. (I.1.2.))	Types of intelligence (S1.1.)	+ +	+ +

[23] These tables can be found in Chapter V.4.2 of the dissertation.

	Summary	Term	AI-def.	Education
2.	Based on the linguistic analysis (I.3.1.), no meaning was found that would need to be incorporated into the new concept. Most layers of meaning are present, and there is no reason to include those that are missing (e.g. understanding, recognition).	root analysis	-	-
3.	I also introduced the concept **of AIRT** (II.1.), which distinguishes the essence of AI from the highly complex systems that "surround" it. This must be validated in the new concept, if not with the acronym I proposed, then in some other way.	AIRT (S2.1.)	++	++
4.	The digital ecosystem (II.1.4.) does not need to be incorporated into the concept, as it is only a consequence.	digital ecosystem	-	+
5.	The divisions of learning models and machine learning do not provide any additional technical information that is necessary for a general definition (II.2.).	neural model types	-	+
6.	**Swarm intelligence** stands out among biologically inspired systems in terms of its importance and uniqueness: it highlights the importance of decentralised AI systems. A reference to the "digitisation of biological collaborations" (II.3.3.) would be more accurate – one of these should be included.	swarm intelligence / biological collaborations (S2.2.)	+ +	+
7.	The hardware background of AI, neural databases, and even NLP are not relevant to the definition (II.3.1&4&5.)		-	+
8.	Including forward-looking laboratory solutions would confuse the currently targeted concept and should be avoided (at the end of list II.3.1.). For example, augmented human solutions would require a rethinking of the word "artificial".	Future possible technologies	-	+
9.	**Synergy** (synergistic or symbiotic AI) refers to the goal of creating systems that are good to live with (I.2.1. & 2.3-4.), and its inclusion is recommended – indicating that such machines are not designed to work against humans.	synergy (S2.4.)	+	+

	Summary	Term	AI-def.	Education
10.	**Autonomy divisions** (III.): It is not necessary, and in fact it is recommended to avoid the term autonomy. Its use is only recommended when mentioned in conjunction with the word automation, in a negative sense.	autonomy	- - (+)	+
11.	**Thin AI is in fact just automation**: this should be pointed out in order to prevent fears (III.3.2. & III.4.1.).	automation (S3.6.)	++	+ +
12.	**AI challenges** (IV.1.): None of these belong to the essence of the definition, only to the description of AI.	challenges	-	+
13.	**AI biases** (IV.2.): Their details are not relevant.	distortions	-	+
14.	**Cognition / cognitive**: the word is recommended to be included as a reference to the limitations of technology (III.5 and IV.2.2.)	cognition (S4.3.)	+	+
15.	**Mapping**: the term is recommended to be used with reference to the limitations of technology (III.5 and IV.2.2.), as well as autonomy based on the definition of information technology (V.3.1.).	mapping (S4.3.)	+	+
16.	**Character of mistake** (IV.3.)**:** In certain cases, it may be useful to include it in a definition, but it is not essential if omitted.	mistake	-	+
17.	**Historical regressions of AI** (IV.4.): not necessary		-	+
18.	**Human acceptability of AI** (IV.5.): Only recommended in cases where the human challenges of learning the technology are to be highlighted at the beginning of a course.		-	-
19.	It is important to mention **non-neural AI solutions** (V.1.) separately, as this clarifies the ambiguity of whether deterministic or non-deterministic AI is meant by the term. It is essential to include one of the suggested prefixes if the definition is to be formulated in terms of safety rules.	neural and non-neural (indeterministic) AI (S5.5-A)	+ +	+

	Summary	Term	AI-def.	Education
20.	**Pseudo-AI** (V.1.4.): This would make the definition cumbersome and require explanation, so it is not necessary. However, I consider it essential in other texts, education and regulations in order to avoid confusion between the concepts.		-	+
21.	**Computer** technology[24] : It is important to mention this when formulating definitions for security rules, as the security of non-IT AI also differs significantly. (It could be omitted, for example, if someone wants an approach that also applies to biological AI, but this is not a reality at present or in the near future.)	Computing technology (S5.6)	+	+
22.	I suggest **avoiding the term "**computer science" because, as I have shown, it is not accurate to use it as a synonym for "computing" in the context of AI (V.3.1.).	informatics (S5.4)	-	+
23.	**The convergence** of **sciences** brought about by AI (V.3.2.), although extremely important, is also conceptually invalid.	science convergence	-	+
24.	**AI classifications** (V.4.1.) do not need to be part of the definition, as they complement it.	AI classifications	-	+

[24] The term "computer science" was included only because of its comparison with the word "informatics", not because it is new.

