

5 December 2019

Role of Standards in Facilitating Innovation while Addressing Ethics and Value in Autonomous and Intelligent Systems (A/IS)

Executive Summary

This think piece describes the roles and benefits of Information Communication Technologies (ICT) Standards and general standardization process, as well as provide benefits of ICT standards. The think piece then moves to provide an overview of Autonomous and Intelligent Systems (A/IS), describe the ethical dimensions and concerns, and express this conviction of OCEANIS members that robust standards are needed to facilitate innovation in these spaces, now and in the future. To encourage the development of socially beneficial applications of A/IS, and to protect the public from adverse consequences of A/IS, intended or otherwise, effective standards and government regulations are needed. This think piece is intended for standards developing organizations (SDOs), governments and organizations contributing to the engineering of dependability and trustworthiness of A/IS as they develop and standardize technical solutions. OCEANIS members will publish the final paper via their formal and informal channels to reach the relevant stakeholders.

Voluntary or open standards in the area of ICT are norms and specifications, in most cases with a world wide application potential. They are essentially designed to enable innovation and open new market opportunities to their users by allowing seamless integration of products, services and processes. Further, they help create ecosystems that promote high functionality, economies of scale and healthy competition. This is essential to help ensure that markets remain open, allowing consumers to have choice and allowing new entrants to successfully enter markets while minimizing risk associated with developing new technologies. Because technology is an enabler for global interconnection and interdependence, mechanisms for unifying technology are foundational to increased economic growth and trade.

A/IS are systems that are capable of adaptation and learning based on feedback and data from their environment. A/IS hold great promise to benefit society in areas as diverse as transportation, health and social care, enterprise productivity, communication network optimization, power grid adaptation and management, agriculture, manufacturing and entertainment, to name but a few application domains. A/IS are key to enabling digital transformation and are already being applied to boost efficiency, solve problems and create scalable individualized experiences. However, digital transformation can only be implemented successfully where trust is achieved through transparency and the process is driven by ethical principles. As the use and impact of A/IS become pervasive and affect many aspects of our lives, the need to establish societal and policy guidelines in order for such systems to remain human-centric, serving humanity's values and ethical principles has emerged.

Effective A/IS policies serve the public interest in several important respects. A/IS policies and regulations, at both the national level and as developed by professional organizations and governing institutions, would protect and promote safety, privacy, human rights, and cybersecurity, as well as enhance the public's understanding of the potential impacts of A/IS on society. Without policies designed with these considerations in mind, there may be critical technology failures, loss of life, and high-profile social controversies. Such events could engender policies that unnecessarily hinder innovation, or regulations that do not effectively advance public interest nor protect human rights.

5 December 2019

Standards are critical components of the national and global technology infrastructure, and are vital to industry and commerce. Many are also crucial to the safety of citizens, and as A/IS technologies become more prevalent, nowhere is the government's role more important for protecting its citizens than in this emerging technology area. In particular, such technologies require a collective response to their rapid implementation in order to preempt possible negative effects of their use. Formal standards bodies, consortia, alliances, and other related organizations and stakeholders must work together to tackle emerging standardization needs in A/IS.

Please see a list of all OCEANIS SC members [here](#).

Special thanks to Adapt Centre, BSI Group, CIO Strategy Council, Fujitsu, KfV, IEC and IEEE for their drafting work.

Introduction

Many discussions about the dependability and trustworthiness of Autonomous and Intelligent Systems (A/IS) are already occurring and are intensifying. Although it is inherently difficult to define "intelligence" and "autonomy", for practical purposes, one could say that computational systems using algorithms and data to address complex problems and situations, including the capability of improving their performance based on evaluating previous decisions, could be perceived as "intelligent". Such systems could also be regarded as "autonomous" in any given domain as long as they are capable of accomplishing their given tasks despite environmental changes within (and sometimes outside) this domain.

In this specific domain of A/IS (and beyond sci-fi sensationalism and speculations around super intelligence), most discussions revolve around a variety of issues, such as explicit ethical values and dilemmas, transparency, accountability, unintended bias, and data governance. Many governments around the world, businesses, non-profit organizations and other groups are currently working on their lists of issues and priorities. All the above are contextual -- meaning they are not purely technical matters and are often subsumed as "ethical" issues, in a broader sense of the term "ethical". A/IS are the key to enabling digital transformation and are already being applied to boost efficiency, solve problems and create scalable individualized experiences. However, digital transformation can be most successfully implemented where trust is achieved through transparency and the process is driven by ethical principles.

