41st International Conference of Data Protection and Privacy Commissioners, Tirana, Albania

Closed Session Topic A: Practical workshop follow-up to the AI and ethics debate at the 40th Conference - From Theory to Practice.

21 October 2019

Opening Speech by Dr. Ing. Konstantinos Karachalios, Managing Director IEEE SA[1]

Ladies and Gentlemen

It is a pleasure to be here with you today. As one key representative of the global techno-scientific communities, IEEE is lending its particular strengths and constituency expertise in a network-of-networks effort to offer collaborative platforms and build understanding of the unfolding AI and data protection domain. We are pleased that through the expertise of the IEEE community—some 400,000 strong—and working in collaborative partnerships across the globe, outcomes of our efforts are influencing the way the technology is developed and used—and more so, we are developing practical approaches, tools and programs to address the challenges. We are engaging also to inform the evolution of regulatory and legal environments, among others, that will ensure that the future development of AI and autonomous and intelligent systems are performed with appropriate care and in alignment with societal values and ethical principles.

The theme of this year's conference is "convergence and connectivity," and I assume that you use these terms mainly in a legislative and regulatory context. Interestingly enough, these terms resonate directly also in the world of technology. The raging speed of technological evolution and convergence in the digital sphere, combined with ubiquitous connectivity gave birth to a "cloud," using networks, search machines, data and algorithms to build a "global system of enormous power and energy,"[2] with a huge impact on the global socio-technical landscape. Data, including our personal digital footprint, are both the fuel and the prize of these machines but at what cost to our privacy and well-being?

Fancy tech gadgets and platforms are integrated in business models that are fed by the fundamental psychological needs of humans for connection, belonging, and recognition, engaging us in mechanisms and processes that emulate a satisfaction of those needs, but structurally do the exact opposite. The paradox is that this virtual connectedness and subtle manipulation disconnect us radically from real time, space and the humans next to us. The more we are connected to these information machines, the less we seem to be open to real knowledge.

This dovetails with the specific focus of this session about how to move from theoretical principles to practice regarding "AI ethics." For the rest of my presentation, I will be using the term algorithmic decision-making systems instead of the marketing acronym "AI." These systems are socio-technical systems, consisting of algorithmic computer programs and data sets as well as of people using them.

---

[1] Speaking here on my own capacity, not necessarily representing official opinions of the organization I work for

[2] James Bridle, New Dark Age: Technology and the End of the Future, Verso, May 2019

They may sometimes be connected to automatic actuators, with a physical impact on their environment.

I guess the reason why your profession increasingly cares about what you regard as ethical aspects of such systems is that they are rapidly becoming the spearhead of the digital technologies that are radically reshaping the data protection and privacy landscapes. For instance, algorithmic decision-making systems could re-personalize a given set of anonymized data. They are also built to nudge us to use gadgets and Internet platforms to the maximum extent, delivering *en passant* all our personal data to second and third parties, for the purpose of a detailed profiling of us all. This peculiar mode of "transparency" opens the gates for – to put it diplomatically – interesting and innovative social engineering, with outcomes that do not thrill all of us the same. So, understandably, many legislators and regulators try to cope with these new forms of socio-technical creativity and to defend the socio-political systems they represent.

While the law makers are trying to become more technically savvy, this appetite to engage is not always reciprocal and technical communities usually do not bother to build bridges with the socio-political actors and to play a role as a neutral advisor; this makes life easy for lobbies with practically unlimited resources to expand their business models based on our involuntary profiling.

So, it is justified to pose the question of a gap between the development of algorithmic decision-making systems and the demands on regulators and policy makers to set some meaningful rules. I'd like to break down this general question into more specific ones and in the context of the principles in the Declaration on Ethics and Data Protections in AI from the 2018 International Conference of Data Protection and Privacy Commissioners:

- Is the technology evolution generally moving faster than the legislative and regulative processes can catch up?
- If yes, is this gap widening in the domain and the era of algorithmic decision-making systems?
- If yes, what could be the societal and political consequences?
- And finally, what should be done and by whom, if we estimate the consequences to be serious?

The gap debate
There are credible analyses showing that there is indeed a rapidly increasing gap between the two processes: the rate of technological acceleration and convergence on the one side, and the delay in time that the political processes take to grasp and react to the socio-political implications of this phenomenon on the other side. In this narrative, the timing gap between the two processes poses a political problem that must somehow be addressed in order to protect our dignity and safety and to mitigate overconcentration of power.

