IEEE SA STANDARDS ASSOCIATION

The Benefits of a Multidisciplinary Lens for Artificial Intelligence Systems Ethics

A Primer for Education Thought Leadership

IEEE

Table of Contents

Committee Members **3**

Abstract **4**

An Overview of IEEE's *Ethically Aligned Design* **5**

Introducing the IEEE Global Initiative Education committee **6**

Domain of AIS Ethics Education **6**

The Role of Multidisciplinarity in AI Ethics Education **10**

Choice of Collaborative Approach and Disciplinarity **15**

Centering AIS Ethics Around Human Design Thinking **19**

Conclusion: The Sanctity of the System **22**

Appendices

Appendix 1: AIS Landscape 24

Appendix 2: AIS and Governments: Supranational and National Level 28

Appendix 3: Figure from IEEE Std 7010™-2020 29

Additional Resources **30**

Committee Members

Lead Authors:

- Melodena Stephens
- Marisa Zalabak

Co-Authors:

- John C. Havens
- Hagit Messler

Contributors:

- Masheika Allgood
- Rosalyn W. Berne
- Sinead Bovell
- Larry Bridgesmith
- Rafael A. Calvo
- Ann Hill Duin
- Jeff Felice
- Jordan Harrod
- Kohei Kurihara
- Isabel Pedersen
- Abel Pena-Fernandez
- David Ryan Polgar
- Edson Prestes
- Jessica Whitaker

Abstract

This paper serves as an introductory overview to provide information and insight on current trends about how to best educate on subjects relating to the ethics of designing and using artificial intelligence system(s) (AIS).[1] As *AI ethics educators* (a term used throughout this paper to designate our committee members working in various academic, corporate, and policy-based realms at the front lines of creating awareness on how to implement AIS in an ethical manner), utilize a globally-adopted approach such as IEEE's Ethically Aligned Design (EAD) and infuse these principles within our coursework.

A key insight leading to the creation of this committee came in the recognition that teaching elements of ethically oriented AIS issues such as bias, fairness, risk, and explainability was only the first step in truly educating students and the world at large about the power of these systems. As some of the first educators teaching these subjects in the modern age of artificial intelligence (after the various lulls in the development of AI technology—for example the AI winters of 1973 and 1988), we have had the good fortune to analyze, identify, and adjust teaching based on how the concepts surrounding these ideas were and are originally received by various students in multiple contexts around the world. In this regard, it was from listening to, communicating with, and observing students that we have gained the greatest perspective on how AIS must be addressed.

There are several basic aspects to consider in the expansion of AI ethics education, beginning with the lack of an agreed-upon definition of *artificial intelligence* (see Appendix 1). The landscape of the efforts in the field of AI ethics education—as well as the diverse perceptions regarding governance and regulation (see Appendix 2)—has also informed this perspective.

Concerns regarding how human data is utilized in a majority of AI algorithms or how AIS often are used to analyze human behavior demonstrate these tools are often mistrusted or are adopted without detailed consideration, simply because they are new to students or users. For most people, in general, values reflect or drive behavior; however, where machines or algorithms identify aspects of their identity that people have not claimed as their own, influencing their decision-making, AI systems can seem invasive or worthy of distrust. These are all key elements for AI ethics educators to be mindful of in their work; we offer many insights along these lines in the pages that follow.

We further acknowledge that in addition to the engineers involved with AIS, the responsibility of design is spread across various stakeholders: the investors, the users, the recommenders, the regulators, and the educators. Therefore, we begin to view AIS through a multidisciplinary lens. In this document, we share some of our achievements, best practices, and concerns with the goal of helping others to develop a global society of ethically aligned design for AIS. Additional resources and papers will follow.

[1] While we may use the term *artificial intelligence* in this paper to honor the historical introduction of the term, we prefer and recommend the term *artificial intelligence systems* (AIS), as defined by the Organisation for Economic Co-operation and Development (OECD): "An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy."

An Overview of IEEE's *Ethically Aligned Design*

Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems is a landmark report created by more than 700 volunteers in a process supported by IEEE from 2016 to 2019. In three versions, with more than 500 pages of feedback, *Ethically Aligned Design* provides more than 300 pages of issues and recommendations, providing a bedrock of methodologies, insights, and tools surrounding how best to honor human values, human rights, and environmental sustainability in the design and use of AIS.

The mission of The IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (or The IEEE Global Initiative, a group within Industry Connections—a program of The IEEE Standards Association—that created *Ethically Aligned Design*) is "to help ensure every stakeholder involved in the design and development of autonomous and intelligent systems is educated, trained, and empowered to prioritize ethical considerations so that these technologies are advanced for the benefit of humanity." The term stakeholder refers to anyone involved in the research, design, manufacture, or messaging around AIS, including the universities, organizations, and corporations making these technologies a reality for society.

Quotes from *Ethically Aligned Design* are offered throughout this paper based on their specific, enduring guidance for educators and technologists alike.

Introducing the IEEE Global Initiative Education Committee

The IEEE Global Initiative Education Committee operates within the broader mission of The IEEE Global Initiative. Over five years, the Committee brought together more than 50 educators and practitioners with diverse backgrounds and experiences from across the globe to share and develop best practices.

Hagit Messer, Chair of the Education Committee and Kranzberg Chair in Signal Processing, School of Electrical Engineering, Tel Aviv University, reflects this focus:

"I see the future of AIS technologies through glasses with two lenses: awareness and responsibility. Users—at all levels—of such systems must be aware of their ethical implications when deciding to use them."

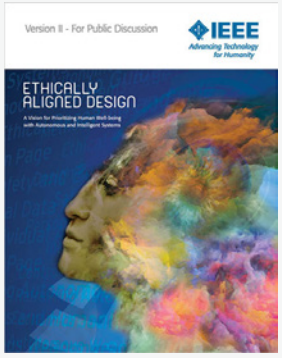
The committee contributors to this paper welcome feedback and experiences from others to enrich the community as we advance the education of today's and tomorrow's AIS influencers and practitioners.

Quotes from committee members are also listed throughout this paper to provide first-hand expertise for educators, trainers, and practitioners on the front lines of AI ethics education.

DOMAIN OF AIS ETHICS EDUCATION

"As engineering is taught as a collection of techno-science, logic, and mathematics, embedding ethical sensitivity into these objective and non-objective processes is essential. Curriculum development is crucial in each approach. In addition to research articles and best practices, it is recommended that engineers and practitioners come together with social scientists and philosophers to develop case studies, interactive virtual reality gaming, and additional course interventions that are relevant to students."

—"Classical Ethics in A/IS," *Ethically Aligned Design* (p. 45)



Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Artificial Intelligence and Autonomous Systems

AIS ethics is not the sole domain of engineers, technologists, or ethicists. The domain of AIS encroaches on various disciplines other than engineering like science fiction, art, education, health, social sciences, law, procurement, management, and even military strategy. It is impossible for one discipline alone to grasp all the possible areas in which AIS ethics may be applicable, as the value chain is complex.

Isabel Pedersen, Committee member and Professor of AI and Ethical Implications of Technology at Ontario Tech University, states:

"We need to use much longer forecasting strategies to identify unintended consequences, and we need to teach our students to also work on strategies for mitigation and reversal of unintended negative effects of AIS."

Innovation is fueled by openness to cognitive diversity. A room full of similar thinkers may not imagine beyond status quo ideas. When creating AIS, ethical design as a methodology has empowered transformative changes in products, brands, organizations, industry, and government due to a form of participatory design analyzing end-user values at the outset of manufacturing in ways not utilized before.

This journey of recognizing the need for and implementing ethically aligned design (the methodologies espoused in *Ethically Aligned Design*) is one which requires listening and actively seeking feedback. Providing accountable methodologies, instituting this feedback is necessary due to the disparity between market demands, investor needs, planetary sustainability, and human rights. A common set of tools is required to identify the metrics of success for AIS creation so systems do not harm people or the planet while value is still provided to customers or end users from AIS products and services.

As Committee member and Professor Edson Prestes, Federal University of Rio Grande do Sul, notes:

"Being part of and contributing to these debates on AIS ethics involves, first, humility to understand we do not have all answers (and of course no one has), and we need to be open to new domains and new ideas, which can be too far from our formal training, to understand the technology under different points of view."

Ethics as a topic has been studied for centuries in non-AIS (and nontechnical) contexts. Mirroring this logic, to benefit from cross-pollination-based expertise, encouraging collaboration among students, instructors, and professionals across generations, disciplines, industries, and countries would involve inspiring camaraderie and innovation while providing methods and tools for challenges to consensus in atypical collaborations.

Professor Rafael A. Calvo, another Committee member and esteemed academic at Imperial College London, states:

"The ethics curriculum for engineers has been about managing risk and conflicts of interest, being diligent, and promoting safety. These are important, but do not show our outcomes as the sociotechnical products they are."

