

AI&Law Breakfasts

7th edition: certification of algorithmic systems


Summary of interventions


(not revised by the authors, only webinar is authoritative)

Guests: **Lord Tim Clement Jones**, former chair of the House of Lords elected Committee on Artificial Intelligence (2017-2018) (United Kingdom), **Arisa Ema**, PhD, Project Assistant Professor at the University of Tokyo (Japan), **Nicolas Economou**, Chief executive of H5 and chair of the Law Committee of the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (United States) and **Yaniv Benamou**, PhD, Of Counsel Attorney, Lecturer at the University of Geneva (Switzerland).





AI and Law Breakfasts
7th edition – Live webinar


Université
de Strasbourg


COUNCIL OF EUROPE
CONSEIL DE L'EUROPE


Certification of algorithmic systems

 **Opening by Lord Tim Clement-Jones**
Former Chair of the House of Lords Select Committee on Artificial Intelligence, 2017-


 **Arisa Ema**
PhD, Project Assistant Professor at the University of Tokyo (Japan)

 **Nicolas Economou**
Chief executive of H5 and chair of the Law Committee of the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (United States)

 **Yaniv Benamou**
PhD, Of Counsel Attorney, Lecturer at the University of Geneva (intellectual property, digital privacy and technology law) (Switzerland)

 **16 October 2020**
13.00-14.30 CET
Live on BlueJeans Events

Public event – no



<https://primetime.bluejeans.com/a2m/liv>

The 7th edition dealt with the certification of algorithmic systems, including those relating to the latest developments in artificial intelligence (AI). At a time when public opinion is very concerned about the concrete consequences of some AI applications, such as discrimination or the weakening of human agency, regulators are looking for concrete solutions to create trust among users. The idea of guaranteeing, through the intervention of an independent third party, the conformity of an algorithmic system with a certain number of rules, including human rights principles, has therefore been raised in academic and institutional literature. The purpose of the webinar on the certification of algorithmic systems was to explore in a practical way the opportunities, but also the practical issues of such a proposal.

Lord Tim Clément-Jones



Former Chair of the House of Lords Select Committee on Artificial Intelligence, 2017

Lord Clement-Jones has long experience of parliamentary affairs in the UK, EU and internationally, and has advised governments, blue chip companies and trade associations operating in highly regulated fields including financial services, utilities, health, pharmaceuticals and the environment.

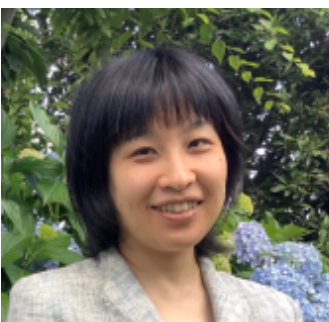
Lord Clement-Jones opened this session of AI and Law Breakfast by underlying the importance of public trust. He noticed that the last two years have seen a flourishing set of ethical principles at national and international level, and a concerted move towards operationalizing those ethical principles by reference to the risks, by application, by sector. This kind of hierarchy of risk approach needs to result in decisions on the appropriateness of a gradual approach to governance such as voluntary codes, corporate governance standards, and mandatory regulation.

However, Lord Clément-Jones is also worried by the idea that the precautionary principle should be the overall principal for risk calibration, idea he qualified as a overcautious approach.

A legal framework would have the added benefit of increasing transparency and promoting trust but Lord Clément-Jones also raises a lot of questions regarding the terms of this framework (auditors, standards, stakeholders, public sector and user roles).

Lord Clément-Jones concluded by saying that he thinks certification quality labelling or kitemarking has great potential but it's part of a wider agenda for ensuring trustworthiness and for ensuring effective regulation of AI.

Arisa Ema



PhD, Project Assistant Professor at the University of Tokyo (Japan)

She is an academic in the field of Science and Technology Studies (STS) with a focus on artificial intelligence (AI) ethics and governance. She leads and participates in various initiatives in Japan and abroad, working to ensure responsible use of AI that is inclusive and beneficial for all.

Arisa Ema is working on how to build an AI governance that includes data and algorithmic control and how to implement AI principles into practices and AI social implications to workplace and lifestyles.