The participants in the Open Community for Ethics in Autonomous and Intelligent Systems (OCEANIS) recognize that understanding the opportunities and challenges associated with such contextual/ethical aspects around A/IS are becoming essential to developing and standardizing A/IS technology that people can trust and further innovate. We also understand that every stakeholder must assume their responsibility in these systems in order to create viable solutions, and that standards developing organizations (SDOs) are an important factor in this equation. Therefore, this think piece provides a broad ecosystem perspective on the role of standardization in facilitating innovation in A/IS, while addressing issues that expand beyond technical solutions to also include addressing ethics and broadly accepted values. It will explore how standards can play a transformational role to support – in its broader sense – the ethical application of A/IS, developing the sector while protecting the end user. This think piece is intended for SDOs, governments and organizations contributing to the engineering of dependability and trustworthiness of A/IS as they develop and standardize technical solutions. OCEANIS members will publish the final paper via their formal and informal channels to reach the relevant stakeholders.

5 December 2019

Benefits and Impacts of ICT Standards¹

ICT standards provide people and organizations with a strong basis for mutual understanding and are critical for promoting, incubating and advancing innovative global technologies by:

- Facilitating the creation and expansion of international markets and minimize market fragmentation;
- Advancing global interconnectivity interdependence and interoperability;
- Affording the opportunity to keep pace in the fast-changing world of technology;
- Decreasing the cost of and enabling the widest choice of goods and services;
- Helping protect the health of people and the environment.

Autonomous and Intelligent Systems (A/IS) Overview

A/IS hold great promise to benefit society in areas as diverse as transportation, health and social care, enterprise productivity, communication network optimization, power grid adaptation and management, agriculture, manufacturing and entertainment, to name but a few application domains. Recent success in machine learning, signal processing, planning algorithms, digital sensing, embedded and cloud computing, and voice, image and pattern analysis have greatly accelerated the application of A/IS. They hold great promise to benefit society, but they also present potential new social, legal, and ethical challenges, with corresponding new requirements to address issues of systemic risk, diminishing trust, privacy challenges, and issues of data transparency, ownership and agency. There is therefore a need for developers and operators of A/IS to maintain awareness of and employ consensus-based global best technical practices and standards. These practices and standards must respect legislation, and recognize and align end-users' and citizen's values when building and deploying A/IS.

As the use and impact of A/IS become pervasive and affect many aspects of our lives, the need to establish societal and policy guidelines in order for such systems to remain human-centric, serving humanity's values and ethical principles has emerged. What has also emerged is the development of a host of frameworks and principles by a great number of institutions and organizations with the ambition to guide development and use of such systems in these directions.

A/IS are specifically designed to reduce the necessity for human intervention in our day-to-day lives. In so doing, these new systems are also raising concerns about their impact on individuals and societies. Current discussions include advocacy for a positive impact, such as optimization of processes and resource usage, more informed planning and decisions, and recognition of useful patterns in big data. Discussions also include concerns about potential harm to privacy, discrimination, loss of skills, adverse economic impacts, risks to security of critical infrastructure, and possible negative long-term effects on societal well-being. While noting that organizations

¹ Please see Addendum One.

5 December 2019

have varying views,² one example of ethical concerns comes from the World Economic Forum asking:³

- *How do we distribute the wealth created by machines?*
- *How will A/IS affect our human behaviors and interactions?*
- *How should we overcome AI bias, including racial sensitivity?*
- *How will we guard against mistakes by A/IS?*
- *How should we guard against unintended consequences?*
- *How do we define the humane treatment of A/IS?*

Ethical Dimensions of A/IS

OCEANIS members agree that A/IS must be created and operated in such a way as to respect, promote, and protect internationally recognized Universal Declaration of Human Rights⁴. Ethical limits should be based on an understanding of end-user values and implementation of their accountable and pre-determined design. Society has not yet established universal standards or guiding principles for embedding human values and norms into A/IS. As these systems are instilled with increasing autonomy in making decisions and manipulating their environment, it is essential that they are designed to adapt, learn, and follow the norms and values of the community they serve. Moreover, their actions should be transparent in signaling their norm compliance and, if needed, they must be able to explain their actions for example, for auditing purposes. This is essential if humans are to develop appropriate levels of trust in A/IS in the specific contexts and roles in which they function.