Given the effects of AIS on humans individually and collectively and on how they choose to interact with each other, it is also vital to consider psychological and social well-being and the ramifications of unethical practices or products. This focus will help engineers and technologists benefit from the social, political, and philosophical expertise they do not possess as part of their professional backgrounds.

As educational psychologist, founder of Open Channel Culture, and Committee Co-Chair Marisa Zalabak says:

"There are many ways that AIS can contribute to the well-being of humanity, yet the psychological and neurological consequences of AIS should be considered more seriously. The issue of racial and social bias built into the systems is one major example with both individual and social negative consequences."

It is apparent that there are significant numbers of people who feel they are excluded from the process of developing and deploying AIS.

David Ryan Polgar, Committee member and founder of AI Tech is Human, says:

"There is still a large disconnect between people educating future responsible technologists and how the industry creates AIS in isolation."

Part of the challenge is that AIS is largely unseen in its algorithmic form; embedded in everyday products, the consequences of the outputs of the machines and systems utilizing those algorithms are not always clear to end users.

Sinead Bovell, Committee member and WAYE founder, explains this further:

"As a youth, I know that the impact of today's inventions will be most strongly felt by tomorrow's users. The outcomes of our autonomous and intelligent systems are a reflection of our priorities. If ethically aligned design is not prioritized, we cannot expect ethically aligned outcomes; and we run an extremely high risk of perpetuating societal harms. We have to actively choose to make our systems fair, explainable, safe, and free from discrimination and societal harm. A passive approach to ethical design won't get us there."

Although ethics is often associated with law, many acts that would be widely condemned as unethical are not prohibited by law, and the complexity of ethics calls for a multidisciplinary approach. In relation to AIS ethics, the responsibility of ethics can lie with the individual or the organizations working with AIS.

As Committee member and CEO of DASH4Law, Larry Bridgesmith, notes:

"The ethical development of AIS applications is a multifaceted problem requiring a multidimensional solution."

Rosalyn W. Berne, Committee member, and Anne Shirley Carter Olsson, Professor of Applied Ethics, School of Engineering and Applied Sciences, University of Virginia, explains this further, pointing out the need for multidisciplinary thinking to avoid the unintended consequences of design from only one group of technologists or designers of AIS.

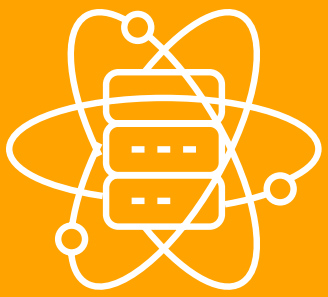
Rosalyn W. Berne, Committee member says:

"Consumers, manufacturers (everyone in the supply chain of a product), funders, policy makers, etc., all have a role to play in the development and use of engineered products. But ethically aligned design requires commitment, engagement, and awareness on the part of engineers. If the engineers don't prioritize ethically aligned design, then the pressures of other priorities such as cost, return on investment, and efficiency can easily take precedence and lead to disregard for nodes of ethical concern."

The Role of Multidisciplinarity in AI Ethics Education

The role of multidisciplinarity in AI ethics education is focused on the need for communication between the various stakeholders creating, marketing, and using AIS. Ann Hill Duin, Committee member and Professor of Writing Studies at the University of Minnesota, highlights the importance of the topic.

We are in the midst of the reinvention of all fields as big data and academic analytics, artificial intelligence, autonomous writing agents, and augmented/virtual reality become commonplace amid pervasive information and tracking. I often share with students that "the technical communication future is not your advisor's workplace." We must instill in students a depth of disciplinary understanding along with ethically aligned design for AIS to enable them to address grand challenges of the present and the future.



In many companies, the methodologies and taxonomies engineers use are wildly disparate from ontologies or practices of various disciplines in the same organization. Even when the same language, such as English, is used by everyone in a company, the multiple vocabularies and definitions from various disciplines (including terms used by each of the countries and governing agencies involved with the organization) can feel as if everyone is speaking in multiple tongues without an interpreter. But it is in these spaces—where various experts communicate aspects of AIS to one another—that they can best identify how to also communicate to those who use the systems. To avoid distrust of AIS, methodological ways of listening to end users or customers must happen in order to truly hear their concerns and needs in ways that can then be reverse engineered into the design of a product. AIS educators in this regard are supreme communicators over and above any subject matter expertise in machine learning, philosophy, or business.

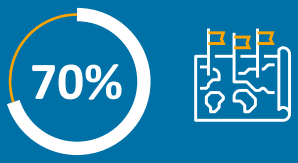
The Organisation for Economic Co-operation and Development (OECD) paper "Recommendations of the Council of Artificial Intelligence" (2019) highlights the importance of the actors and stakeholders utilizing AIS, indicating the need for a wider scope of stakeholder input. This leads to consideration of several disciplinary approaches to the design and development of AIS. Most educators fall under the category of stakeholder as AIS educators influence engineers, managers, and policy makers that invent, work, deploy, and operate AIS. AIS application fields are large (World Intellectual Property Organization, 2019), and until recently these fields—medicine, education, city planning, finance, law, agriculture, and so forth—have also been educational specialization areas (see Figure 1).

Figure 1. The diversity of AI application fields (adapted from WIPO, 2019)

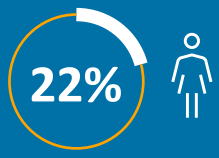
During the first quarter of the global pandemic, when most education went online, there were 936 million education apps downloaded.[2] Who vetted these apps or certified their data privacy terms (especially as students are too young to give consent)? On one hand, there was little choice, as many education ministries mandated that education must continue, which left teachers and the parents on their own to make informed choices. This exposed many gaps and disparities in AIS literacy, in availability and access to reliable information and resources, as well as the inequities in access to technologies. Rampant misuse of student and citizen data was also common, further hindering the trust regarding AIS technologies and furthering the need for sound ethical design and use of AIS. With the scale of AIS adoption increasing exponentially, this precedent revealed the need to quickly fill these gaps in our understanding of AIS.

As AIS continues to become part of daily life for a majority of people around the world, it is critical to embrace an approach toward AIS ethics that incorporates a variety of disciplines and ways of thinking, forming a strong intercultural perspective across geographic boundaries, and a nested approach to sustainable development that embraces a human-centric and planet-sustainable approach. While sustainability and AI may seem unrelated, methodologies using deep learning have been shown to produce high carbon usage, harming the planet as much as multiple round-trip flights circumnavigating the globe. On the positive side of this equation, AIS can also provide statistical and other benefits, improving supply chain and other issues related to increasing sustainability, in turn, highlighting the vital importance of educating emerging engineers (traditionally and non-traditionally trained) to reframe innovation with a regenerative mindset. The key to any of these discussions from an ethical standpoint is the urgency of the critically important issues of healing the planet for the healthy and equitable flourishing of all people.

The United Nations Educational, Scientific and Cultural Organization (UNESCO) paper "Recommendations on the Ethics of Artificial Intelligence" (2021) highlights disparities in the field where AIS growth is unequal. For example, North America and China dominate in terms of economic impact (70%). There is a gender gap (only 22% of AIS professionals are women).



North America and China dominate in terms of economic impact (70%).



There is a gender gap (only 22% of AIS professionals are women).

Recommendations on the Ethics of Artificial Intelligence by UNESCO

We saw an increasing number of AI journal publications in 2020, which is 5.4 times greater than published in 2000, according to the "AI Index 2021 Annual Report" by Stanford University. However, there is more emphasis on commercialization versus scientific papers—as the ratio of scientific papers to inventions decreased from eight to one (8:1) in 2010 to three to one (3:1) in 2016 according to the WIPO Technology Trends 2019: Artificial Intelligence (Chap 3), raising concerns. The good news is that ethics is increasingly becoming an important topic as shown by the increased number of publications documented by the "AI Index 2021 Annual Report." Therefore, it is critical that education or educators step in, either formally or informally, to make all stakeholders aware about AIS ethics.

[2] Statista, "Global Mobile Education App Downloads by Platform 2020," 2021, https://www.statista.com/statistics/1128262/mobile-education-app-downloads-worldwide-platforms-millions/

Parallel with the advancement of emerging AIS technologies today, regulation and other governance forms are struggling to keep up with the rapid growth and deployment of these systems. Fortunately, widespread regulations like General Data Protection Regulation (GDPR), principles for AI from UNESCO, developing tools on AI risk from the National Institute of Standards and Technology (NIST), and standards from IEEE focused on ethical principles are helping in this space. Largely unaddressed, however, is that without formal fines or genuinely stringent consequences for violations, future attempts by governments and industry to bring oversight will be difficult to maintain with the pace of innovation.