However, there are other voices saying that the gap is a good thing, that legislators and regulators should, as a rule, hold back as to not stifle innovation. The proponents of this theory often use notorious stories of failed attempts to over-regulate technological progress (the famous "man-with-a-flag-walking-in-front-of-the-car" example). They implicitly assume that technological innovation is always something positive and one should not meddle with it.

In the Algorithmic Decision-making Systems era
This classical controversy has been re-energized through the burgeoning "AI ethics" debates, regarding certain "ethical" aspects of algorithmic tools and the underlying datasets.

- What causes this Cambrian explosion of technology ethics?
- Are there specific circumstances related to algorithmic decision-making systems that should guide policy makers to take action?
- Are there plausible scenarios where legislative inertia is a bad option?

I am sure there are other, equally or even more pertinent questions to our theme, but these ones are sufficient for me to deploy a narrative.

The threats
Algorithmic Decision-making Systems are becoming increasingly pervasive and intrusive in our everyday life. They bring many benefits in terms of supporting scientific progress, promoting human wellbeing, creating economic value, and offer the potential to support the exploration of new solutions to major social and environmental problems. They can be used *i.a.* to detect behavioral patterns that may help improve urban planning and help medical scientists understand the genetic context of diseases.

However, as alluded to in my introduction, such powerful and complex technologies also present potential new social, legal, and ethical challenges, with corresponding new requirements to address concerns over explainability, perceptions of systemic risk, and issues of data transparency, privacy, ownership and agency over our identity.

The latter is both a precondition for and the main result of self-determination and freedom. Without this agency we are not able to participate in democratic political processes as *πολίτες,* ergo autonomous political actors. We should not forget that the term *political* itself is directly derived from and intimately linked to the concept of active citizenship within a *πόλις*. So, the guardians of a democratic society cannot afford to remain idle when the very foundations of active citizenship, a pre-condition to any democratic system, are being massively undermined.

Evidently, we are rapidly losing ground. Lord Anthony Giddens says that it is the mastering of the techniques of storage (food, weapons, information) and of their "transportation" that gives rise to empires. Currently, the most precious goods to be stored and transported are information and knowledge. Moreover, the totality of our individual and collective digital footprint has become part of the "cloud" inventory.

It is therefore not surprising that the emperors of the cloud declare privacy as an anomaly in human history and let us know that what would be left of privacy is what they – in their magnanimity - are willing to concede to us. Not very much, I'd guess. It is misleading to reduce this to a problem of an alleged free choice between "privacy" and "convenience," because it is not about our individual taste or exposure. As Paul Nemitz argues in his essay "Democracy and Technology in the Age of Artificial Intelligence,"[3] it is about the capacity of our societal collectives to acquire political autonomy and to establish and nurture democratic spaces and processes.

---

[3] Nemitz, Paul, Democracy and Technology in the Age of Artificial Intelligence (August 18, 2018). DOI 10.1098/RSTA.2018.0089 - Royal Society Philosophical Transactions A. Available at SSRN: https://ssrn.com/abstract=3234336

And while some say we deserve what we are getting, our children are factually deprived of their most fundamental and encoded rights, by having all kind of data about them gathered ever since they were born, and by being treated as consenting adults in the cyber-space. How are they supposed to have a chance in their fight for dignity and political self-determination in a world where their desires and intentions will be transparent from the moment of their birth, in a hyper realization of Foucault's Panopticum?

Action
It is evident that in some areas action should had happened yesterday; there is no reason to wait anymore, the damage is taking place in front of our eyes. Beyond undermining our democracies, we are exposing our children to very significant mental risks.

The segments of the three communities of policy makers, enterprises and technologists that believe this is an untenable situation have to join their forces and work together to establish clear laws/rules and implementable technical standards, giving rise to platforms that are compatible with socio-politically sustainable business models.

To start with, we, techno-scientists can reduce as much as we can feeding the forces creating the vicious cycle. To understand these forces, we must add a layer of self-reflection, individually and collectively, about what we are doing. We should not bring ourselves to the very unpleasant position to ask ourselves again and again "what have we done? We have created a monster."[4] We must challenge the self-serving perception of inherent innocence or benevolence of our acts. The time of innocence is over. It is time to mature, at individual and collective levels. Several organizations, even some of the largest companies, have already begun to address ethical aspects of our profession and to encourage their members and employees to assume their share of responsibility for a human-centric design and use of the technologies and systems they produce.