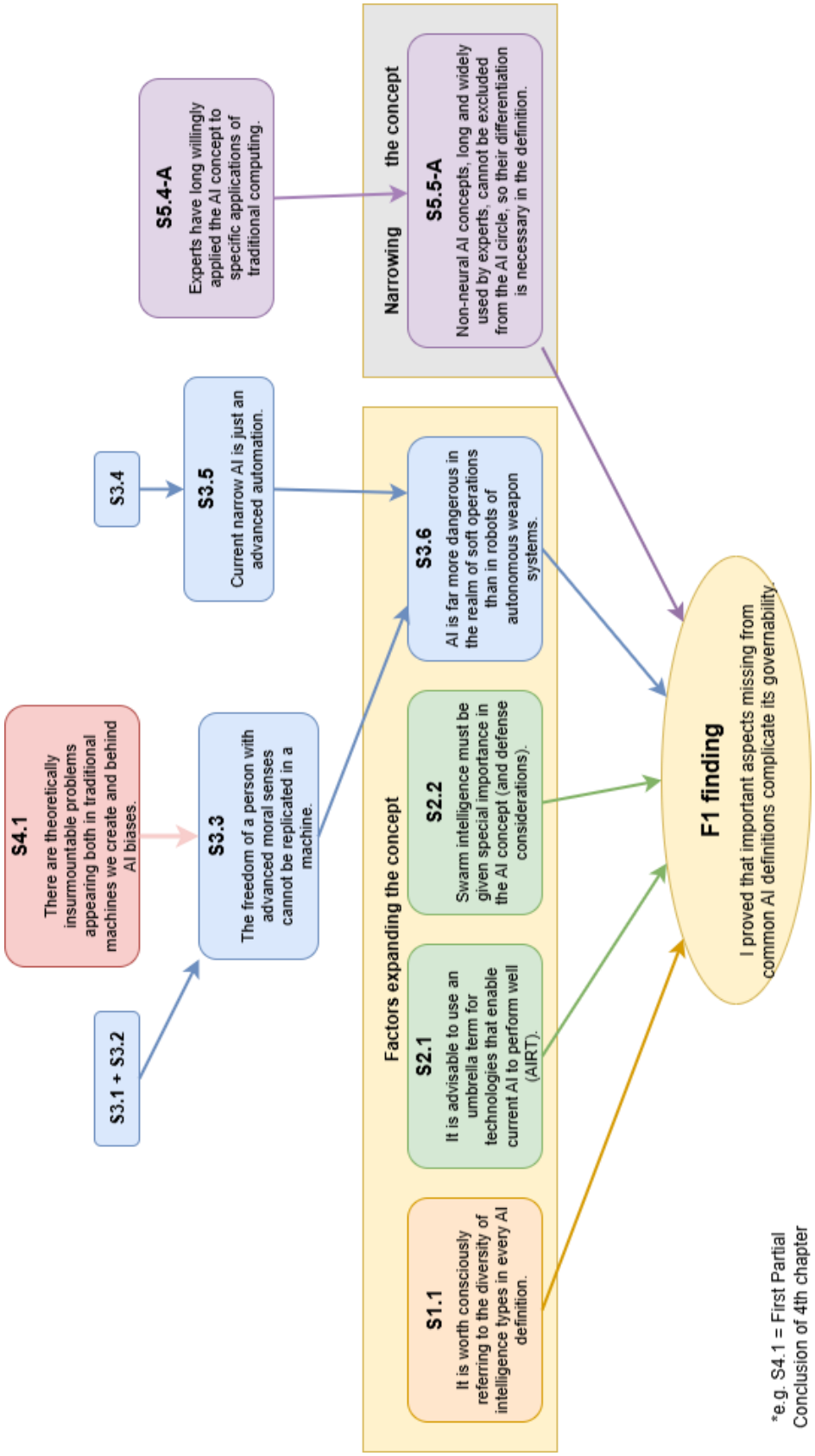
II. table: Why is the concept of AI ambiguous without certain terms? (own creation)