The resulting ethical gap between unregulated AIS and its continued creation is currently being filled by corporations, organizations, and individuals. It is becoming clear that past precedent and use cases alone cannot help us wade through the ethics quagmire we face. This tenuous environment, along with nonregulatory drivers such as accessibility and equality, has amplified the need to develop a culture of ethics by design in the AIS global community. The UNECSO paper, "Preliminary Study on the Ethics of Artificial Intelligence" explains one of the dilemmas.

It is important not to equate experience with intelligence, even though some experts have suggested that recent developments in AI might also be a reason to re-examine the importance of this experience or awareness for being human. If experience is at the core of being human, ethical considerations must ensure that this is protected and enhanced through the use of AI rather than side-lined or disempowered. (UNESCO, 2019, p. 6)

In addition to the form and approach for collaboration, as education is a universal human right, there is an urgent need for all stakeholders and users to understand AIS. There are profound disparities in AIS literacy, highlighting a lack of access to technology for many. Ideally, we should learn using both formal and informal channels using a multidirectional flow—top-down, bottom-up, and across generations (see Exhibit 4).

In the United States (US), there has been a drop in the number of doctoral students who enter into academia—from 42.1% in 2010 to 23.7% in 2019[3]—suggesting that we need to find other methods to get critical perspectives of AIS into the public sphere. Furthermore, top professors in the AIS space are leaving, suggesting a brain drain impacting not only the specialized streams but those fields that need an interdisciplinary understanding of AIS as well.[4]

[3] Zhang, Daniel, Saurabh Mishra, Erik Brynjolfsson, John Etchemendy, Deep Ganguli, Barbara Grosz, Terah Lyons, James Manyika, Juan Carlos Niebles, Michael Sellitto, Yoav Shoham, Jack Clark, and Raymond Perrault, "AI Education," Chap. 4 in "The AI Index 2021 Annual Report," AI Index Steering Committee, Human-Centered AI Institute, Stanford University, p.13, Mar. 2021, https://aiindex.stanford.edu/wp-content/uploads/2021/03/2021-AI-Index-Report-_Chapter-4.pdf.

[4] Gofman, M., and Z. Jin, "Artificial intelligence, education, and entrepreneurship," *Journal of Finance*, Forthcoming, 2020, http://dx.doi.org/10.2139/ssrn.3449440

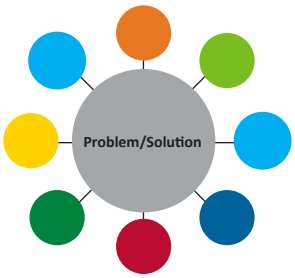
Committee member Jessica Whitaker highlights the importance of EAD in today's world:

"Prioritizing EAD in AIS solutions will prevent offensive, exclusionary, and harmful products. The design outcomes will be more beneficial to a larger population, thus resulting in higher profits. The leadership financing the work must understand these benefits and provide a budget to incorporate EAD. Even as an engineer who has personally experienced the negative impacts of AIS without EAD, it is challenging to convince my peers and other decision-makers of the value. Personal ethics carry over into employment ethics and prevent those who have not experienced adverse circumstances due to the lack of EAD from accepting that it is needed or acknowledging that it is an industry problem."

Choice of Collaborative Approach and Disciplinarity

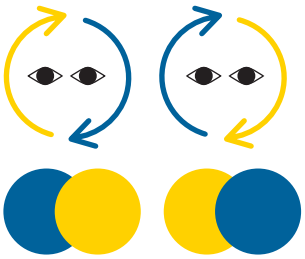
There is a variety of approaches for collaboration between disciplines, within systems, and between systems when discussing AIS ethics. Given the complexity of ethical inquiry involving technical and social considerations, and the distinct differences between the disciplines and perspectives involved, it is important to clarify definitions and present the rationale for the collaborative approach selected by this committee.

Multidisciplinary



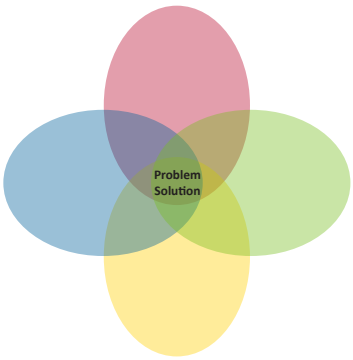
Multidisciplinary describes an additive approach; people from several different disciplines staying within each of their boundaries, working together to involve and consider solving problems alongside disciplines of professional specializations and perspectives.

Cross-disciplinary



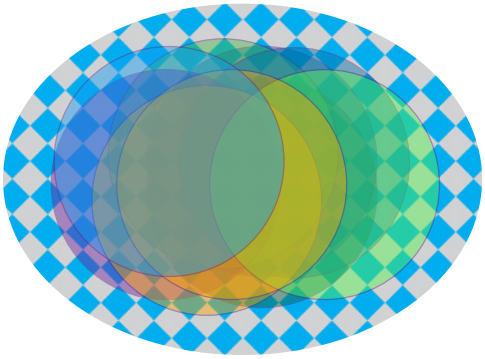
Cross-disciplinary describes an approach that engages cognitive empathy, viewing, and considering from the perspective of another discipline.

Interdisciplinary



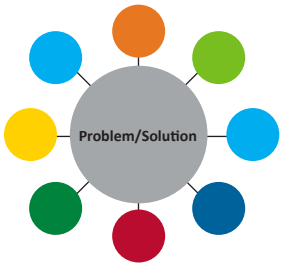
Interdisciplinary describes an approach combining two or more disciplines. Like a Venn diagram, meant to cross boundaries by synthesizing and creating a new level of integration while remaining within disciplinary frameworks. This approach goes beyond the addition of parts or disciplinary perspective-taking, by recognizing that each discipline can transfer methods that affect the output of the other disciplines.

Transdisciplinary



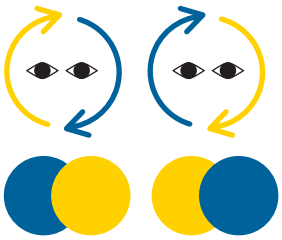
Transdisciplinary—also referred to as a *xenogenesis* or "between, across, and beyond disciplines"—is a term often used in the science of team science. It describes an approach that relates two or more disciplines and branches of knowledge that unify individual disciplines and intellectual frameworks to form an entirely new approach unlike any of the contributing parts. The approach transcends boundaries, and hierarchies, integrating technical, natural, social, and health sciences, as well as informal stakeholder groups, with a goal of understanding the present world (see Figure 2).

DISCIPLINARY COLLABORATION



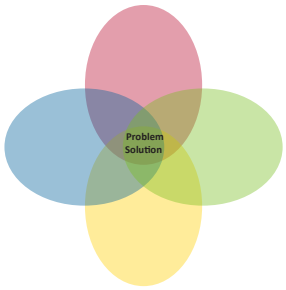
Multidisciplinary

- Integration: Separated, juxtaposing, coordinating
- Perspective: 22 disciplines, individual messages
- Team's Goals: Project
- Leadership: Varied leadership for individual information exchange & coordination



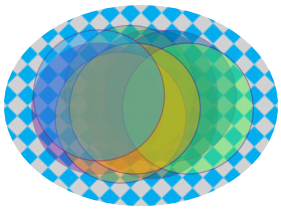
Cross-disciplinary

- Integration: Separated, interacting, coordinating
- Perspective: 2+ disciplines, exchange of disciplinary lenses
- Team's Goals: Project, learning
- Leadership: Varied leadership for individual information exchange & coordination



Interdisciplinary

- Integration: Integrated, interacting, linking, blending
- Perspective: 22 disciplines, combining and blending disciplinary perspectives
- Team's Goals: Learning, new ideas
- Leadership: Varied leadership for common information exchange



Transdisciplinary

- Integration: "Become One," transcending, transgressing, transforming, systematic integration of components
- Perspective: Includes disciplines + stakeholders
- Team's Goals: Problem oriented, creation of new knowledge
- Leadership: Rotating leadership

Collaboration between engineers and social sciences is not new and has proven highly effective and increasingly necessary in AIS design and development. We begin with this multidisciplinary approach to collaboration as a foundation and entrance, moving toward a transdisciplinary approach, increasingly considered to be an effective choice for the emerging future. Key ways an organization can move toward collaboration between and with multiple disciplines, including tools and steps for each group (e.g., academic engineers, educators, business developers, and policy makers), are covered in the following sections.

One of the discussions engineers can have is: How can we reverse engineer technologies to incorporate various AIS ethics concerns? This includes the kind of methodologies found in IEEE Std 7000™.[5]

Committee member and Co-Chair Marisa Zalabak states:

"The complexity of the wicked problems, possible solutions, and risks involved in AIS ethics and practices begs more holistic consideration than is currently found today in most disciplinary fields. The benefit of diverse perspectives and expertise boosts the quality of innovation through the collective intelligence that unfolds in the process… yet, specific and multimodal approaches for collaboration are needed to move beyond individual disciplinary and psychological boundaries to effectively realize multidisciplinary and transdisciplinary collaboration."

