TDL | Trust in Digital Life

Building Trust in Artificial Intelligence

11:30-13:30 , 20 March 2019

The Representation of the State of Hessen to the EU
rue Montoyer 21, 1000 Brussels

Trust in Digital Life ran a two-hour roundtable on 'Building Trust in Artificial Intelligence', hosted by the Representation of the State of Hessen to the EU.

In April 2018, the European Commission published its strategy on the opportunities offered by AI, followed in December by a 'Coordinated Plan on Artificial Intelligence'. At the time of the event, the high level group of experts appointed by the EC was expected to publish the AI ethical guidelines at the beginning of April.

The roundtable participants were asked to address some of the issues identified in the strategy for Europe. In particular:

- Making the compelling case for a successful AI-driven economy
- Addressing the perceived trustworthiness of AI systems
- Developing skills and technologies to further the widespread adoption of AI-based solutions
- Facing the ethical, cybersecurity and privacy-related challenges
- Helping Europe become a global leader in AIT

The roundtable comprised representatives from academia, industry, and the European institutions. The opportunities and challenges associated with all the topics under discussion stimulate a lively debate, sparked ideas and generated insights into how we can move forward into an AI future.

Claire Vishik - Senior Director, Trusted Technologies, Intel (Moderator)
Claire opened the session stating that its goal was to start a multi-disciplinary community of practice on AI industry, academia, economics, legal and ethics, involving people from different backgrounds with the intention of producing policy papers and whitepapers.

Going round the room in turn, participants were asked to introduce themselves as well as their interest or involvement in AI.

Irina Orssich - AI & Digital Industry, DG CNECT
Irina is working within the EC on AI, providing strategic coordination of activities with Member States and on ethics. She observed that there are many definitions of AI requiring a framework, a wide focus for the future. “Good to start broad and home in on definitions”.

Florent Frederix – Principal Administrator, Trust and Security Unit, DG CNECT
Florent is working in Unit H1 and since last year has had a new responsibility for looking at AI in the context of Horizon Europe (the research and innovation framework programme that will succeed Horizon 2020). As a Continuing Policy Fellow at the Centre for Science and Policy in the University of Cambridge, he is responsible student knowledge exchanges on AI. “Everybody is promising everything, the challenge is to discover the hype”.

Jens Jeppesen – Director, European Affairs Centre for Democracy & Technology
In his role, Jens covers a broad range of technical issues and digital positions with the goal of adopting best practices for tools and applications, and correcting bias to avoid concerns on issues. Jens introduced Vincenzo Tiani who recently joined the Centre for Democracy & Technology.

Amardeo Sarma – General Manager NEC Laboratories Europe
Amardeo is responsible the security issues associated with 5G, and, within NEC’s European labs, also for AI. His primary interest is in the trustworthiness of AI systems.

TDL | Trust in Digital Life

Milan Petković – Department of Mathematics and Computer Science, Eindhoven University of Technology
Milan is also the Head of Data Science Department at Philips Research, with expertise in Security, Privacy and Trust in Modern Data Management where his main goal is to create innovation projects n machine learning. At EIT he is chair of trustworthy AI. On the definition of AI: it's really coming from 50 years ago, whereas today's focus is more on machine learning, although we shouldn't ignore the other areas.

Paul Timmers – Senior Advisor, European Policy Centre
Paul is also a visiting research fellow at the University of Oxford where he studies cybersecurity policy and digital transformation, with an interest in AI, particularly protection against autonomous weapons. Until 2017 he was Director at the EC's Digital Society, Trust & Cybersecurity, responsible for policy, legislation and innovation in cybersecurity, digital privacy, digital health & ageing, e government, and smart cities/mobility/energy and was also a member of the management board of ENISA. Within Digital Enlightenment, he has done work on labour policy and taxation within industry and manufacturing, having an interest in anything where human decisions are involved.

Karina Marcus – Science Officer, COST Association
The European Cooperation in Science and Technology (COST) has been a funding organisation for the creation of research networks, called COST Actions, since 1971. Karina is responsible for a number of actions as applied to health from forensics, to bacteria and the process of consciousness, most of which is highly multi-disciplinary. AI is a very broad term that will ultimately be derived from what is not known – what is explainable and what is not. Karina's interest is in the latter.

Natalie Bertels – Centre for IT & IP Law (CiPiT), KU Leuven
Natalie is a senior researcher in privacy and data protection, focussing on the data protection challenges of an AI/big data/IoT/HPC setting. She is also the vice chair of the Policy and Societal Task Force of the Big Data Value Association (BDVA). She is a lawyer, and although there is no AI law yet, there may be in the future. And it's not that it's not regulated: it is mostly based on data. The challenge is how to interpret these laws – not only the GDPR but also competition law, product law et al. There will be limits and the question becomes whether to adopt a new regulatory approach or to create a regulatory sandbox. It will be very important to have multiple stakeholders from multiple disciplines debate the issues.