She shared an overview of the different discussions in which she is involved. Dr. Ema explained that not only several institutions, including the Council of Social principles of Human-Centric AI in the Cabinet office, have issued guidelines for the use of AI, but also global and startup companies. The University of Tokyo highlights the need to look at regional differences in culture, customs and institutions while discussing AI governance, while the Japan Society for AI underlines the fact that both the AI researchers and the AI system must abide to those policies. According to Dr. Ema, some of the companies locate AI ethics and governance not as corporate social responsibility but as a framework for the management strategy. Dr. Ema also cited the Japan Deep Learning Association, that organizes two certification tests to educate on how utilize and implement ethical AI that includes questions of ethical, legal and social implication.

However, it is difficult to estimate the quality of AI services based only on one company. This is the reason why she, and her colleagues of the University of Tokyo are creating a model for a risk assessment with multi stakeholder's discussion including used cases in Japan. In this model, the basic concept is that AI risks must be considered as layers of AI systems, service providers, and users.

Dr Ema also highlights the fact that industry structure matters as well. Japan has, indeed, many businesses to business companies, making a long supply chain which raises issues about distribution of shared responsibility. Therefore, the JDLA started a study group on AI governance that aims at considering the network of ecosystem involving insurance and auditing companies as well as a whistleblowing system and a third-party committee for incident investigation. So far, their conclusion is that the nature of AI, which is constantly learning, is making difficult to produce a quality assessment. Also, the differences of impacts between sectors must be considered.

Nicolas Economou



Chief executive of H5 and chair of the Law Committee of the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems (United States)

Nicolas was a pioneer in advancing the application and governance of AI in legal systems. He leads the Law Committees of The Future Society and of the IEEE Global Initiative on Ethics of Autonomous and Intelligent Systems. He is a member of the Council on Extended Intelligence (CXI).

Nicolas Economou presented IEEE SA body and its “Global Initiative on Ethics of Autonomous and Intelligent Systems”. IEEE SA is one of the largest technical organizations working on the issue of trustworthy AI. The term “Autonomous and Intelligent Systems”,

rather than simply AI, is chosen due to its broader notion that can include all the chain of AI and also because the Association work is focused globally on emerging technological systems.

IEEE built, in collaboration with CEPEJ and the Data Agency Security, a framework, namely the “Informed Trust Framework” which provides an answer to the major question: “how to operationalize a trustworthy AI”?

In doing so, the first step to take is focusing on a proper definition of trustworthiness: it must be simple, uniform, broadly applicable and able to adapt to different cultures and environments, as well as applicable to AI future innovations.

The “Informed Trust Framework” can help to put in practice the rights enshrined in the ECHR and can provide a framework that can establish what evidence we actually need for determining the extent to which AI systems in the law (and beyond) are trustworthy, the extent to which they are effective, the extent to which the operators are competent, the extent to which people are accountable and finally the extent to which we can really have evidence of transparency.

Nicolas enshrined the four constituent elements (Principles or Trust Conditions) contained in the IEEE framework:

- Effectiveness
- Competence, which is referred not only to the operators, but also to the users. The users need to understand how AI interact and how AI systems can influence her/his perception in the light of the final decision.
- Accountability for identifying who has responsibility when an AI system fails.
- Transparency, which includes two elements: 1) access to information and 2) access to an explanation that must be adequate for different stakeholders.

In addition, Nicolas provided a case study from U.S. about the use of Technology assisted review (TAR) utilized in the civil and criminal field. Specifically, in this latter case AI is used for the review of vast group of documents. Studies conducted from 2008 to 2011 by U.S. National Institute of Standards and Technology demonstrated two things:

A) TAR technology assisted review in the US can perform more effectively than humans

B) these technologies are an effective method to properly conduct fact finding.

What Nicolas had shown is that, in the considered context of fact findings, AI can be qualified as trustworthy when the four trust conditions, included in the IEEE Framework, are properly assessed.

He concluded highlighting the mission of IEEE: develop instruments that regulators, courts, lawyers and others can use to ensure that these four trust conditions are met.

