THE WASP-ED AI CURRICULUM: A HOLISTIC CURRICULUM FOR ARTIFICIAL INTELLIGENCE

H. Lindgren[1], F. Heintz[2]

[1]*Umeå University (SWEDEN)*
[2]*Linköping University (SWEDEN)*

Abstract

Efforts in life-long learning and competence development in Artificial Intelligence (AI) have been on the rise for several years. These initiatives have mostly been applied to Science, Technology, Engineering and Mathematics (STEM) disciplines. Even though there has been significant development in Digital Humanities to incorporate AI methods and tools in higher education, the potential for such competences in Arts, Humanities and Social Sciences is far from being realised. Furthermore, there is an increasing awareness that the STEM disciplines need to include competences relating to AI in humanity and society. This is especially important considering the widening and deepening of the impact of AI on society at large and individuals.

The aim of the presented work is to provide a broad and inclusive AI Curriculum that covers the breadth of the topic as it is seen today, which is significantly different from only a decade ago. It is important to note that with curriculum we mean an overview of the subject itself, rather than a particular education program. The curriculum is intended to be used as a foundation for educational activities in AI to for example harmonize terminology, compare different programs, and to identify educational gaps to be filled. An important aspect of the curriculum is the ethical, legal, societal aspects of AI and to not limit the curriculum to the STEM subjects, instead extending to a holistic, human-centred AI perspective.

The curriculum is developed as part of the national research program WASP-ED, the Wallenberg AI and transformative technologies education development program.

Keywords: Artificial Intelligence, Human-Centred Artificial Intelligence, Artificial Intelligence Education, Curriculum.

1 INTRODUCTION

New technologies are needed that can contribute to addressing societal challenges relating to health and managing diseases, education, democracy, environment, food production, energy supply and transportation, promoting a sustainable society. Artificial Intelligence (AI) has become a fundamental component of technology infrastructures and transformational change in society, leading to the increased need of AI knowledge and competence in a broad range of disciplines. Efforts in life-long learning and specifically competence development for professionals in AI have been on the rise for several years. However, these initiatives have mostly been applied to Science, Technology, Engineering and Mathematics (STEM) disciplines in order to provide industry with the competences that are required to advance the AI-supported technology infrastructures and innovation. AI technology is increasingly affecting people's daily work and leisure life, becoming more and more interactive and proactive. Complementary competences other than STEM is needed to meet the societal and individual needs.

Even though there has been significant development in Digital Humanities to incorporate transformative technologies in higher education, including AI methods and tools, the potential for such competences in Arts, Humanities and Social Sciences is far from being realised. The medical field relies heavily on data and fundaments of medical and health knowledge and expertise, where AI techniques are only beginning to be applied to provide decision support and personalised treatment. Furthermore, there is an increasing awareness that the STEM disciplines need to include competences relating to how AI will be integrated in society and affect society and individuals.

The aim of the presented work is to provide the foundations for educational activities in AI in Sweden by identifying and agreeing on the subject matter content and based on this developing a WASP-ED[1]

[1] The Wallenberg AI and Transformative Education Development Program (WASP-ED) is a research program aimed at advancing AI education through a national collaboration across universities: https://wasp-ed.org/.

curriculum for AI in the style of the ACM Computer Science Curriculum[2]. The curriculum is intended to be used as a foundation for educational activities in AI to for example harmonize terminology, compare different programs, and to identify educational gaps to be filled. An important aspect of the curriculum is the ethical, legal, societal aspects of AI [1,2,3]. Therefore, we frame this extended scope by referring to a *Holistic* curriculum for AI, including *Human-Centred AI* (HCAI). Definitions of HCAI have recently been proposed, such as the following by Nowak and colleagues [4] who term HCAI as AI that:

"…*focuses on collaborating with humans, enhancing their capabilities, and empowering them to better achieve their goals*."

While this definition covers the human perspective on AI, we aim in this work adopt an even more holistic perspective. Consequently, it is important to make explicit that the curriculum is intended to be much broader than what can fit in a particular education program and should rather be the union of everything that is covered in different educations [5].

Today most universities have courses in AI, but the content in these vary greatly [5]. This makes it hard for both students and employers to understand what a course contains, or a person knows based on course names. This also makes it challenging to compare what is being offered at different universities and to identify gaps and opportunities in the existing offerings. This becomes even more challenging when including courses that aim to provide knowledge in the broad topics of humans and AI, AI in society, and similar. The proposed curriculum is intended to address these concerns.