As an example, IEEE's Code of Ethics was recently improved and includes following commitments:[5]

We, the members of the IEEE, in recognition of the importance of our technologies in affecting the quality of life throughout the world, and in accepting a personal obligation to our profession, its members, and the communities we serve, do hereby commit ourselves to the highest ethical and professional conduct and agree to hold paramount the safety, health, and welfare of the public, to strive to comply with ethical design and sustainable development practices, and to disclose promptly factors that might endanger the public or the environment and to improve the understanding by individuals and society of the capabilities and societal implications of conventional and emerging technologies, including intelligent systems.

Beyond words, we are dedicating a lot of energy and resources, and have established several large ecosystems in order to address the issues I have been presenting to you. We have created the monumental oeuvre of Ethically Aligned Design; as a matter of fact we have invented this expression. We are going from principles to practice by tackling the issues of governance of algorithmic systems and of data sets through new series of standardization projects and through certification programs. We have started specific projects for children data governance and for a safer and more empowering cyber space

---

[4] https://www.vanityfair.com/news/2018/08/silicon-valley-engineers-fear-they-created-a-monster
[5] https://www.ieee.org/about/corporate/governance/p7-8.html

for our children. We are engaging very intensively and resourcefully with policy makers who ask for our advice and contribution, at municipal, national and international levels. Our thinking and terminology are accepted and adopted worldwide, because they resonate with the concerns and the hopes of so many people.

I hope that in the following panel discussion and during this conference I can be more specific about how a basically technical organization can walk the talk and move from "nice sounding" mission statements and principles to concrete action.

Of course, we cannot do this alone. We have to ally ourselves with those actors, like you, who also fight for regaining our identity and dignity in the digital space, and for an Internet in the service of democracy and enlightenment.

In the *Iliás* there is a remarkable episode about the fate of King Télephos, who was wounded by Achilles. His wound was not healing with time, and the oracle told him "ο τρώσας και ιάσεται,", which can be translated as "the one who hurt you will heal you." Only if we, techno-scientists, wake up and divert techno-science from being a Deleuzian *machine de guerre,* only then will we stop infecting the wound and be part of a healing process, enabling our children to live on a sustainable planet and in democratic societies.

I welcome anyone who is interested to know more about or to participate in our projects and programs to contact me or any of the IEEE staff present in this conference.

Thank you and I look forward to a day of sharing, learning, and action.

SOME INFO THAT COULD BE CONVEYED DURING THE PANEL DISCUSSION

A practical "tool" in the form of a standard that can help to achieve such goals is IEEE's standard project P7000 for *Model Process for Addressing Ethical Concerns during System Design*. This standard, currently in an advanced phase of development, aims to support organizations in "creating shared value" through technology. Organizations are becoming increasingly aware of the need to demonstrate socially responsible behavior when dealing with stakeholders, customers, regulators, and society in general. Socially responsible organizations recognize that their decisions and actions affect not just their financial bottom line, but also society and the environment. One of the principles of social responsibility is ethical behavior. IEEE P7000 builds on the awareness that organizations are societal players with great power, responsibility, and leverage. This standard is designed to work for all organizations—large, small, for profit, non-for-profit- to build better products with a much more refined and nuanced value proposition and with less risk.

There are other means to help form a foundation of trust that organizations and governments can use, such as certification programs. IEEE is currently developing such a program, known as the Ethics Certification Program for Autonomous and Intelligent Systems. Comprised of leading smart urban centers such as Vienna and Espoo, governmental agencies, such as Finland's Ministry of Finance, and multinational companies, such as Accenture, the program focuses on providing greater algorithmic transparency, accountability, and reduction of bias in autonomous and intelligent systems. The certification program offers a process and defines a series of marks by which organizations can seek certifications for the processes around the autonomous and intelligent systems products, systems, and services they provide.

Another trust challenge regarding intelligent technologies and systems is data governance—or how an individual's data is handled. Trust can be gained if organizations and governments are transparent about their data usage policies. For example, if data is needed to help AIS make better decisions, it is important that the person providing the data is aware of how their data is handled, where it is stored, and how they are used.