Yaniv Benamou



PhD in Law, *Of counsel* attorney, lecturer at the University of Geneva, (intellectual property, digital privacy and technology law) (Switzerland)

He is appointed as expert by WIPO for copyright and museums, executive director of the Digital Law Summer School and Expert Committee Member of the Swiss Digital Initiative. One of his current research relates to self-regulation, including certifications mechanisms and their interface with liability and public participation.

As expert in the field of digital governance and self-regulation, **Yaniv Benamou** observed a great increment in the use of certification and labelling which are self-regulation instruments.

Concerning this rise, three legal challenges were selected by Yaniv:

- i) **The first challenge is about the different types of self-regulatory models.** Firstly, Yaniv explained the difference between co-regulation, which requires a state intervention/state approval, such as in the case of GDPR or in the case of codes of conduct, vs. label which is a pure private self-regulation instrument that does not require a state intervention. As previous academic expert involved in the foundation, Yaniv brought the example of a case study: the so-called swiss digital trust label arising from the swiss digital initiative. The swiss digital trust label, as usually happens for labels, is combined with a certification trademark and a certifier or an auditor for compliance verification. Secondly, Yaniv wonders how to ensure democratic legitimacy when these models are applied. Indeed, these norms are often drafted by few decision-makers with no control and no state intervention and maybe used sometime as marketing tool or even ethics washing. The democratic legitimacy could be better ensured with procedural norms that allows the participation of all stakeholders, including consumer organizations. This kind of transparent process was followed by the swiss digital trust label.
- ii) **The second challenge is the clarification of accountability and liability of all parties involved.** The accountability/liability must be ensured for the norm drafters, as well as for the auditors and the certified users. For the norm drafters the proposal is that they shall remain accountable for compliance with laws. With respect to the auditors, they may be considered liable when the audit is not properly conducted, as several cases study demonstrated: among these cases Yaniv choose a decision of the U.S. Court in which a private certifying company is found liable for insufficient audit to companies, such as the New Times or Apple (FTC Final Order, No. C45-12, 12 March 2015, Trust-e). Finally, with respect to the certified users, they may be held liable even when they comply with the code of conduct. For instance, in 2011 ALSTOM (a transport company) has been

condemned for not preventing bribery despite the implementation of the code of ethics.

The issue concerning the second challenge is the clarification of the normative effects of the label or the certification when assessing liability of the labelled company. These normative effects may range from a mere interpretation guidance for courts to a real compliance presumption. The suggestion made by Yaniv is that for legal certainty, the better choice is to rely on certification when possible.

iii) **The next and final challenge presented concerns the way to audit algorithmic systems.** Yaniv highlighted the need for each certification or labelling to contain different auditable criteria to be used by an entity/auditing firm. Two are the possible approaches: a) a multi-layer approach with one layer contained simple descriptions understandable from standard user perspective and b) an editable layer containing detailed specification for the auditors and for the certifiers.

Another way to audit algorithmic systems is the use of most binary criteria as possible or criteria that refer to recognized norms. Yaniv acknowledges that the challenge in this latter approach is to deal with less binary criteria, such as those subjected to interpretation, for instance best practices or criteria that are evolving. In this light, criteria to consider are for instance explainability or ethics, which are used in the swiss digital trust label initiative.

Finally, Yaniv faced the issue of mandatory vs. voluntary nature of certification for high-risk technologies. The core questions are matter of politic and technical legislations that define the material scope of high-risk technologies along the all lifecycle of the AI based systems (for instance which exact item during the whole AI lifecycle shall be certified). Another question is: who is the most adequate certifier for bringing trust among the public? Is an international label certifier or a private company?

Concluding his presentation, two are the key outputs reached by Yaniv: 1) Co-regulation shall be preferred to ensure democratic legitimacy and to ensure the legal certainty with the so-called compliance presumption; 2) Auditability of algorithms is feasible but legislative action is required, because for a legal standpoint existing legislations are too limited in their scope, for instance the GDPR has the upside of being a co-regulated model with the associated enhanced democratic legitimacy and the legal certainty of compliance presumption, but it has a downside that is the circumstance of being limited to data protection and it is characterized by a lack of coordination between multiple certification bodies.

In the end, in Yaniv's view self-regulation, in particular co-regulation, is an appropriate regulatory model as it provides the necessary flexibility for technical fast evolving technologies, but only as long as the democratic control and the legal certainty are ensured.