[5] IEEE Std 7000, IEEE Standard Model Process for Addressing Ethical Concerns during System Design, is available from The Institute of Electrical and Electronics Engineers (http://standards.ieee.org/).

UNIVERSITY ACADEMICS

University academics can invite colleagues from departments outside of their own to join lectures on AIS and ethics. So that colleagues are not made to feel they are speaking outside of their comfort zone, this does not have to be overly formal.

For a colleague from the engineering school, discussions can start with easy questions such as:

- Based on what you've heard today in our discussions on AI ethics, how do you feel engineers would respond to the need for applied ethics methodologies in their work?
- What made sense to you today?
- What should these students know about engineering when they are thinking about ethics?

STUDENTS

Students can be exposed to a curriculum for ethics and AI technologies. For example, Harvard University has a foundation course Tech Ethics: AI, Biotech, and the Future of Human Nature (Gen Ed 1058). The purpose of such a course is to create awareness of the limitations of technology and the gray areas we have not yet resolved.

Students can address the challenges of AIS ethics presented as cases in existing curricula or discuss the decision flow in the value-chain that lead to poor AI ethics. There are many sources available, e.g., the AI Incident Database.

Book or movie clubs can help augment discussions of AIS ethics. For example, to start the conversation or create debates, you can use science fiction (like Snow Crash by Neal Stephenson), newspaper reports, and nonfiction books (*Weapons of Math Destruction* by Cathy O'Neil) or even watch movies like the Terminator series.

SYMPOSIUMS

Large, diverse groups can be reached by hosting a symposium with experts from multiple schools or from different areas including community groups and agencies on the subject of AI ethics. At the symposium, sample cases with studies surrounding the launch of a fictional AIS product or service can be provided and the symposium experts can address what they feel are the most critical areas of concern for the issues involved. What will likely happen is that a legal expert will discuss areas around compliance or risk, while a communications or marketing expert will discuss how a company's brand may be viewed by customers based on a crisis relating to AI. A moderator in this example can point out the thought process of each expert based on their areas of expertise. Moderators can also ask questions like: How would you describe your approach, views, or expertise to someone with no understanding of your field?

PRACTITIONERS

When deploying the multidisciplinary practice for designing the environment to create ethical decisions, goal setting is the initial process for achievement. Following the goal-setting stage, designing the environment for psychological safety provides opportunities for the participants to share their thoughts or deepen their insight by listening to others. A multidisciplinary approach should include acceptance of any silence of those participants who are not speaking, respecting silent reflection and non-verbal communication (e.g., text, motion, shared images), while paying attention to any verbally undisclosed insights from the subject.

TRAINING

Training provided for groups within disciplines should be optimized for constructive communication as well as with others outside of individual disciplines (e.g. psychological safety, multimodal and social/cultural communication, collaborative agreements, and processing conflict). This includes training in emotional and social intelligences (e.g., self-awareness, empathy, compassion, relational awareness, perspective taking, the ability to navigate differences, and tolerance for ambiguity). It is important to understand that this kind of training is ongoing individually and collectively, to build the sustainability of human interactive skills.

PUBLIC POLICY OFFICIALS

Public policy officials can participate in human-centered design thinking workshops that focus on how the end user perceives the AIS. Various stakeholders from diverse backgrounds can brainstorm or discuss answers collected through a guided interview questionnaire list. This helps identify possible issues and vulnerabilities, such as:

- Is the AIS inclusive?
- What are ethics issues on data?
- What training is required?
- What is the level of awareness and competence on digital skills?
- What is the deviation between intended impact and actual impact?

This can be followed up with foresight on AIS failures: Whose responsibility is it? What is the impact (direct and indirect)? How will you mitigate?

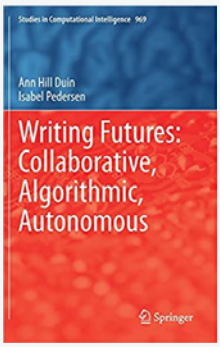
Another method is reframing. We can look at the AIS solution or problem from multiple frames of reference: societal, environmental, health, security, and so forth. This will help create robust policies. Last but not least, we can explore scenarios where the AIS system can fail as well as the consequences for each and roles and responsibilities for mitigation or prevention.

Centering AIS Ethics Around Human Design Thinking

Empathy is a critical part of building AIS that is relevant and beneficial for the humans for which they are designed. The impact of AIS has grown so rapidly and become so pervasive in every aspect of people's lives that when they do not perform as expected, they raise distrust in the technologies. An example of the scale of adoption is the airline industry versus the augmented reality game Pokemon Go. While the airplane took 68 years to reach 50 million customers, Facebook took three years, and Pokemon Go 19 days. As Jeff Felice, Committee member and president of CertNexus says, "The challenge with any technology is that it often moves more rapidly than the laws and regulations that attempt to govern it." To create an AIS that resonates with its users, we need empathy, and one of the best techniques is using human-centered design, which considers the potential effects of the technology on the user physically, mentally, and emotionally.

Citizens will also play a significant role. Human-centered design means encouraging collaboration across disciplines to amplify a diversity of voices, opinions, and practices. One of the difficult challenges lies in comprehension. Jordan Harrod, scientist, defines AI literacy as the ability for a person to confidently understand and interact with AI-based systems.

Duin and Pedersen, in their book *Writing Futures: Collaborative, Algorithmic, Autonomous,* promote the idea of AI literacy for global citizens, students, and professionals in both technical and nontechnological fields. They say that at the heart of "an AI literacy lies the issue of changing roles and infrastructures. Understanding technological agency developed out of human and non-human collaboration involves looking to future landscapes." Finland's Ministry of Economic Affairs and Employment, for example, encourages its citizens to take the Elements of AI course to gain a degree of literacy as a means of civic engagement.



***Writing Futures: Collaborative, Algorithmic, Autonomous* by Ann Hill Duin and Isabel Pedersen**

In addition, the reality of how many citizens have no access to the technology should be considered. For example, even in a city like New York, many public schools lack the resources to provide access to technology, let alone accompanying AIS literacy education. It is vital to highlight the responsibility of ethical human design on the designer, managers, and producers of the AIS products since they are the beginning of the value chain.

As Kohei Kurihara, Committee member and CEO and Co-Founder of Privacy by Design Lab, Japan, notes:

"Engineers and technologists will have much stronger power, not to produce technology but social direction unintentionally. Unless they have an ethical mindset, their actions may sacrifice some parts of our society based on their inherent biases reflected in the outputs of what they create."

Masheika Allgood, Committee member and Founder of AllAI Consulting, continues on the issue of trust and says:

"I think the issue with these efforts is that they don't get to the core issue of distrust with AI systems—control. We don't trust AI systems because no one outside of the developers has any control over how they make decisions. People don't want to just understand why the system made the decision it made. We want the ability to influence the decision before it's finalized to play an active role in decision-making."

Governments are now among the primary investors in AIS research and development. Governments began the digitalization journey in the early 1990s, expanding products in government-to-government (G2G), government-to-business (G2B), and government-to-citizen (G2C) portfolios. Governments, through their inherent function, work in the public value space by default. Governments, therefore, need to be inclusive, giving voice to the minority—those without power—and to the voiceless—the future generations or refugees who are outside the national context. In today's interconnected world, AIS design must be globally inclusive. This means looking beyond national interests and embedding global citizenship and embracing the Universal Declaration of Human Rights.

Abel Pena-Fernandez, Committee member, senior IEEE member, and Co-Founder of Code Explorers World Wide states:

"As a topic for diplomacy, AIS without ethically aligned design is a recipe for diplomatic crises. AIS without ethically aligned design can very quickly become a tool of intellectual domination and modern colonization. Cross-cultural cooperation is essential if AI is to bring about benefits worldwide."

Governments work with diverse stakeholders to solve "wicked problems" that take generations to mitigate. To work in this complex public value space, governments began using the concept of human-centered design thinking. One of the early pioneers of this methodology was MindLab, Denmark, which began in 2002. They were not necessarily looking at AIS but data to understand the nature of human problems to find where proactive solutions could work. For example, the United Arab Emirates (UAE) government introduced the Ministry of Possibilities, a virtual space to handle wicked problems using design thinking. The ministry works on anticipatory service, incentivizing behavior, and improving procurement. While the focus on such initiatives is on experimentation and empathy to achieve scale, such design methods allowed governments to create more adaptive AIS solutions that were relevant for the user yet efficient from a resource point of view. This process is critical as massive digital projects like smart cities, electronic IDs, digitalization of health, and blockchain projects have now become the norm. Though digitalization has accelerated with the COVID-19 pandemic, up to 80% of transformation projects fail.