Effective A/IS policies serve the public interest in several important respects. A/IS policies and regulations, at both the national level and as developed by professional organizations and governing institutions, would protect and promote safety, privacy, human rights, and cybersecurity, as well as enhance the public's understanding of the potential impacts of A/IS on society. Without policies designed with these ethical considerations in mind, there may be critical technology failures, loss of life, and high-profile social controversies. Such events could engender policies that unnecessarily hinder innovation, or regulations that do not effectively advance public interest nor protect human rights.

² IEC - <https://ethicsstandards.org/new-iec-white-paper-on-artificial-intelligence/>

IEEE - <https://ethicsinaction.ieee.org/>

ISO - <https://www.iso.org/news/ref2336.html>

ITU - <https://news.itu.int/ai-in-2019-where-are-we-and-whats-next/>

European Commission High-Level Expert Group on AI presented Ethics Guidelines for Trustworthy Artificial Intelligence -

<https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai> OECD -

<https://legalinstruments.oecd.org/en/instruments/OECD-LEGAL-0449>

UNESCO -

<https://unesdoc.unesco.org/ark:/48223/pf0000367422?posInSet=2&queryId=325cbca9-7ad3-4265-8118-88c3dc451766>

³ <https://www.weforum.org/agenda/2016/10/top-10-ethical-issues-in-artificial-intelligence/>

⁴ <https://www.un.org/en/universal-declaration-human-rights/>

See as an example, Ethically Aligned Design, Edition 1: <https://ethicsinaction.ieee.org> and OECD AI Principles:

<https://www.oecd.org/going-digital/ai/principles/>

European Commission High-Level Expert Group on AI presented Ethics Guidelines for Trustworthy Artificial Intelligence:

<https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>

5 December 2019

A/IS, Ethics, Digital Transformation and Standards

Artificial Intelligence Systems (AIS) are the key to enabling digital transformation and are already changing many aspects of daily life. Related technologies are being applied to boost efficiency, solve problems and create scalable individualized experiences. Finding answers to the many ethical dilemmas raised — that take into account issues such as privacy, security and integrity for the widest possible benefit — is vital to the development of innovative A/IS technologies.

Digital transformation is not only about reimagining business in the digital age to deliver greater value to customers. It is essential that ethical considerations shape the design process in order to maximize public good while limiting the risk of inadvertent harms or unintended consequences. International standards developed by multiple stakeholders should ensure the right balance is struck between the desire to deploy A/IS rapidly and the need to study their ethical implications.

Above all, digital transformation can be most successfully implemented where trust is achieved through transparency and the process is driven by ethical principles. AI relies on data sets, including personal information. How this data is collected, managed and used is also an ethical issue. All stakeholders must have a clear understanding of what organizations hope to achieve and how they will use the data. Full permission must be granted to use personal information with adequate understanding of the likely consequences.

A key issue is bias and fairness of Algorithmic Decision Making Systems (ADMS). While it may be relatively easy to detect and mitigate bias, it is often difficult to get to the bottom of how ADMS are making decisions in order to solve the problems, as more often than not algorithms operate within a ‘black box’. It is one of the most important challenges we face, as algorithms are increasingly at the centre of our daily lives, from search engines and online shopping to facial recognition systems and booking flights. There are ethical concerns, for example, about the use of data collected by facial recognition applications, including development bias.

ADMS are already making decisions that can have a major impact on human lives and safety in areas ranging from banking, employment services to the management of critical infrastructure. In the medical care sector, for example, ADMS technologies are playing many different roles in the delivery of healthcare treatments, surgeries and services, from robotically-assisted surgery, virtual nursing assistants, dosage error reduction and connected devices to image analysis and clinical trials. Such technologies are improving diagnostics and helping doctors to make better decisions for patients. In manufacturing, they are improving efficiency and cutting costs across the whole value chain and life cycle of a product, from idea to order, construction and development, delivery, recycling and all related services. The list goes on. Organizations can now look to standards such as CIO Strategy Council's CAN/CIOSC 101:2019,⁵ a standard helping organizations design and implement responsible ADMS solutions. This standard sets an international benchmark for the ethical design and use of such solutions.