missing term	misunderstanding caused
types of intelligence	A mistake in human interpretation (in the field of philosophy and anthropology) is to narrow intelligence down to cleverness, which makes the perspectives of human imitation ambiguous – even though more and more types of intelligence are being analysed in greater depth, as these too can be developed in every human being, not just their intellectual abilities. (This goes beyond technology, as it suggests that human dignity is directly proportional to intelligence, and may generate contempt for less intelligent citizens in the future.
AIRT	Misunderstanding that a given system forms a unit with other technologies poses a security risk and leads to poor assessment of possibilities in use.
swarm intelligence	The structure and usability of a set of intelligent entities differs from an AI model running on a central server to the same extent that an anthill differs from a tiger. Of course, both are "animals" – but they are fundamentally different in terms of their functioning and goal achievement, just as their resource requirements need to be managed differently.
synergy / synergistic AI	The use of this term is intended to dispel (or at least mitigate) misunderstandings about AI. Technology that complements humans in a complementary way is a goal that should be a fundamental aspiration for all stakeholders, whether developers, users, regulators, supervisors or evaluators.
automation	Fear of machine autonomy is common, but this may be counterbalanced by emphasising this other term.
cognition / cognitive abilities	The scientific term referring to the world perception behind intelligence clearly indicates the direction in which technology is heading, namely towards machines being able to process as comprehensive a spectrum of human perception as possible.

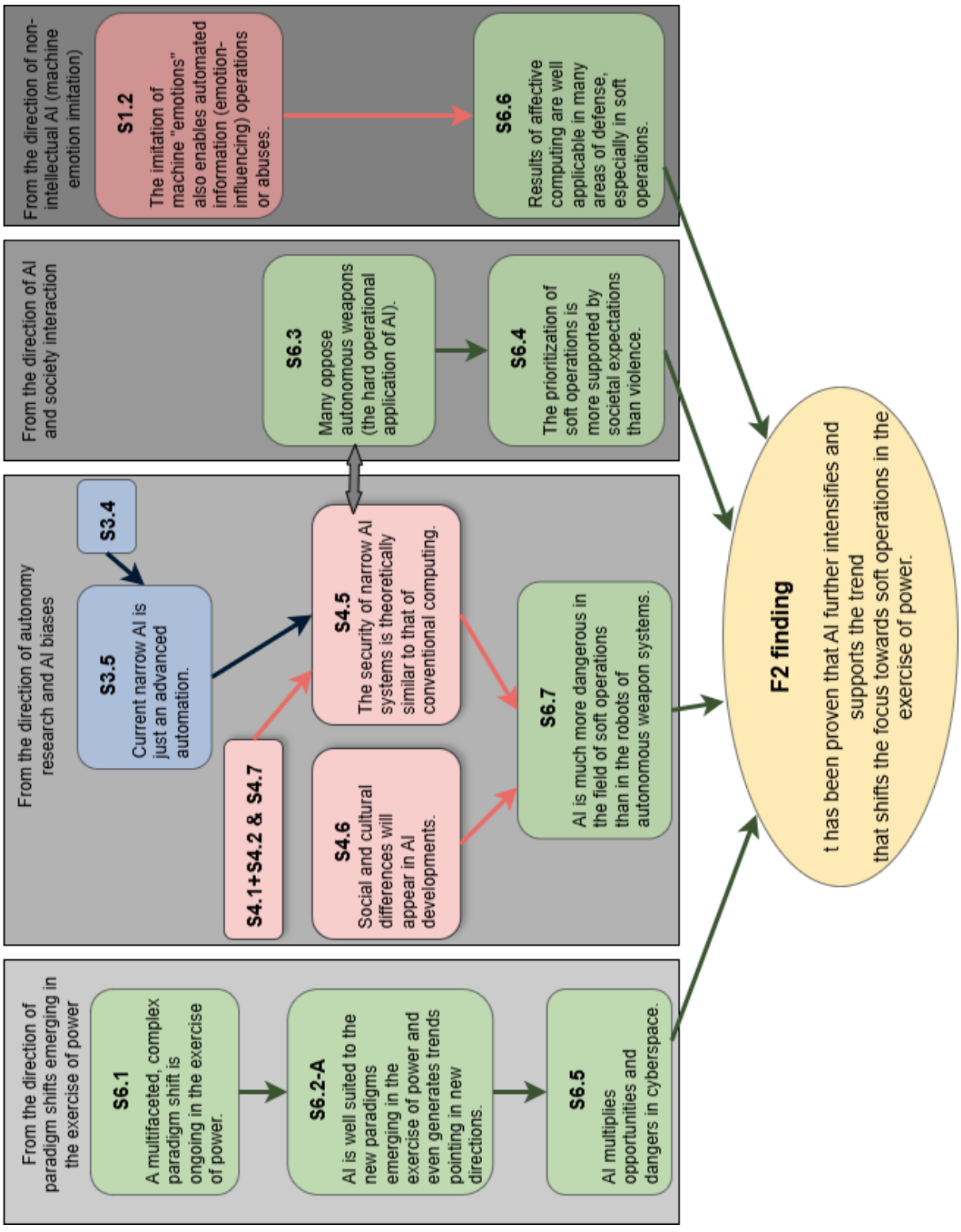
missing term	misunderstanding caused
mapping	It is important to note that computer science and AI map the world in some way into machine models, so the term refers both to the theoretical difference between machines and humans and to the limits of mechanisation – thus, its use is useful both in terms of dispelling fears and ensuring safety.
computer science	It may generate security misunderstandings in the future. Although this word is not missing from most of the currently used concepts, its omission may arise in connection with other artificial cognition enhancement technologies – for which, however, a different term should be sought due to significantly different controllability.
non-neural AI (deterministic AI)	The existence or role of neural black box agents in a system is often unclear, which can cause security misunderstandings, so their absence must be indicated in some way (e.g. with one of the two proposals).