Furthermore, governments own massive amounts of citizen data and this will become an area of future conflict—as seen with the European Union's GDPR and its 2021 Proposal for a Regulation on a European approach for Artificial Intelligence and with the California Consumer Privacy Act (CCPA) of 2018. The challenges of data paucity, scarcity, and incompatibility must all be considered when using human-centric design principles.[6]

Kohei Kurihara, Committee member and CEO and Co-Founder of Privacy by Design Lab, Japan, says:

"AIS has to consider inclusive design, enlarge the scope of dataset based on human data that won't be excluded from AIS application and service."

This human-centered thinking at the governmental level has multiple layers, including mandates for what projects to fund; which technologies to adopt or commission (e.g., robotics); how to regulate these technologies; how to monitor deployment and impact; how to determine purchasing agreements that ensure accountability and transparency; and how to communicate and educate the population. This ethos applied to AIS is reflected in the General Principles outlined in *Ethically Aligned Design*.

Melodena Stephens, Committee member and Professor of Innovation Management, Mohammed Bin Rashid School of Government, Dubai, co-author of *AI Enabled Business: A Smart Decision Kit*., says:

"The need for human-centered thinking extends broader than the design of products or services, to policy instruments and tools—product specification and approvals, laws and regulations, and grant approvals. These changes may need to be done in tandem with greater awareness of rights (most people are unaware of data rights, for instance), and global citizenship values."

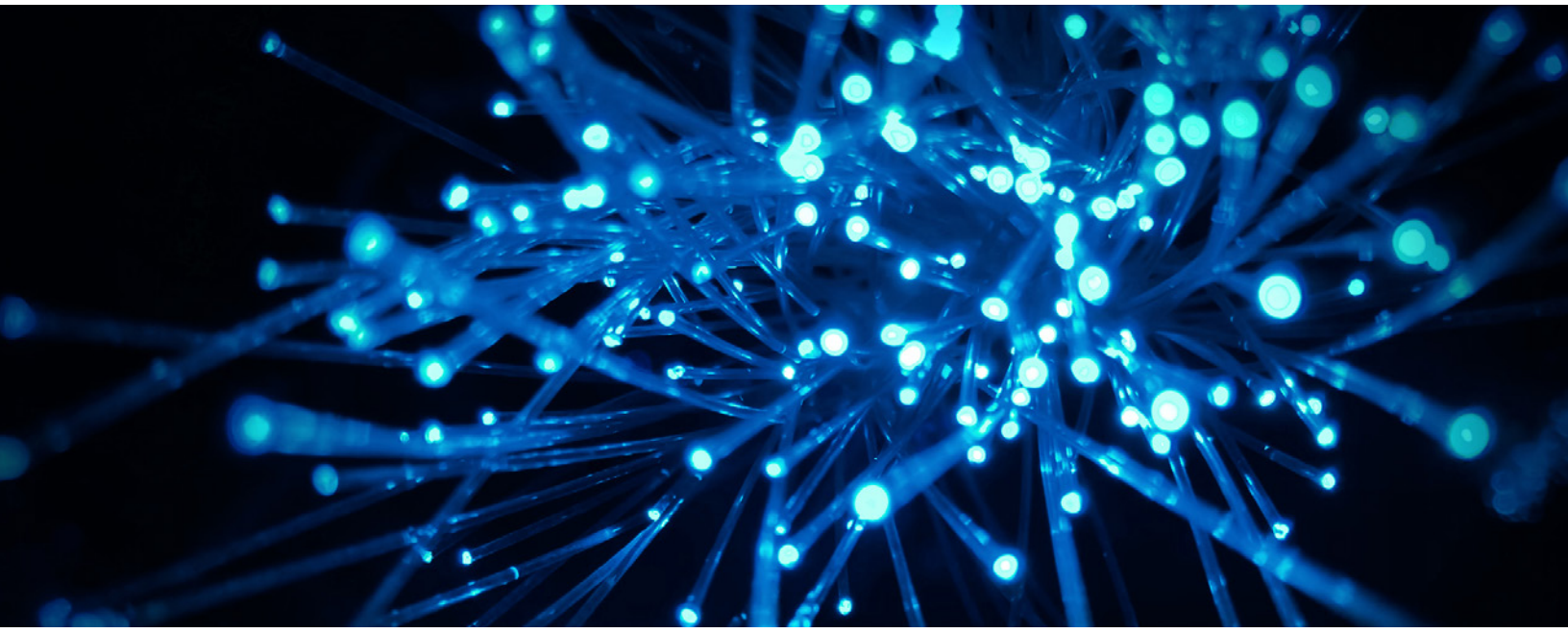
[6] For further information, refer to Appendix 2: AIS and Governments: Supranational and National Level.

Conclusion: The Sanctity of the Systems

It is easy to forget that AIS are driven by humans, and the decisions humans make are based on their personal and professional understanding and the data they use. The mystique around these AI technologies espoused by media, including films, has assigned either a dystopian or utopian importance that belies the fact that these systems are imperfect reflections of ourselves—our human knowledge, behaviors, and identities.

When analyzed in aggregate form, it can also be easy to assign a scientific importance to analyses that may provide information but not necessarily wisdom or actionable insights. The exponential expansion, speed, and volume of AIS design, development, and mobilization will require even more alignment with standards and goals for well-being, which will call for deeper methodological scrutiny of how a technology or device affects our sense of self or our loved ones. In this regard, the AIS and the technologies they comprise are less important than AI ethics educators reminding students, corporations, policy makers, and end users that their values are not only worth studying but should form the basis of any responsible design and innovation. This is to help ensure that AIS technologies reflect values we as humans embrace rather than values that an AIS has assigned based on our words, sentiments, and behaviors.

Simplistic as it may sound, bringing together multiple disciplines to discuss AI ethics means hearing from various areas of expertise about the nature of humanity. Lawyers with a focus on justice will have different values than the users or engineers fascinated with solving design challenges or creating technology from inspiration and mechanical knowledge. In addition, as the effects of AIS reach beyond traditional disciplines consulted, stakeholders and users (intended and unintended) must be considered. In this vein, the myriad ways technologies are perceived is in the "AI of the beholder." For example, a technically savvy adult will perceive a voice assistant differently than a child who thinks the voice belongs to a human in another room. Both of these audiences deserve human rights and agency. To consider all aspects, practitioners in a full range of human sciences and disciplines must also be engaged. End-user values or actual participation in design must be prioritized; otherwise, all unintended consequences and risk can be classified as what they actually really are—the denial to overtly include the culture, opinions, values, or feedback of end users. Working myopically is irresponsible and inherently brings risk for end users not involved during creation; furthermore, the myopic view is simply arrogant and amounts to poor design. The opportunity for innovation and the intention to serve humanity come with a mandate regarding communication—how these tools are perceived cannot be about any designer's, developer's, or producer's intention in isolation.



The challenge and call of AI ethics educators is to move beyond bias and hold to the sanctity of systems—the human systems that interact and communicate knowledge and wisdom with one another. This goes beyond the need to simply provide endless information (resulting in overload or misinterpretation) by recognizing people's need to feel that their voices are being heard and their values are being honored; otherwise, they will not trust those putting inventions in their midst. This practice for educators, then, is one of being modern-day soothsayers or interlocutors between multiple groups of stakeholders, allowing them the opportunity to recognize the core of our humanity—with its fallibility and beauty—and to provide a universality leading to empathy. It is this desire to impart wisdom that mirrors the empathy educators naturally possess, stemming from their need to improve the lives of those they teach.

The intention of this initial work of the committee is to do just that; to improve lives. To begin this ongoing process, we are inviting AIS educators to expand the understanding of ethical considerations in order to build processes of critical thinking. This in turn becomes the foundation to scaffold inclusive AIS ethics education, in which we also invite the participation of non-traditionally trained AIS designers, AIS developers and producers, and the consumers and users (intended and unintended) of AIS. Eventually, we hope to move even further into the education and preparation of younger generations. It is with this desire, and honoring the expertise of the committee's contributing practitioners and resources, we share this work focused on AI ethics education now and in the future.

Appendices

APPENDIX 1: AIS LANDSCAPE

Artificial intelligence systems (AIS) have become ubiquitous.

As the development of AIS accelerates, there is a growing need to ensure these solutions are not causing any harm, and there is hope that they are doing good. UNESCO's 2021 recommendation highlights that AIS can help in education, create an additional 3.3 million jobs, and add US $4 trillion in added value by 2022.