2 METHODOLOGY

An investigation was made of categorizations of AI topics in the field of AI and related applied areas. The most influential international research conferences and the Computer Science Curriculum developed by the Association for Computing Machinery (ACM) were reviewed. The ACM Computing Classification system[3] is also developed by ACM and is used in scientific conferences organized by ACM. The ACM classification contains more than 200 topics and was reviewed along with topic classifications of other international AI conferences. More than 320 topics were identified, which were further classified in this work. Among these, there was a substantial number of overlapping concepts, or specializations of a more generic topic.

Further, course curriculum of AI courses spanning across education programs and universities in Sweden were reviewed, selected based on an inventory made by extending the Lund University inventory [5].

The data was analyzed along four dimensions:

1. subject matters and their classifications;
2. target groups and advancement levels;
3. fundamental vs. applied perspectives; and
4. interdisciplinary vs. disciplinary education.

An initial and preliminary description of a high-level curriculum was formed. The preliminary curriculum was discussed broadly along the four dimensions with researchers, teachers, and developers of higher education across disciplines and universities in Sweden. An inventory of AI courses across 10 universities in three countries was also used to evaluate the initial version [5], by mapping courses along the themes of the curriculum, and the curriculum was adjusted based on the evaluation.

For the purpose of framing the topics within an understanding of how AI can be, and already is, embedded in human activities and in societal and technical infrastructures, a theoretical framework is defined, which is summarized in the following section.

2.1 Theoretical framework for understanding and developing AI technology in a humane world

To place AI technology and related knowledge and skills in a human-centred framework, which provides the fundamental aspects of the human, the society, and the environment in which the human is active, we adopt a model comprising four perspectives, which were formed based on theories and models of

[2] https://www.acm.org/education/curricula-recommendations

[3] http://dl.acm.org/ccs.cfm

human function, activity, occupation, and motivation [6, 7, 8, 9]. The theories are commonly applied in domains where technology is studied in a human context, e.g., human-computer interaction, medicine, health, and in education.

Starting with the human at the centre as the first perspective, fundamental aspects affecting how AI technology is developed and used relate to the following four domains of meaningful human activity: *Being*, *Doing, Relating*, and *Becoming* [8]. On a functional level, the human applies *physical*, *social*, *cognitive*, and *mental* functions and abilities (e.g., [10]), which makes the human necessarily *bodily connected* to their mental processes (i.e, *intelligence*) as well as environments. Bodily connected are also human emotions and creativity. These capabilities of being and doing meet the basic human needs *Autonomy* and *Competence* [6]. Further, they are tightly intertwined and cannot be separated if human activity is to be understood by e.g., an artificially intelligent system embedding a *Theory of Mind* [11] similar to how humans reason about and understand other humans' intentions and motives. These basal functions of a human are also blueprints and inspiration to develop *artificial* intelligence, embodied as robotic artefacts.

Further, from a second perspective, the human as member of a society is surrounded by immediate family, friends, other human beings, animals, things and nature – things to care for and to take care of. This relates to being socially connected, enacting relationships and fulfilling the basic human need *Relatedness* [6].

Beyond this as the third perspective, there are social constructs - things to care for and shape as citizen/member of organisations and society, and which function as constraints for human and non-human activity. Some are abstract, mental constructs such as society, formal public and private organisations (e.g., NGOs, industry), informal sub-communities, laws, values and norms (e.g., democracy), scientific and other knowledge (e.g., disciplinary fields). Other are physical, such as infrastructures including streets, public transportation, communication channels, Internet, IoT. Some social constructs can be both physical and abstract, for instance, money and innovations (another kind of value we associate with a thing).

A fourth perspective is needed, to encompass the perspective of *Becoming* from a personal and societal perspective. This perspective is crucial to capture the human's continuous development, and their development of tools and shaping their environments while doing [7, 8]. On the individual level *becoming* relates to knowledge, dreams, fantasies, scenarios of the future, values, which are rooted in and sprung from their past experiences. In the societal perspective, the becoming is intertwined with the history of society and mankind, our tools and collective visions and memories.

Relating AI technology to these human-centred perspectives could be correspondingly referred to as the following matching layers:

- Layer 1 - Core functionality: The artificially intelligent agent, including its functionalities, capabilities and limitations: physical, computational (cognitive), social (technically and computationally connected);
- Layer 2 - Socially and personally embedded: Surrounded by human and artificial agents, animals, things, nature – things to take care of, and to different degree being socially connected and enacting relationships;
- Layer 3 - Socially affecting and socially constrained: Constrained and guided by social constructs such as things to care for and shape as constructed part of organisations and society. These function as constraints for human and non-human activity, and in this context the AI technology:
    - Abstract: Society, formal public and private organisations (e.g., NGOs, industry), informal sub-communities, laws, values and norms (e.g., democracy), scientific and other knowledge (e.g., disciplinary fields) (socially and computationally connected with and part of);
    - Physical: Infrastructures such as streets, public transportation, communication channels, Internet, IoT (technically and computationally connected and part of).
- Layer 4 – Emerging/becoming: How and why AI technology develops over time.