David Goodman (Rapporteur) – Senior Consultant, Trust in Digital Life
David said that he had written explanatory analysis on machine learning and was co-authoring a book on the impact of AI on the workplace from a numerous different perspectives, with a particular insight into the pivot points where individuals, companies, regulators have choices and decisions to make.

Eric Badiqué – Adviser for Artificial Intelligence, European Commission
Eric is newly appointed to this advisory post.

Nineta Polemi – Cyber Security, Technologies and Capacity Building, DG CNECT
Nineta contrasted how AI can be used to defend against severe attacks but also be deployed to create global attacks that are unthinkable.
We need to re-define AI in terms of automated decisions, deep learning and a bucket of concepts to handle data. The issues are:
1. How to translate/transpose regulations to engineering regulations?
2. How to aim Europe to become leaders in human-centric AI: there are many technical issues such as how to implement the quality of data. The GDPR made the EU famous – we can do the same with AI.

Leonardo Lucarno – Consultant, APCO Worldwide

Catherine Chronaki – Secretary General, HL7 Foundation
Catherine admitted to having been around long enough to see technology hypes come and disappear two-three times. In the context of heathcare, there has been a significant impact of high-quality data, as well as an increase in productivity, including shorter waiting lists and an improving use of hospitals. AI helps make clinical decisions and supplement (not replacing) the role of individuals which puts humans in the centre. How can this be achieved in an ethical way? Catherine suggested it requires standards.

Svetla Nikova – Research Expert COSIC, KU Leuven
Svetla interests are in the areas on cryptography and privacy. She started AI at KU Leuven and is following the AI HLEG (high level group). What can Europe do differently, given that we are usually behind the US? We can take a lead in privacy, ethical guidelines, avoiding bad AI scenarios. AI is a very good thing but it can also be very dangerous.

Julia van Best – Permanent Representation of the Netherlands to the EU

Theodoros Karapiperis – Head of Unit, Scientific Foresight Unit (STOA), European Parliament
Theo is head of the unit responsible for providing policy advice to the European Parliament on the future of science and technology, a remit which embraces every possible technology and horizontal. He was involved in the working group that drafted the robotics report that was published in May 2016. In September 2017, he organised an event that asked the question, how rational is to be in favour of AI, that invited Stephen Pinker as guest speaker.
This was followed, six weeks later, by a media function looking into information and disinformation in the media and new studies being published on the polarisation impact of fake news and associated accountability. He also mentioned a current workshop taking place in the Parliament entitled, "Is AI a human rights issue?"
Riccardo Masucci (Moderator) – Global Director of Privacy Policy, Intel
Riccardo's focus is on privacy policies worldwide. In looking at new technologies, and the associated aspects of policy, he seeks to combine many perspectives – one of the really good things about TDL! Starting from questions about ethics and privacy, he wants to understand what are the biggest challenges with AI?

Catherine Chronaki said that the biggest challenge is profiling, dealing with outliers, which has to be dealt with in an ethical way. Healthcare requires more regulation and sensitivity. Europe can make a difference in this respect. Are there more technologies, such as encryption and homophoric encryption that can help?

Milan Petković wanted to stress the importance of ethics, giving as an example the use of AI in US Courts to assess whether someone is likely to re-offend, noting that it tends to bias against blacks. It's in hand, the use of AI to benefit industry to move forward but the ethic has to be addressed to achieve the right balance.
Some tech projects, involving Phillips and EIT, are working on secured multi-part computation, for example the SODA project (Scalable Oblivious Data Analytics). He referred to privacy-preserving systems, that do not reveal sensitive date in healthcare, insurance and other verticals.

TDL | Trust in Digital Life

Natalie Bertels gave some examples to broaden the positive aspects of AI, such as data protection, bots, mentioning the SPECTRE (Smart city Privacy: Enhancing Collaborative Transparency in the Regulatory Ecosystem) and PRiSE (Privacy by design Regulation in Software Engineering) projects that was looking into bias and fake news.

The ethical principles of AI – what is building trust for citizens? – is part of the AI High Level Group (HLEG) and "AI for People". Are we looking at distributed environments or critical infrastructure? Experimentation is important but what are the rules? There is harmonisation as exemplified by the differing approaches to driverless cars in France and Belgium.