Yet, despite the popularity of the term AIS, there seems to be very little alignment on what the term means (see Exhibit 1). Definitions vary, some focusing on human intelligence parameters and others on technical specifications—hardware and software. The differences in definitions and uses of the term lead us to explore commonalities and examine possible operational definitions that support responsible development and deployment of AIS.

What is common across all the definitions is the ability of the machine systems to mimic human intelligence at some level. The fact that there is no agreement on what human general intelligence specifically is adds to the complexity of this definition and our perception of what AIS is or is not.

Marvin Minsky (2000)[7], Co-Founder of the Massachusetts Institute of Technology's AI laboratory, said:

"Only a small community has concentrated on general intelligence. No one has tried to make a thinking machine. The bottom line is that we really haven't progressed too far toward a truly intelligent machine. We have collections of dumb specialists in small domains; the true majesty of general intelligence still awaits our attack...We have got to get back to the deepest questions of AI and general intelligence."

AI was initially fueled by governments during World War II and the Space Race. However, most government AIS strategies do not start with a definition of what they mean by AIS, nor is there much overlap on what is meant by responsible AIS or governance of AIS principles. Only 42 countries signed the OECD Principles of AI; and 14 governments along with the European Union (EU) joined the Global Partnership on AI Initiative in 2020. This is despite the fact that more than 172 countries use AIS according to Oxford Insights' 2020 Government AI Readiness Index. The number of AI engineers who can inform the public is surprisingly small. The 2019 State of AI report puts this number at about 22,000 highly trained AI specialists and perhaps 300,000 AI researchers and practitioners within broader technical teams.[8] Yet almost every company and government are making decisions on AIS and, without a detailed understanding of what AI can and cannot do, we may have a complicated situation on our hands. Recent challenges that highlight this problem include the 2018 and 2019 Boeing Max 737 plane crashes (pilots were unable to get control over the autonomous decision-making capability of the planes); the 2020 United Kingdom government's decision to use AIS to determine marks for the A-levels; and, in the United States, the 2021 Appriss software scandal on prescription drugs in the healthcare system.

There is a lack of a common understanding of AI among experts and in the common press. Even in Europe there are some concerns that they may be using the wrong definition of AI. For example, because the definition may present loopholes in the legal framework, some intelligent systems are at risk of being excluded from oversight in the European Union's proposed legislation.[9] This is bad for both businesses and citizens. As illustrated in Table 1, definitions vary, creating challenges in how to approach AIS ethics education.

[7] Stork, D. G., Hal's Legacy. 2001's Computer as Dream and Reality, Cambridge, MA: The MIT Press, 2000, p. 27.

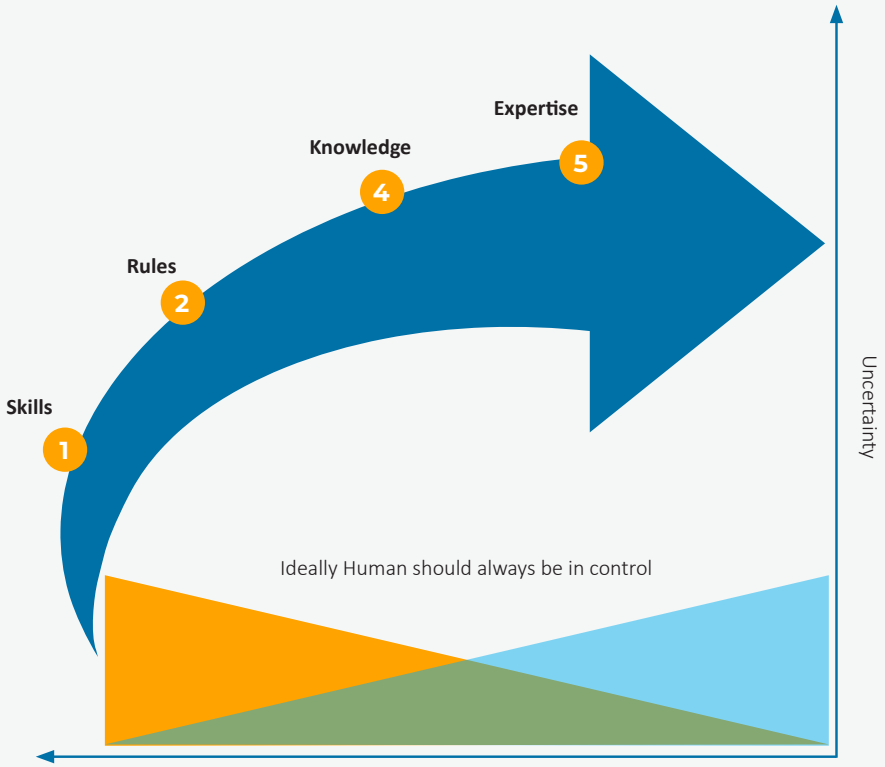
[8] MMC Ventures, The State of AI 2019: Divergence.

[9] Bryson, Joanna J., Europe Is in Danger of Using the Wrong Definition of AI, WIRED, March 2, 2022.

TABLE 1: DEFINITIONS OF AI

Organization	Definition of AI
Ethically Aligned Design (EAD (2021) Glossary. Note the EAD document does not define AI.	"The capacity of computers or other machines to exhibit or simulate intelligent behaviour" (Oxford English Dictionary). The definitions differ depending on whether they are meant for a general audience or those in specific fields such as computational sciences, engineering, economics and social sciences, ethics and philosophy, or international law and policy.
Future of Life Institute (2021). Note Asilomar AI Principles do not define AI.	"Artificial intelligence today is properly known as narrow AI (or weak AI), in that it is designed to perform a narrow task (e.g., only facial recognition or only internet searches or only driving a car). However, the long-term goal of many researchers is to create general AI (AGI or strong AI). While narrow AI may outperform humans at whatever its specific task is, like playing chess or solving equations, AGI would outperform humans at nearly every cognitive task."
WIPO (2021)	"There is no universal definition of artificial intelligence (AI). AI is generally considered to be a discipline of computer science that is aimed at developing machines and systems that can carry out tasks considered to require human intelligence. Machine learning and deep learning are two subsets of AI. In recent years, with the development of new neural networks techniques and hardware, AI is usually perceived as a synonym for 'deep supervised machine learning.'"
European Commission on AI (2019)	AI system refers to any "AI-based component, software and/or hardware," focusing largely on rationality (p. 1).
OECD (2019)	"An AI system is a machine-based system that can, for a given set of human-defined objectives, make predictions, recommendations, or decisions influencing real or virtual environments. AI systems are designed to operate with varying levels of autonomy."
UNESCO (2019) 40th Session of the General Conference, 12–27 November 2019.	Looks at AI as a cognitive technology (p.3). Acknowledges the multitude and diversity of definitions of AI (p. 5). The theoretical and scientific definition adopted is "using AI concepts and models to help answer questions about human beings and other living things" (Boden, 2016, p.2).
UNESCO (2022)	The pragmatic or technological definition of AI is "engineering-oriented," interdisciplinary, and directed "in order to create machines or programs capable of independently performing tasks that would otherwise require human oriented intelligence and agency" (pp. 5-6).
AI Guide (2021), from UAE	"AI defines a collection of technologies enabling a machine or system to comprehend, learn, act, and sense like a human."
DARPA (2020) (Slide 2)	"*Artificial intelligence* is a *programmed* ability to process information." Specifically, DARPA looks at perceiving, learning, abstracting, and reasoning.
Singapore's Model Artificial Intelligence Governance Framework (2019). (p. 18)	"AI refers to a set of technologies that seek to simulate human traits such as knowledge, reasoning, problem solving, perception, learning and planning, and, depending on the AI model, produce an output or decision (such as a prediction, recommendation, and/or classification). AI technologies rely on AI algorithms to generate models. The most appropriate model(s) is/are selected and deployed in a production system."
Britannica	"*AI* is designated as "the ability of a digital computer or computer-controlled robot to perform tasks commonly associated with intelligent beings."
McCarthy (2004)	"It is the science and engineering of making intelligent machines, especially intelligent computer programs. It is related to the similar task of using computers to understand human intelligence, but AI does not have to confine itself to methods that are biologically observable."
IBM (2022)	"Artificial intelligence leverages computers and machines to mimic the problem-solving and decision-making capabilities of the human mind."

AIS are currently being used for various types of enhanced machine and software technology that mimic human intelligence. The key challenge in using, adopting, designing, or promoting these technologies is to understand the level of control the user and beneficiary are delegating to the AIS. This depends on the skills, rules, knowledge, and expertise the system requires and needs to be mapped against the uncertainty of the situation (see Figure 1 in this appendix)[10]. The world, over the last few years, has become increasingly uncertain; yet some of the most automated systems to which we relinquish control are in these situations! In this context, education (both formal and informal) has a vital role to play.