AI technology with its core technology developed on Level 1, can be further developed and applied to build infrastructures and things such as autonomous systems or knowledge in a scientific field on Level 3, or applied to serve a human or social purpose on Level 1 or 2, constrained by Level 3. To span such a broad range of purposes puts particular demands on the contents of an AI curriculum. For instance, *How deep understanding of the core technology on Level 1 is required to evaluate the effects of the technology on the levels 2 and 3? How deep and broad knowledge of the levels 2 and 3 is required to develop core technology that will be useful on these levels?* All people are humans, with personal experiences of the four layers, however, to contribute professionally to the development, implementation and evaluation of AI technology in society, some expertise needs to be developed encompassing deeper knowledge in a broad range of fields.

3 RESULTS: MAPPING OUT AN AI CURRICULUM

The relationships between the human and the AI technology can vary depending on the situation. Why we need to speak of relationships in the context of AI technology is evident in the discussions and envisioned use of AI technology as presented in research and popular media since AI was termed. The roles of AI technology have been described as a tool, prosthesis, symbiotic agent, companion, partner, coach, co-worker, and for the (co-)designer and developer - object of construction and innovation [12,8,4].

The agent could enact relationships on the functional level (Level 1), through invasive techniques such as Brain-Computer Interfaces or robotic prostheses connecting to muscles or nerves. The social relationships are typically enacted on Level 2, and by enacting relationships, the basic humane functions and abilities are affected also on Level 1, the same way as they are by enacting relationships with other humans and things found on Levels 2 and 3. Increasing interest and focus is seen in society as well as research in a broad range of scientific domains on such effects, e.g., on human decision making, attention, and social behaviours. Likewise, the AI agent may be affected through enacted relationships, for instance, through reinforcement learning techniques. Moreover, AI technologies also enacts roles in connection to Level 3, typically as part of infrastructures, but also increasingly often in situations where it needs to be governed by societal regulations and informal social norms.

How do we organise scientific fields and subject matters along these levels? In a simplified perspective, the medical field focuses on the first level, the individual and their functions and dysfunctions, pathology and diseases, and how these are affected and spread through Level 2. Mainly, STEM disciplines concern Levels 1 (biological matters) and 3 (abstract mathematical constructs or physical/digital infrastructures). Humanities and social sciences concern Level 1, 2 and the abstract level 3, relating to societal and humane aspects and perspectives.

Another factor to embed when developing new knowledge through research and when educating students, is the scientific traditions and methodologies, which can be divided into two major strands:

i) the empirical studies of matters as they function in today's society, e.g., existing AI technology as it is used today, and

ii) the constructive studies of mechanisms for development and change, by intervening, and study the effects of the intervention, e.g., the engineering disciplines where new AI technology is being developed and its application is studied.

Across both strands and fields of research, frameworks of how to address phenomenon are developed and applied, for instance, critical perspectives to challenge norms, structures and phenomenon. Epistemology, methodology and theoretical frameworks are captured in the curriculum as a basis that is embedded across all topics, and which is expected to blend in the transdisciplinary endeavours.

A third factor is whether AI technology is approached as a tool or a subject of study and development. In the analysis of existing courses, it was found that a substantial part was developed for the purpose to introduce AI technology as a method, or a tool in particular disciplines, or research fields.

The proposed AI curriculum is organised along the theoretical framework into the following complementary layers (Fig. 1):

- Layer 0 represents the fundaments of knowledge such as epistemiology, methodology, and theoretical frameworks, which are applied across all education.
- Layer 1a includes the core AI theory and technologies, which are already well covered in existing AI education, summarised into nine key AI topics.
- Layer 2a includes topics relating to Humans and AI from the individual's perspective, in particular *Human-AI Interaction and Collaboration*. This also includes an individual's participation in the development of new AI systems.