Irina Orssich stated the importance of harmonisation of testing in Europe, that she had been discussing that week with Member States. The Commission is putting aside 1.5 B EUR, with some Member States adopting a broad scope for testing whereas others are more discrete. She emphasised that we don't need 27/28 testing facilities: sectors should harmonise their efforts.

She said that there are many threats and challenges. Many companies have invested in algorithmic bias which is a very tricky subject and there are many well-known examples. We need, as citizens, to feel that we are still in control of our privacy whether it be email or social media, and in the race for AI-ready surveillance cameras. But we, as Europe along with other like-minded jurisdictions, do have ethical principles and laws.

Paul Timmers said that we are at a very interesting stage with AI with a combination of policies in, for example, immigration and financial investment. Healthcare is a particularly interesting case. Babylon Health is a diagnosis application which works better than humans do but there is push back because it's excluding access to health trends to the AI literate, who tend to be in the 20-35 year age demographic. Also explainability and risk management is better understood in healthcare than it is in, say, driverless vehicles: in the one, millions die whereas in the other it's in the thousands. It's contentious whether algorithms are explainable; and then what about intellectual property rights? We should healthcare to derive real use cases.

Claire Vishik asked the question: algorithms are developed by humans, to what extent do they reflect human bias, giving as an example insurance claims made when there is an air crash? Shouldn't humans be in the loop to remove any bias?

Florent Frederix said that the main challenge is engineering: in the case of the recent Boeing 737 Max crashes the problems were caused by the design. We need to set the boundaries of regulations.

Ghassan Karame said that there are many good applications of AI. But first of all, we should understand what AI can and can't do. For example, innovation can be impacted by a few dots in a spread.
1. AI so far is still not safe: access control, encryption are harder to break than AI
2. Algorithmic bias depends on the quality of data and how and who provides the content.
Our focus should be on what AI can do today and not tomorrow.

Theodoros Karapiperis asked about explainability: how do we connect to the real code, at what level, with whom and to what depth? He stressed that across the board legislation on AI is possible for which we would need 'soft instruments'. He also asked:
1. The guidelines from the AI HLEG, who is going to use them, how do they fit?
2. How could a regulatory agency be useful?

Nineta Polemi said that the Commission already had prepared calls and topics.
She asked what are the reasons for bias?
1. With reference to GDPR, data needs to be qualified, using data auditing techniques and tools. Unit H1 already has some project working in this area.
2. Who writes/certifies the algorithms? Under the recently passed Cybersecurity Act, ENISA is responsible for specifying certification schema, which is a big step towards a certification schema of systems, including AI, that will go a long way to creating the trustworthiness of systems which is a very important step towards the trustworthiness of AI. The GDPR provides the right to question, but to whom? Both the quality as well as the right amount of data is important. For example, in personal finance, it might not be enough to know just what happened in the last three years or limit the scope to Greece (say) rather than worldwide. The GDPR covers policies, the rest is open to interpretation.

Irina Orssich affirmed that the AI HLEG is very independent, and that yesterday agreed guidelines on ethics which they will be presenting in April to the European Parliament. From 9 April, it will go out to the rest of the EU and beyond. To get this far, there has been a lot of mutual exchange, having had a consultation process running from December to February amassing 3000 comments covering:

1. Ethical questions
2. Requirements, technical and non-technical, transparency and responsibilities
3. A list of assessments to transform into reality which would be a 'novelty' and could lead to a piloting phase or testing. People and companies are signing up.

The document will be revised at the end of the year and go back into an assessment/review phase.

Amardeo Sarma said that AI guidelines should be enshrined in our laws, although not having guidelines is not an excuse to break the law.

- China has a different set of laws
- With AI systems built on correlations, decisions can be bad – and at this point, we are only at the correlation stage and we are some way before the real impact will happen.

Florent Frederix observed that cybercrime use a lot of AI and would be worth talking to cybercriminals about the limitations of the GDPR.

Natalie Bertels asked how are we developing the evolution of systems with regulations and governance. It should not only be personal but with a higher level of control. Do we perhaps need a supervisory agency?

Paul Timmers said that the one mistake with the GDPR is that it missed out on innovation. We're going to be moving the needle from regulation to innovation. A key question is whether the EU is a leader in the ethical sense and also with innovation?

Milan Petkovic supported Paul and Amardeo on the matter of existing law vs new regulations. We're using a lot of AI already: Microsoft in its spam filters, Google in creating photo albums. We're making use of AI for certain applications only – why? For example, there is a great need in healthcare, amongst, say, radiologists.

TDL | Trust in Digital Life

Milan doesn't understand how MR works (even though Philips makes them) The point being that, if AI is proven in critical trials, there's no need to understand.
To achieve optimum success, we need to provide/facilitate access to enough data. What is sufficient – patient data?