The IEEE has launched a number of standards projects to help organizations address such issues. Some examples include a standards project on *Personal Data Agents*—which gives guidance about the design of systems that can organize and share personal information on their own; a project on *Machine Readable Personal Privacy Terms*—which provides individuals with means to proffer their own terms respecting personal privacy in ways that can be read, acknowledged, and agreed to by networked machine agents acting on their behalf; a project for Child and Student Data Governance –that defines how to access, collect, share and remove data related to children in any educational or institutional setting where their information will be accessed, stored, or shared; an Employee Data Governance standard and several others. We are also supporting the work of and have submitted very comprehensive inputs to the German Data Ethics Commission and to the UK ICO's draft Code on "Age Appropriate Design."

IEEE, under its mission to advance technology for humanity, does not work in isolation; we are persistently lending our strengths and constituency expertise in a global network-of-networks effort to help build understanding of the challenges in this space, to inspire, and develop practical solutions.

Marisa Tschopp, an organizational psychologist said "trust is the social glue that enables humankind to progress through interaction with each other and the environment, including technology." We believe in this

function of trust in the area of technology and that ultimately only well-designed autonomous and intelligent systems in combination with data sets that fulfil certain quality criteria will earn the trust of the public. The question is, how we make sure that autonomous and intelligent systems are well designed and the data sets are of high quality. It is precisely to address this need that IEEE, in partnership with public and private sector partners from all around the world, establishes consensus-based, globally open standards and certification schemes in a transparent and inclusive manner. We are also very involved in collective efforts to establish global commons in autonomous and intelligent systems and related data sets.

IEEE SA ACTIVITIES SUPPORTING ETHICAL DATA GOVERNANCE AND AI SYSTEMS

RAISING THE WORLD'S STANDARDS

SUPPORTING ETHICAL DATA GOVERNANCE AND AI SYSTEMS

IEEE Consensus-Building Activities Relevant to the ICDPPC Declaration (1)

Guiding Principle	Theme	IEEE Programs
1	Human Rights and Fairness Principle	• Global Initiative on Ethics of Autonomous and Intelligent Systems • Global Initiative to Standardize Fairness in the Trade of Data • Digital Inclusion, Identity, Trust, and Agency (DIITA) Program
2	Continued Attention and Vigilance	• Ethically Aligned Design: A Vision for Prioritizing Human Well-being with Autonomous and Intelligent Systems (EAD) • EAD for Law/Parenting/Industry/etc. Series • The Ethics Certification Program for Autonomous and Intelligent Systems (ECPAIS) – Accountability track • IEEE P7009™ Draft Std for Fail-Safe Design of Autonomous and Semi-Autonomous Systems
3	Systems Transparency and Intelligibility	• ECPAIS – Transparency track • IEEE P7001™ Draft Std for Transparency of Autonomous Systems • IEEE P7008™ Draft Std for Ethically Driven Nudging for Robotic, and AI Systems • IEEE P7012™ Draft Std for Machine Readable Personal Privacy Terms

SUPPORTING ETHICAL DATA GOVERNANCE AND AI SYSTEMS

IEEE Consensus-Building Activities Relevant to the ICDPPC Declaration (2)

Guiding Principle	Theme	IEEE Programs
4	Ethics by Design and Privacy by Design/Default	• IEEE P7000™ Draft Std for Model Process for Addressing Ethical Concerns During System Design • IEEE P7002™ Draft Std for Data Privacy Process • IEEE P7004™ and IEEE P7005™ Draft Std for Child/Student Data and Employer Data respectively • IEEE P7007™ Draft Std for Ontological Standard for Ethically Driven Robotics and Automation Systems • IEEE P7014™ Draft Std for Ethical Considerations in Emulated Empathy in Autonomous and Intelligent Systems
5	Empowerment of Every Individual	• IEEE P7006™ Draft Std on Personal Data AI Agent • IEEE P7008™ Draft Std for Ethically Driven Nudging for Robotic, Intelligent and Autonomous Systems • IEEE P7010™ Draft Std Wellbeing Metrics Standard for Ethical Artificial Intelligence and Autonomous Systems
6	Reduction of Biases or Discriminations	• ECPAIS – Algorithmic Bias track • IEEE P7003™ Draft Std for Algorithmic Bias Considerations • IEEE P7013™ Draft Std Inclusion and Application Standards for Automated Facial Analysis Technology