In the connected world in which we live, international standards that are open and inclusive address the entire A/IS ecosystem, rather than focusing on individual aspects, can play a key role in addressing the ethical aspects of the technologies we encounter in daily life. Enhanced collaboration with multiple stakeholders is essential to

⁵ See CIO Strategy Council standard CAN/CIOSC 101 ‘Ethical design and use of automated decision systems’:
<https://ciostrategyCouncil.com/standards/technical-committees/>

5 December 2019

avoid competing approaches that lessen trust and discourage cooperation. International standards reached by consensus are our best hope for ensuring that the process of digital transformation is underpinned by ethical principles, including trustworthiness and fairness.

Moreover, wider adoption of such technologies and systems will depend to a large extent on effective risk management and OCEANIS members are already active in this area. For instance, the joint technical committee set up by IEC and ISO (JTC1/SC42) is carrying out important standardization work.⁶ A new ISO/IEC standard will provide guidelines on managing risk faced by organizations during the development and application of AI techniques and systems. It will assist organizations in integrating risk management for AI into significant activities and functions, as well as describe processes for the effective implementation and integration of AI risk management.

Disruptive technologies like artificial intelligence pose both challenges and opportunities across all sectors. For this reason, the joint ISO and IEC technical committee is liaising with a number of committees, in both organizations, that focus on different technologies and industries, as well as external organizations and consortia.

Another example is the IEEE P70xx standardization projects series that directly addresses most of the contextual issues and aspects of A/IS systems. The flagship project P7000 itself is establishing a model process by which engineers and technologists can address ethical considerations throughout the various stages of system initiation, analysis and design. Three years into its development it is entering a mature phase and the process requirements include management and engineering view of new IT product development, computer ethics and IT system design, value-sensitive design, and stakeholder involvement in ethical IT system design.

Concluding, A/IS systems must be developed and should operate in a way that is beneficial to people and the environment, beyond simply reaching functional goals and addressing technical problems. This approach will foster the heightened level of trust between people and technology that is needed for its fruitful and natural use in our daily lives.

Some of the areas where standards can help maximize A/IS for the benefit of all humanity include:

- Addressing ethical concerns during system design,
- Addressing the expectations by the public for system transparency and accountability,
- Identifying and mitigating algorithmic bias,
- Documenting and protecting users interests in personal data,
- Evaluating reliability of online messaging,
- Protecting privacy and especially children's personal information online,
- Assuring the safety of A/IS,
- Addressing requirements and methods for "human augmentation".

⁶ <https://www.iso.org/committee/6794475.html>

5 December 2019

*Governmental and Intergovernmental Activities*⁷

Governments throughout the world, in both developed and developing countries, are introducing - or have introduced - plans and strategies to address A/IS and their implications on their home country, region and the world. These considerations include a wide array of topics and concerns including:

- Investing in pertinent research and development,
- Setting governance standards,
- Building pertinent workforce,
- International engagement,
- Ethical norms,
- Safety and security,
- Economic, ethical, policy, and legal implications.

The International Telecommunication Union (ITU) is the United Nations' (UN) specialized agency for information and communication technologies, and has become one of the key UN platforms for exploring the impact of AI.⁸ ITU has stated that it "will provide a neutral platform for government, industry and academia to build a common understanding of the capabilities of emerging AI technologies and consequent needs for technical standardization and policy guidance."⁹

"ITU organized multiple events entitled, 'AI for Good Global Summit' in 2017, 2018, and 2019. These Summits have focused on strategies to ensure trusted, safe and inclusive development of AI technologies and equitable access to their benefits. Teams at the Summits highlighted the ability of AI to help achieve the Sustainable Development Goals (SDGs)¹⁰ through activities such as mapping poverty, aiding with natural disasters using satellite imagery, and helping achieve Universal Health Coverage."¹¹

Perspectives

To encourage the development of socially beneficial applications of A/IS, and to protect the public from their potential adverse consequences, intended or otherwise, effective standards and government regulations are needed. Effective A/IS standards serve the public interest in several important respects. A/IS standards and regulations, at both the national level and as developed by standards organizations and governing institutions, protect and promote safety, privacy, human rights, and cybersecurity, as well as enhance the public's

understanding of the potential impacts of A/IS on society. Standards policies designed without these considerations in mind may lead to critical technology failures, loss of life and high-profile social controversies. Such events could engender governmental policies, regulations and laws that unnecessarily hinder innovation, and do not effectively advance public interest and protect human rights.