- Layer 3a includes topics relating to societal and humane aspects on AI, and critical perspectives relating to these, which we summarise as *ELSEC* (ethical, legal, social, economic and cultural aspects of AI), and *Trustworthy AI, f*or which an educational initiative is being developed based on the work of the EU High-Level Expert Group [3,13].
- Layer 1b, 2b, 3b complements the three layers 1a, 2a and 3a, and includes AI technologies that are in today's society mainly applied to contribute to technical infrastructures (*Distributed and Edge AI*, and *Robotics, Control and Autonomous Systems*), and which are also well covered in existing AI and CS education.
- Layer 3c includes AI technologies applied as method for i) conducting research in Medicine, Social Sciences, Humanities; ii) addressing societal needs and challenges: improving health, decision making, warfare/defence, etc.; iii) managing societal infrastructures: business, transactions, public transportation, etc.
- Layer 4 addressing the history of AI, and the futures of AI.

The emerging concepts *Hybrid AI* and *Sustainable AI* are increasingly being used, however, have different meanings depending on in which contexts they are used.

Hybrid AI can in one hand, refer to Level 1 where combinations of AI theories and techniques are fused to build more intelligent systems, and in the other hand, refer to when the levels are merged to include both artificial and human intelligence [14].

Sustainable AI is often discussed in the context of the UN sustainability goals (e.g., [15]), and is in this context referring to all layers to different extent depending on the sustainability goal, however, the outset is based on purposes defined in Layer 3.

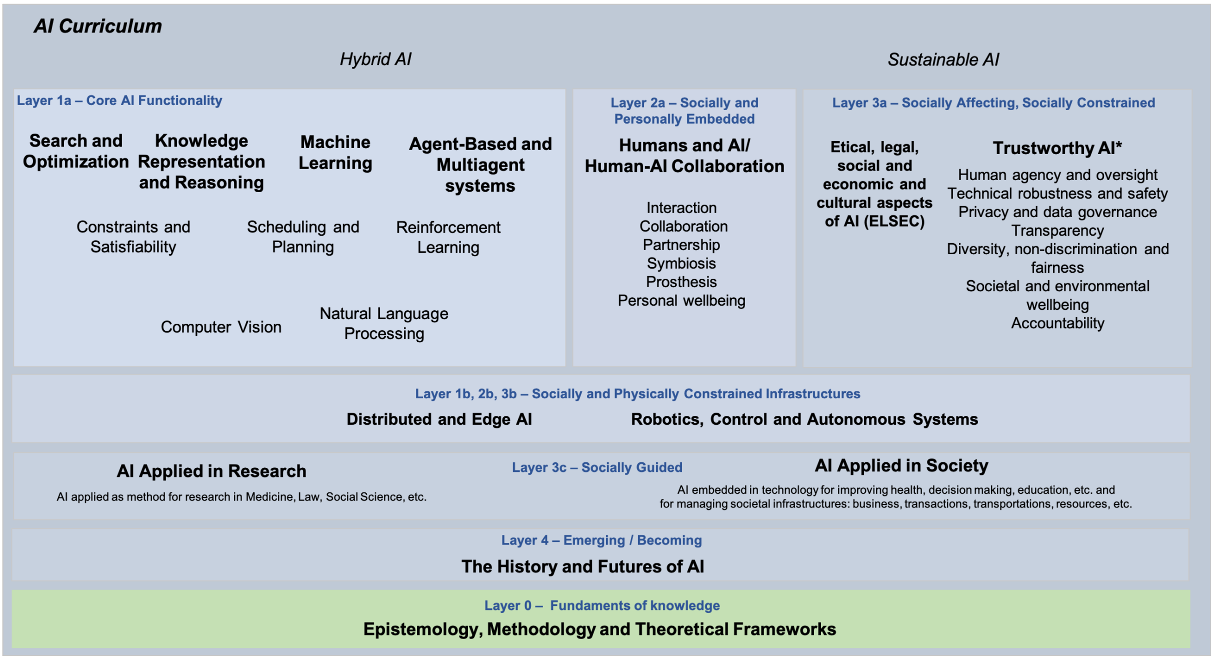

*Figure 1. The layers of perspective on AI, with identified topics to be included in a holistic curriculum for AI. * The content of Trustworthy AI builds on the educational initiative [13] stemming from [3].*

To summarise, as illustrated in Fig 1, 18 topics were selected as key topics to span the different human-centred AI perspectives outlined in this work. Additional 14 concepts were included to exemplify more specific topics in the human and societal perspectives. Two concepts were also included that illustrate the emergent character and general need for transformative and sustainable AI.