Relative Strengths of AIS versus Human Information Processing

Level of Automation	Automation Description
1	The AIS offers no assistance, human must take all decisions and human action
2	The AIS offers a complete set of decision/action alternatives OR
3	AIS narrows the selection to a few alternatives OR
4	AIS suggests one alternative
5	AIS executes the suggestion that human approves OR
6	Allows the human a restricted time to veto before automatic execution OR
7	Executes automatically, then necessarily informs the human AND
8	Informs the human only if asked OR
9	Informs the human only if, the AIS, decides to.
10	The computer decides everything and acts autonomously, ignoring the human.

Source: Adapted from Cummings (2014) and Parasuraman, Sheridan, and Wickens (2000).[11]

Figure 2: Role allocation for information processing behaviors (skill, rule, knowledge, and expertise) and the relationship to uncertainty

[10] Cummings, M., "Man versus Machine or Man + Machine," *IEEE Intelligent Systems*, 2014, pp. 1541-1672.

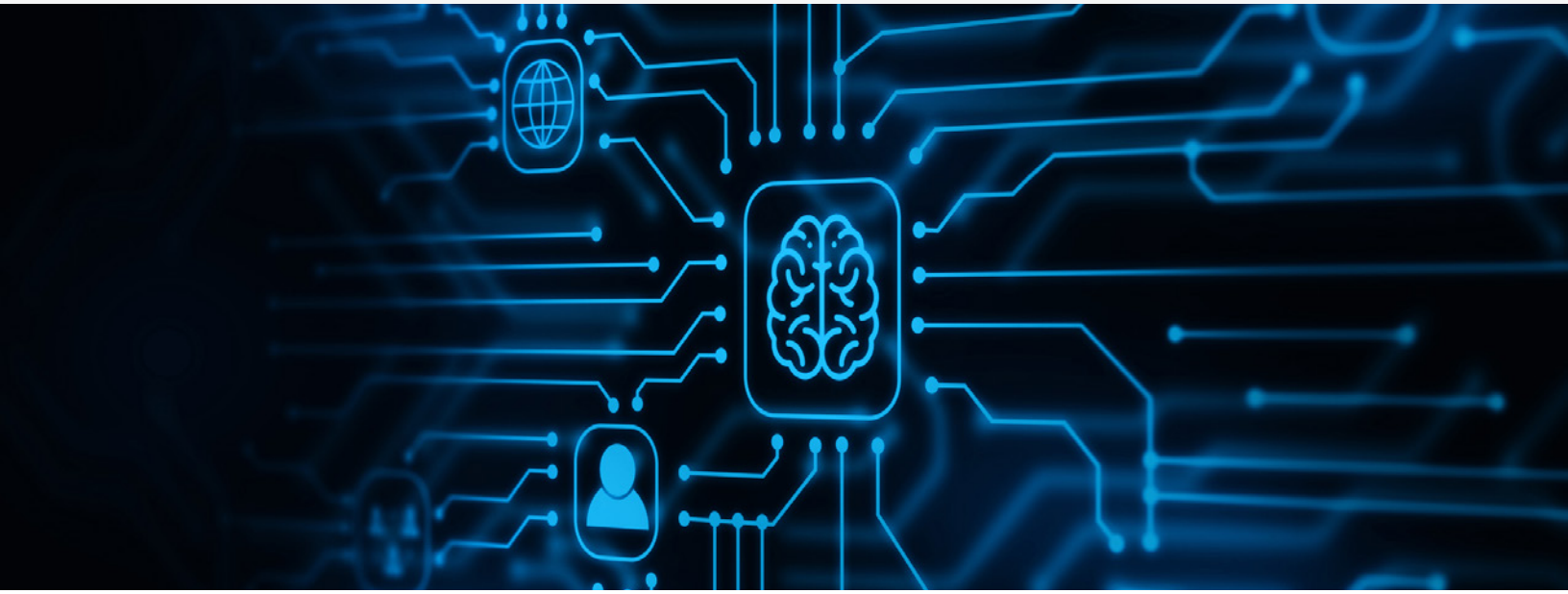
[11] Parasuraman, R., T.B. Sheridan, and C.D. Wickens, "*A Model for Types and Levels of Human Interaction with Automation*," IEEE Transactions on Systems, Man, and Cybernetics—Part A: Systems and Humans, vol. 30, 2000, pp. 286–297.

Autonomous and intelligent technical systems are often designed to reduce the necessity for human intervention in our day-to-day lives. In doing so, these new systems are also raising concerns about their impact on individuals and societies. Current discussions in the EU, OECD, Asilomar,[12] Canada, UNESCO, and in the private sector include advocacy for a positive impact, optimization of processes and resource usage, more informed planning and decisions, and recognition of useful patterns in big data. Discussions also include warnings about potential invasion to privacy, discrimination, loss of skills, adverse economic impacts, risks to the security of critical infrastructure, and possible negative long-term effects on societal well-being. For example, many of the documents talk of the ethics responsibilities within the lifetime of the AIS, but what happens if it is part of a legacy system on which a new AIS is built? (Stephens and Munoz, 2021) What is the impact of AIS adoption on future generations, not just the user of today?

Because of their nature, the full benefit of these technologies will be attained only if they are aligned with society's defined values and ethical principles. We acknowledge that societies are diverse and so are the people who represent them. Through this work we intend, therefore, to establish frameworks to guide and inform dialogue and debate around the nontechnical implications of these technologies, in particular those related to ethical concerns. We understand ethical to go beyond moral constructs and include social fairness, environmental sustainability, and our desire for self-determination.

As John C. Havens says when he mentions his concerns about innovations:

If I bring this up, people often say, "don't hinder innovation" to which I always answer, "I think you mean 'don't mess with my business model.'" I wish people would just say that, because that's the problem—the business model. Or as he ponders the power of new tech's ability to write, "As a long-time writer, when I first realized this was happening, I didn't say, "WOO HOO! No more WRITING! Let's go to the BEACH!" I sat in a long painful silence thinking, "How much harder will it be for my voice to be heard if I'm competing against machines that can put out content non-stop in every language based on people's sentiment" etc. It was not fun."



[12] More information can be found at: https://www.techtarget.com/whatis/definition/Asilomar-AI-Principles.

APPENDIX 2: AIS AND GOVERNMENTS: SUPRANATIONAL AND NATIONAL LEVEL

Through intergovernmental organizations like the United Nations, the Organisation for Economic Co-operation and Development (OECD), and the European Union (EU), countries align regulatory practices and set a vision for governance, for example, Universal Basic Human Rights or Sustainable Development Goals. However, implementing these at the national level through the various stakeholders is not that easy. The government plays an important role in the AIS ecosystems. Canada was one of the first countries to launch an AI strategy in 2017. The United Arab Emirates (UAE) was the first country to appoint a Minister of AI. By January 2020, 27 countries and the EU had published national AI strategies, and 18 more were in the process. A report by CIFAR (2020) (p. 16) finds that "policy areas with the most specific measures across all countries are data and digital infrastructure, talent development, and industrial policy. On the other hand, the policy areas with the fewest specific measures across published strategies are AI in government and inclusion."[13] According to the 2020 AI Government Readiness Index (p. 21), which looks at 36 countries and nine indicators over the four dimensions of inclusivity, accountability, transparency, and privacy (a heatmap is available in the CIFAR 2020 report), the following areas of work, ethics, data, government, and inclusion need more prioritization. The large AI superpowers have much to improve: United States (score 50.008/100), Canada (65.005/100), China (34.475/100), India (41.190/100), United Kingdom (54.566), and Russia (39.124).

Governments impact the trajectory of AIS via research funding, regulations, and policy. For example, DARPA (Defense Advanced Research Projects, a USA National Defense Agency), has been supporting research in AIS for more than 50 years, and its focus has been "AI technology creation to define and to shape the future." DARPA's focus has been to create the third wave where machines will use AIS in such a way that they become colleagues rather than tools, with a much greater level of human-machine symbiosis. Governments are also large purchasers of AI; this is a growing trend since the early digitalization movement of governments in the 1990s.

Jeff Felice, Committee member and President of CertNexus says:

"The challenge with any technology is that it often moves more rapidly than the laws and regulations which attempt to govern it. This issue has been exacerbated by AIS as the technology is not 'controlled' by humans as is it designed today. If we are to pursue these undertakings with the urgency that is necessary, we will a decade from now have technology completing many tasks, research, and other computations that are unimaginable to many today while at the same time benefiting the earth and its inhabitants."

With new initiatives like smart city, big data, electronic medical records, digital identities, and national security, this will just become more complex. Though employees in public service may not all be engineers, they can influence deployment or adoption of AI.