ENLARGED MIND MAPS

I. Figure: Mind map illustrating the proof of H1 (enlarged)



II. Figure: Mind map of the proof of H2 (enlarged)



ABOUT THE AUTHOR

LIST OF PUBLICATIONS

1. FEHÉR, András Tibor: A tanulás 2.0 védelmi vonatkozásai – Milyen védelmi kihívásokat támaszt az emberi tanulással szemben a gépi tanulás paradigmaváltó hatása? In A mesterséges intelligencia hatásainak specifikus vizsgálata, szerkesztette Magyar Sándor, Budapest, KNBSZ, 2025. pp. 173-209., https://www.knbsz.gov.hu/kiadvanyok#1fd723c5-92d3-4dff-bc47-3edf4cb1b38a
2. FEHÉR, András Tibor: A mesterséges intelligencia természetes telei - Az MI technológia lassulását leíró forgatókönyvek és azok védelmi vonatkozásai. In A mesterséges intelligencia (MI) hatásainak specifikus vizsgálata, szerkesztette Magyar Sándor, Budapest, Katonai Nemzetbiztonsági Szolgálat, 2025. pp. 86-125, https://www.knbsz.gov.hu/kiadvanyok#1fd723c5-92d3-4dff-bc47-3edf4cb1b38a
3. FEHÉR, András Tibor: A mesterségesintelligencia-alapú hidegháború etikai háttere. Szerkesztette Kovács Zoltán. A mesterséges intelligencia és egyéb felforgató technológiák hatásainak átfogó vizsgálata, 2023., pp. 355–92., https://unipub.lib.uni-corvinus.hu/8690/
4. FEHÉR, András Tibor: Lépések egy biztonságos mesterséges erkölcsi érzék felé. In A hadtudomány és 21. század tanulmánykötet, szerkesztette Szelei Ildikó, pp. 114–34., Budapest, Doktoranduszok Országos Szövetsége, Hadtudományi Osztály, 2022., https://hdi.uni-nke.hu/document/hdi-uni-nke-hu/hadtudomany-es-a-21-szazad-kotet-2022.pdf
5. FEHÉR, András Tibor: Mesterséges intelligencia a kibervédelemben. In A hadtudomány aktuális kérdései napjainkban, szerkesztette Szelei Ildikó, II. kötet: pp. 123–38., Budapest, Ludovika Egyetemi Kiadó, 2022. https://tudasportal.uni-nke.hu/xmlui/handle/20.500.12944/17628.
6. FEHÉR, András Tibor - NÉGYESI Imre: A digitális visszaélések katonai vonatkozásai és az orosz modell. Honvédségi Szemle – Hungarian Defence Review , 151, sz. 5., 2023., pp. 3–15. DOI: 10.35926/HSZ.2023.5.1, https://real.mtak.hu/182029/
7. FEHÉR, András - NÉGYESI Imre: A gépi érzelmek a fegyveres erőknél és az autonóm rendszerekben. Hadtudományi Szemle, XIV, sz. 3., 2021., pp. 163–76. DOI: 10.32563/hsz.2021.3.12, https://folyoirat.ludovika.hu/index.php/hsz/article/view/5454/4759