4 CONCLUSIONS

A holistic AI curriculum was developed based on human-centred perspectives on how AI may be integrated in human and societal activities and organisations. The proposed curriculum includes the following: Core AI functionality; how AI is or can be socially and personally embedded; how AI affects society, and is, could or should be socially constrained; socially and physically constrained technical infrastructures embedding AI; how the development of AI is socially guided to become a useful tool for research and society; and finally, how AI has emerged, and may continue to develop in different potential AI futures. In total, these layers contain 18 distinct top-level topics. Most of these are topics that can be found both in existing education programs and in the scientific conferences and journals of the field, indicating that the curriculum is covering the content of what is considered AI today, which is significantly broader than only a decade ago.

The next step is to describe the content of the topics, the dependencies and progression between the topics, and to identify subsets of topics providing suitable educational pathways for different practitioners and professions.

The curriculum will be used as a foundation for educational activities in AI to harmonize terminology, compare different programs, and to identify educational gaps to be filled.

ACKNOWLEDGEMENTS

Research is funded by The Wallenberg AI and Transformative Education Development Program (WASP-ED).

REFERENCES

[1] B. Friedman, P. H. Kahn, and A. Borning. “Value Sensitive Design and Information Systems, ” *Human-Computer Interaction in Management Information Systems*: Foundations. ME Sharpe, Inc: NY, pp. 348-372, 2006.

[2] M. van der Velden and C. Mörtberg. “Participatory Design and Design for Values, ” In: (van den Hoven, J., Vermaas, P., van de Poel, I. eds) *Handbook of Ethics, Values, and Technological Design*. Springer, Dordrecht, pp. 1–22, 2021.

[3] High-Level Expert Group on AI (AI HLEG). 2018. “Ethics Guidelines for Trustworthy AI, ” Report. European Union. https://ec.europa.eu/futurium/en/ai-alliance-consultation.1.html

[4] A. Nowak, P. Lukowicz, and P. Horodecki. “Assessing Artificial Intelligence for Humanity: Will AI be the Our Biggest Ever Advance? or the Biggest Threat [Opinion], ” *IEEE Technology and Society Magazine*, vol. 37, no. 4, pp. 26–34, 2018.

[5] M. Ekblad and M. Sundström. 2021. AI-education at Lund University. Report 20211210. Lund University. https://www.ai.lu.se/sites/ai.lu.se/files/2022-01/AI%20education.pdf

[6] R. M. Ryan and L. M. Deci. “Self-Determination Theory and the facilitation of intrinsic motivation, social development and well-being, ” *American Psychologist, vol.* 55, pp. 68–78, 2000.

[7] L. S. Vygotsky. “*Mind in society: The development of higher mental processes*, ” (M. Cole, V. Jolm-Steiner, S. Scribner, & E. Souberman, eds.). Harvard University Press, 1978.

[8] V. Kaptelinin and B. A. Nardi. “*Acting with technology: Activity theory and interaction design*, ” MIT press, 2006.

[9] K. W. Hammell. “Belonging, occupation, and human well-being: An exploration, ” *Canadian Journal of Occupational Therapy*, vol. 81, pp. 39 – 50, 2014.

[10] World Health Organisation. “The International Classification of Functioning Disability and Health, ” https://www.who.int/standards/classifications/international-classification-of-functioning-disability-and-health.

[11] D. Premack and G. Woodruff. “Does the chimpanzee have a theory of mind?” *Behav. Brain Sci.* vol. 1, pp. 515–526, 1978.

[12] T. W. Bickmore and R. W. Picard. “Establishing and maintaining long-term human-computer relationships, ” *ACM Transactions on Computer-Human Interaction* (TOCHI), vol. 12, no. 2, pp. 293–327, 2005.

[13] J.C. Nieves and A. Aler Tubella, “Framework for Trustworthy AI Education V1.0,” The Trustworthy AI Project, September 2021. https://www.trustworthyaiproject.eu/framework-for-trustworthy-ai-education/

[14] Z. Akata, D. Balliet, M. de Rijke, F. Dignum, V. Dignum, G. Eiben, A. Fokkens, D. Grossi, K. Hindriks, H. Hoos, H. Hung, C. Jonker, C. Monz, M. Neerincx, F. Oliehoek, H. Prakken, S. Schlobach, L. van der Gaag, F. van Harmelen, H. van Hoof, B. van Riemsdijk, A. van Wynsberghe, R. Verbrugge, B. Verheij, P. Vossen, and M. Welling. "A Research Agenda for Hybrid Intelligence: Augmenting Human Intellect With Collaborative, Adaptive, Responsible, and Explainable Artificial Intelligence," *Computer,* vol. 53, no. 8 (aug), pp. 18–28, 2020.

[15] S. Umbrello and I. van det Poel. "Mapping Value Sensitive Design onto AI for Social Good Principles," *AI and Ethics*, vol. 1, pp. 283–296, 2021.