⁷ See in particular OECD AI Guidelines for suggestions in this area: <https://www.oecd.org/going-digital/ai/principles/>

⁸ <https://futureoflife.org/ai-policy-united-nations/>

⁹ <https://www.itu.int/en/ITU-T/AI/Pages/default.aspx>

¹⁰ <https://www.un.org/sustainabledevelopment/sustainable-development-goals/>

¹¹ <https://futureoflife.org/ai-policy-united-nations/>

<https://ec.europa.eu/digital-single-market/en/news/ethics-guidelines-trustworthy-ai>

5 December 2019

The many forms of standardization serve different distinct purposes, especially in ICT, including A/IS. There is a need for stability (provided by the arena of formal and open standards bodies), coping with rapid change (provided by specialized consortia and alliances), specific intellectual property and marketing environments, and the need for robust community involvement (provided also by tools such as Open Source software development). To tackle the vast emerging standardization needs for A/IS, the groups within each fora will need to more effectively work together in order to achieve the public good characteristics of standards through open systems and open, inclusive standardization processes.

Integrated global markets as well as new requirements related to cultural values such as privacy and ethics, are bringing a new era of standardization to the forefront. This new era needs to be sensitive to local context and, at the same time, be global and inclusive. Standardization processes must therefore be sufficiently nimble to effectively address the development and commercial application of rapidly evolving and potentially intrusive technologies such as A/IS, and they must be capable of addressing ethically aligned design concepts from the onset. In the future, the development and use of certification programs could provide greater transparency and accountability in commercial applications utilizing A/IS.

In particular, the development of A/IS standards needs to be inclusive of diverse communities of experts and users, including economists, ethicists, legal professionals, philosophers, educators, policy makers, regulators, civil society, and community representatives, in addition to technologists and scientists. As many of these new participants have not traditionally directly engaged in technical standards development, there will be a need for significant investments by the SDOs in training and mentoring programs to hasten understanding of and engagement in the process of developing standards. Standards are critical components of the national and global technology infrastructure, and are vital to industry and commerce. Many are also crucial to the safety of citizens, and as A/IS technologies become more prevalent, nowhere is the government's role more important for protecting its citizens than in this emerging technology area. In particular, such technologies require a collective response to their rapid implementation in order to preempt possible negative effects of their use.

Conclusion

Open and inclusive standards development facilitates innovation in A/IS, addresses ethics and values, as well as explores technical solutions to multifaceted issues. Formal standards bodies, consortia, alliances, and other related organizations and stakeholders must work together to tackle emerging standardization needs in A/IS. To help meet this need for improved coordination and collaboration, OCEANIS was formed in 2018 as a Global Forum for discussion, debate, and collaboration for organizations interested in the development and use of standards to further the development of A/IS. The members of OCEANIS are tackling such challenges as how to navigate the process of voluntary standardization in the domain of A/IS.

The standardization ecosystem is complex, with hundreds of pertinent SDOs, SSOs (Standards Setting Organizations) and consortia globally engaged in A/IS. Lacking efficient coordination mechanisms, the gaps that need to be filled are not always evident and the resulting "standards" (industry specifications, guides, open standards, regulations, etc.) are not always mutually consistent or free from contradiction. With this, OCEANIS has developed¹² an integrated, centralized, transparent notification system for new standards work, works in

¹² <https://ethicsstandards.org/repository/>

5 December 2019

progress, and standards published by OCEANIS members and others in the A/IS space to support cooperation, collaboration, and global alignment. Integrating standards development work with the goal of achieving the greatest practicable degree of commonality between standards will help reduce duplication in the preparation of standards, assist in advancing trade, and will help make the most effective use of all resources available.