8. FEHÉR, András Tibor – KERESZTHEGYI István: A gépi nyelvfeldolgozás védelmi felhasználási lehetőségei és kockázata. Szakmai Szemle, XXI., sz. 4., 2023., pp. 75–93, https://www.knbsz.gov.hu/hu/letoltes/szsz/2023_4_szam.pdf#page=75
9. FEHÉR, András - NÉGYESI Imre: Mesterségesintelligencia-alapú kibertértámadási modellek. Műszaki Katonai Közlöny, 31, sz. 3., 2021., pp. 73–87. DOI: 10.32562/mkk.2021.3.5, https://folyoirat.ludovika.hu/index.php/mkk/article/view/5495
10. FEHÉR, András - NÉGYESI Imre: A mesterséges intelligencia alapú kibertámadások lehetőségei. Studia Mundi – Economica, 8, sz. 5., 2021., pp. 25–34. DOI: 10.18531/Studia.Mundi.2021.08.05.25-34
11. FEHÉR, András - NÉGYESI Imre: Artificial intelligence in cyberspace (Part I) -AI based cyber-attack capabilities. RED - American Journal Of Research Education And Development , 2 2019., pp. 31–46, https://pubhtml5.com/lwrb/xykg
12. FEHÉR, András Tibor: Artificial intelligence in cyberspace (Part 2) - realizing cyber attacks with AI. RED - American Journal Of Research Education And Development , 4 2019., pp. 27–41, https://online.pubhtml5.com/lwrb/eudn/#p=27
13. NÉGYESI, Imre - FEHÉR András: The relationship between human intel-ligence and artificial intelligence. RED - American Journal Of Research Education And Development, sz. 3., 2020., http://red.devlart.hu/issues/2020_3.pdf
14. FEHÉR, András - NÉGYESI Imre: Cyber security of the Controller Pilot Data Link Communications (CPDLC) system. RED - American Journal Of Research Education And Development, sz. 2., 2023., pp. 29–36, http://red.devlart.hu/issues/2023_2.pdf

AUTOBIOGRAPHY

Lieutenant Colonel András Tibor Fehér was born on 28 November 1972. He grew up in Szentendre, then settled in the village of Kosd near Vác, where he lives with his wife and three children.

His qualifications are in line with his chosen subject, as he mentions in the introduction to his dissertation draft. He obtained his military qualification as a locator technician at the Miklós Zrinyi National Defence University, János Bolyai Faculty of Engineering in 2003. At that time, he was already a certified IT teacher, having graduated from the Faculty of Natural Sciences at Eötvös Lóránd University in 1998. In parallel with his fifth year at ETLE, he began studying theology at Pázmány Péter Catholic University, where he obtained a Baccalaureatus Theologiae degree in 2002, followed two years later, in 2004 – while serving as a lieutenant – he obtained a Licentiatus Theologiae degree (this was a special academic degree preceding a doctorate in

theology, but at that time his studies were interrupted due to the birth of his children and his work). Based on his qualifications, he is a military officer, IT teacher and Catholic theologian. His thesis also reflects the interests arising from these qualifications.

He began his teaching career in 1998 as an IT teacher at a 12-year private school, then from 2004, alongside his military duties, he taught IT as a part-time lecturer at the Budapest College of Communication and Business, and then, when it was renamed, at Metropolitan University, until 2018. Meanwhile, he took up a practical teaching position here in 2015 and began his doctoral studies at the Doctoral School of Military Science in 2019, becoming a teaching assistant later that year. He has an intermediate C (complex) language exam in English and a basic military Arma language exam in French and German.

His military career began more than twenty years ago, in 2003, at the Hungarian Defence Forces Budapest Garrison Command, where he served as an IT officer in the command staff. With the merger of the unit in 2007, he became a planning officer in the Hungarian Defence Forces Support Brigade, where he was responsible for IT tasks, primarily the unit's IT and security responsibilities, until 2011. He then served as the head of communications and IT at the Ministry of Defence's Institute and Museum of Military History until 2015, where he was responsible for the museum's IT operations until he took up his current position. He has been teaching college and university students, as well as course participants, at the HHK IT Department for 10 years now, offering various courses on programming, databases, hardware, operating systems, military IT and security.

In addition to his publications, he is also a regular contributor to the NKE Governance and Science Blog. He is invited to speak at numerous venues and has given 18 conference presentations over the past few years at the National University of Public Service, the Budapest University of Technology and Economics, Pázmány Péter Catholic University , SzTAKI , and other venues, as well as at various workshops. He has also given interviews, e.g. to the honvédelem.hu portal and on Ludovika TV.

Although his interest has recently shifted towards the technical implementation of AI models, his published scientific work has so far focused on the theoretical aspects of artificial intelligence, always taking defence considerations into account. Some of these did not fit into his dissertation. Examples include his monograph on the relationship between pedagogy and machine learning, and his publications on the use of artificial intelligence in cyberspace for attack and defence purposes.