Only 42 countries had signed the OECD Principles on AI by 2019.[14] This fact also highlights the importance of educating civil servants on the importance of AI ethics, especially as the government plays a key role in educating the public. Principle 5 of the Canadian government's "Responsible Use of AI" is to "provide sufficient training so that government employees developing and using AI solutions have the responsible design, function, and implementation skills needed to make AI-based public services better."

[13] CIFAR, Building an AI World: Report on National and Regional AI Strategies Second Edition, by Johnny Kung, PHD.

[14] OEDC.org, Forty-two countries adopt new OECD Principles on Artificial Intelligence, May 22, 2019.

APPENDIX 3: FLOWCHART FROM IEEE STD 7010™-2020

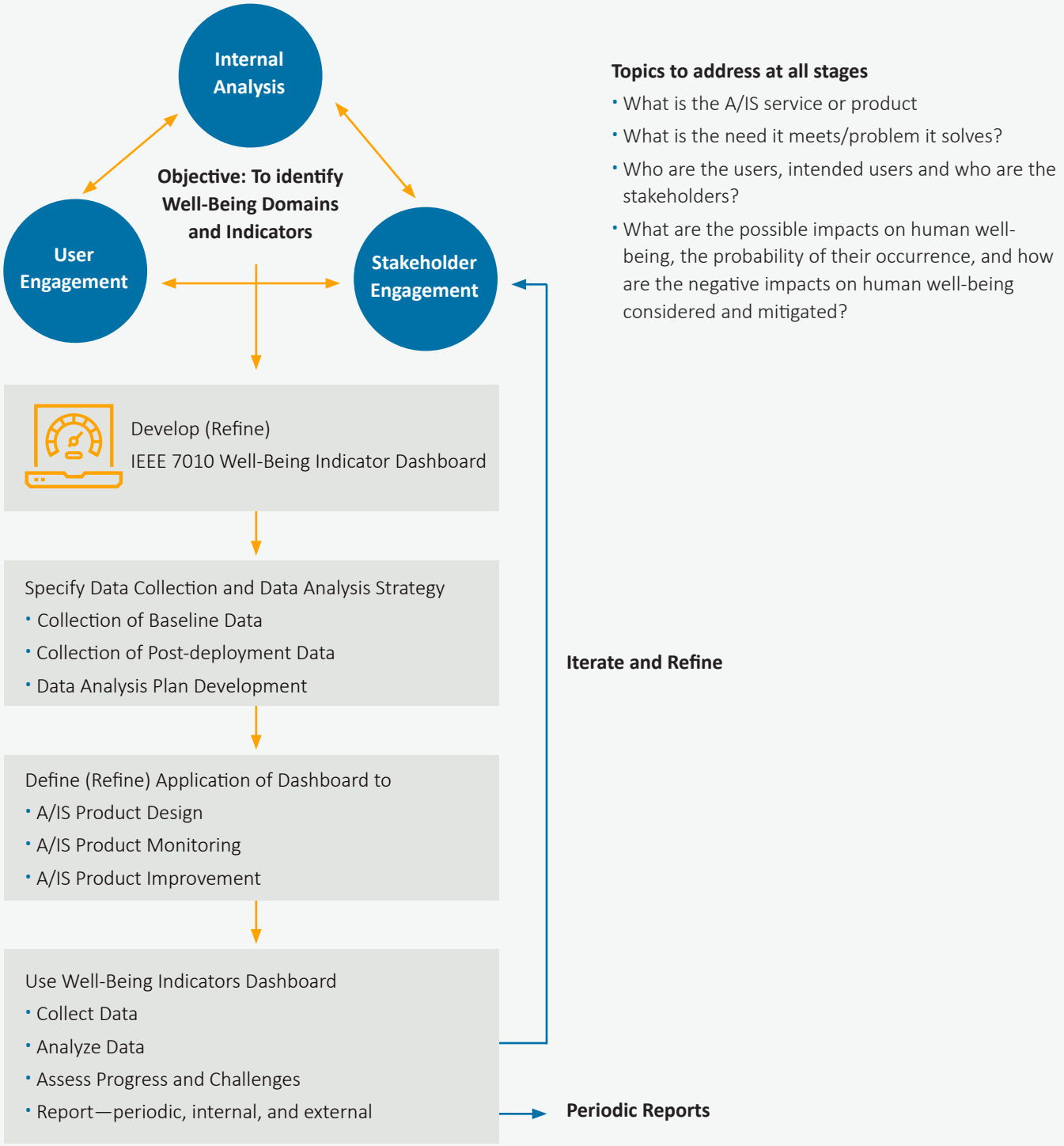


Figure 3 Flowchart of the iterative and adaptive nature of IEEE 7010 Wellbeing Impact Assessment (WIA)

ADDITIONAL RESOURCES

[1] Anderson, John L., "Why Social Science? Because Engineering Is Intended to Benefit Society," PennState Social Science Research Institute, https://ssri.psu.edu/news/why-social-science-because-engineering-intended-benefit-society.

[2] Crawford, K., "Atlas of AI: Power, Politics, and the Planetary Costs of Artificial Intelligence," 2021.

[3] Duin, A.H., and I. Pedersen, Writing futures: Collaborative, Algorithmic, Autonomous, Cham, Switzerland: Springer Nature Switzerland, 2021. https://www.springer.com/gp/book/9783030709273

[4] Dubber, M., Pasquale, F. and Das, S. (eds), "Oxford Book of Ethics of AI," Oxford University Press, 2020.

[5] Franks, B. "97 Things About Ethics Everyone in Data Science Should Know: Collective Wisdom from the Experts," O'Reilly Media, 2020.

[6] Havens, J.C., "Heartificial Intelligence: Embracing Our Humanity to Maximize Machines," Tarcher Penguin, 2016.

[7] Jilani, Zaid, "The Social Responsibilities of Scientists and Engineers," American Association for the Advancement of Science, Nov. 17, 2021. https://www.aaas.org/news/social-responsibilities-scientists-and-engineers

[8] Jones, Christopher R., and Allan R. Jones, "Two blind mice: It is time for greater collaboration between engineers and social scientists around the RDD & D of industrial technologies," Journal of Carbon Research, Jun. 21, 2016, https://www.proquest.com/openview/e5f36b88e0f8c6591fd614fc270e12a1/1?pq-origsite=gscholar&cbl=2055438

[9] Kerry, Cameron F. Meltzer, Joshua P., Renda, Andrea, Brookings Institute Report, AI cooperation on the ground: AI research and development on a global scale. https://www.brookings.edu/research/ai-cooperation-on-the-ground-ai-research-and-development-on-a-global-scale/

[10] Klein, Lisl, "On the Collaboration Between Social Scientists and Engineers," in The Social Engagement of Social Science, a Tavistock Anthology, vol. 2, Philadelphia, PA: University of Pennsylvania Press, 1993. https://www.degruyter.com/document/doi/10.9783/9781512819052-023/html

[11] Lennon, Michael & Zalabak, Marisa & Dajani, Lubna. (2020). Activating Collective Intelligence to Engineer Transdisciplinary Impacts. 10.1109/SMC42975.2020.9283085. https://www.researchgate.net/publication/344751904_Activating_Collective_Intelligence_to_Engineer_Transdisciplinary_Impacts

[12] O'Neil, C., "Weapons of Math Destruction: How Big Data Increases Inequality and Threatens Democracy," Crown, 2016.

[13] Rathenau Instituut, Human rights in the robot age: Challenges arising from the use of robotics, artificial intelligence, and virtual and augmented reality. https://www.rathenau.nl/sites/default/files/2018-02/Human%20Rights%20in%20the%20Robot%20Age-Rathenau%20Instituut-2017.pdf

[14] Roberts, S. T., Behind the Screen: Content Moderation in the Shadows of Social Media, Yale University Press, 2019.

[15] Stephens, M., H. Vashishtha, and Wagner, D. AI Enabled Business: A Smart Decision Kit, Charlotte, NC: Information Age Publishing, 2022/23 (forthcoming).

[16] AI for the Planet: Highlighting AI innovations to accelerate impact. https://www.unesco.org/en/articles/ai-planet-highlighting-ai-innovations-accelerate-impact ?TSPD_101_ R0=080713870fab2000d2eb03f92f5ab469c287e7df5e270a69977f95ca231c86871f0005eec0810246087 e9070e5143000e548c2c0a2a9d095c564f0465483ed27e10c8ff7c817ced185ad092ef8185c4a008ec89a00abcefb8411f7f 5c91b40e6.

IEEE SA STANDARDS ASSOCIATION

Visit us at **standards.ieee.org** and **connect** with us on:

Facebook: facebook.com/ieeesa

Twitter: @ieeesa

Linkedin: linkedin.com/company/ieee-sa-ieee-standards-association

YouTube: youtube.com/ieeesa

Instagram: instagram.com/ieeesa

Beyond Standards blog: beyondstandards.ieee.org

IEEE