Beyond OCEANIS, at its annual meeting in March 2019, the Global Standards Collaboration (GSC), an unincorporated voluntary organization dedicated to enhancing global cooperation and collaboration regarding communications standards and the related standards development environment, recognized the importance of addressing A/IS and agreed to expand its scope to include addressing the ethical and societal implications of ICT systems, services, and technologies.

As A/IS join other ICT technologies that have become ubiquitous in our lives, an essential aspect in their development is the need to address ethical dimensions and trustworthiness issues, including robustness, resiliency, reliability, accuracy, safety, security, and privacy from the onset with standards and best practices.

We look forward to continued discussion and collaboration. We welcome ideas, proposals and other input from the broad community involved in developing, standardizing, and using A/IS technologies.

*** **

5 December 2019

Addendum One

Information Communication Technologies (ICT) Standards and Standardization Processes¹³

Voluntary or open standards in the area of ICT are norms and specifications, in most cases with a world wide application potential. They are set by standards development organizations (SDOs) and consortia (this paper uses simply “SDOs” to mean both categories of organizations) based on the consensus of the participants in the process who intend to utilize these standards, such as public authorities, nonprofit organizations, researchers, practitioners, policymakers, educators, representatives from industry, including SMEs, etc. and that are driven by market and societal expectations (interoperability, safety, security, reliability, accountability, etc.). These standards embody voluntary cooperation for the development of technical and other normative deliverables that memorialize the agreements of parties regarding specific technology, systems, practice, etc. with the goal of enabling the development of ecosystems, for instance driven by the desire for multi-vendor interoperability. They are essentially designed to enable innovation and open new market opportunities to their users by allowing seamless integration of products, services and processes. Further, they help create ecosystems that promote high functionality, economies of scale and healthy competition. This is essential to help ensure that markets remain open, allowing consumers to have choice and allowing new entrants to successfully enter markets while minimizing risk associated with developing new technologies.

In particular, in the ICT area standardization processes that are universally open and transparent from the beginning can help establish trust in world-wide platform adoption and use, which can strengthen the “network effect” of digital platforms. The voluntary standards development paradigm also serves as a best practice model - grounded in a bottom up, decentralized, informal and market-led approach - to provide open access and respond quickly to the accelerating pace of technology changes and shifts in markets.

Any voluntary standards landscape that enable digital platforms must be expansive and complex, as there is a plethora of standards specific to the technologies of almost every digital platform, such as social media, cloud computing, mobile, big data and the Internet of Things, as well as for the technical underpinnings of networking and communications technologies and protocols.

Benefits and Impacts of ICT Standards

There are a variety of potential benefits and impacts of ICT standards, and these may affect the wide range of standardization stakeholders differently. While considering the differences in the standard development processes in varying jurisdictions, some SDOs can eliminate the cost of developing standards by governmental agencies, therefore eliminating the associated administrative burden on government. Standardization can also decrease the costs of goods and services procured by governments.

Standards promote, incubate, and advance global technologies through facilitation of standards development and collaboration. Standards also promote innovation, enable the creation and expansion of international markets, and help protect health and public safety. In addition, they drive functionality, capabilities and

¹³ EC Rolling Plan for ICT Standardisation: latest issue 2019 - <https://ec.europa.eu/docsroom/documents/34788>

5 December 2019

interoperability of a wide range of products and services that transform the way people live, work and communicate. Because technology is an enabler for global interconnection and interdependence, mechanisms for unifying technology are foundational to increased economic growth and trade. By enabling global interoperability, standardization also provides building blocks for innovation, helps facilitate economies of scale, and lowers the risk of single vendor lock-in. Further, standardization facilitates the ability to stay current with technology as it advances. When a regulation is needed, by referring to the most recent standard(s), that regulation can be as current as the state-of-the-art standard in that field.

On an international scale, market fragmentation is minimized, and the burden on industry of complying with agency-specific standards that may differ on a regional/country basis is decreased. Efficiency and economic competition is promoted through the global harmonization of standards, and global interoperability is fostered. By participating in global standardization activities, each individual country's stakeholders are provided with opportunities and economic incentives to establish standards that serve national needs, and are afforded the opportunity to keep pace with current technologies developed and deployed world-wide.