Claire Vishik asked what do you do to translate high level principles into guidance on how to achieve practical aims? Not enough is being done to make AI pragmatic. There are two levels of bias:

- How to define more efficient translation. We need to use security models (as proposed by Ghassan). More common criteria tasks aren't going to help much. There are similar practices in other areas, for example, the particular types of semantics built into existing systems. Old AI neuro-networks have become much faster, but not next generation – yet.
- How to translate from principles to practicalities. Innovation will bring with it new jobs, new areas of innovation

Eric Badiqué said that there was more mileage on traditional approaches and went to differentiate between conscious and unconscious AI. Should we focus/invest on conscious AI, can we go towards data-less systems?
There is a perception that the EU is too focussed on ethics. But in fact there is also considerable investment in SMEs and hubs. There is at least one hub, one in each Member State, and not just focussed on AI. After 2020 the level of investment will rise to 20 BEUR. So an innovation strategy must be stressed.

Irina Orssich replied that it should be a holistic strategy. The first tranche should have a socio-economic dimension with international cooperation, with considerable investment – and not just ethics.

Florent Frederix said that 30 years ago, a new chip development was announced. Today that chip can provide a whole neural network. We have to act now.

Riccardo Masucci said that today we have access to so much data but how can public/private players improve the current situation?

Svetla Nikova said that data provides a lot of opportunities, not only for the good guys but also for malicious players. However, we have privacy by design as well as security by design.

Catherine Chronaki told that HL7 has done a lot of work on FHIR (Fast Healthcare Interoperability Resources), a horizontal on digital health

- A risk assessment on each algorithm, 'computable ethics', which could be incorporated into an API to make a federated model
- Regulation – 'data cooperatives' in a transparent framework

Natalie Bertels came back to access to data and talked about creating supportive elements to data and also creating structures such as data commons, platforms and governance models. There is no one legal framework that is able to manage personal with non-personal data.

Florent Frederix referred to the question of liability with cars, comparing a person who is used to driving in the country then driving in Paris. This is an example of context sensitive learning and there needs to be context sensitive AI – can we develop appropriate guidelines?

Karina Marcus said that not only data but also labelling must be correct. The order of data in a training context can impact output.

Claire Vishik summed up the roundtable session saying that we'd heard quite different perspectives based on different types of expertise. The technical aspects had been aspirational, whereas the discussion had covered more abstract issues multiple stakeholders challenging each other's objectives and identifying where some of the gaps are.

For **TDL**, this was the first activity of its new AI working group and looks forward to other opportunities to meet, network, share ideas and work towards an ethically-sound and trustworthy AI future.

Claire Vishik - Senior Director, Trusted Technologies, Intel (Moderator)

Claire opened the session stating that its goal was to start a multi-disciplinary community of practice on AI industry, academia, economics, legal and ethics, involving people from different backgrounds with the intention of producing policy papers and whitepapers.

Going round the room in turn, participants were asked to introduce themselves as well as their interest or involvement in AI.

Irina Orssich - AI & Digital Industry, DG CNECT

Irina is working within the EC on AI, providing strategic coordination of activities with Member States and on ethics. She observed that there are many definitions of AI requiring a framework, a wide focus for the future. "Good to start broad and home in on definitions".

Florent Frederix – Principal Administrator, Trust and Security Unit, DG CNECT

Florent is working in Unit H1 and since last year has had a new responsibility for looking at AI in the context of Horizon Europe (the research and innovation framework programme that will succeed Horizon 2020). As a Continuing Policy Fellow at the Centre for Science and Policy in the University of Cambridge, he is responsible student knowledge exchanges on AI. "Everybody is promising everything, the challenge is to discover the hype".

Jens Jeppesen – Director, European Affairs Centre for Democracy & Technology

In his role, Jens covers a broad range of technical issues and digital positions with the goal of adopting best practices for tools and applications, and correcting bias to avoid concerns on issues. Jens introduced Vincenzo Tiani who recently joined the Centre for Democracy & Technology.

Amardeo Sarma – General Manager NEC Laboratories Europe

Amardeo is responsible the security issues associated with 5G, and, within NEC's European labs, also for AI. His primary interest is in the trustworthiness of AI systems.

TDL | Trust in Digital Life

TDL's vision is that trust must become an intrinsic property of any online transaction involving personal information, incorporating legal, business, and technical advances, supporting cyber security policies, and integrating societal considerations so that citizens and end users will recognize trustworthy services, transactions, and data, and be prepared to pay for them. Trustworthy ICT will increase confidence and trust in modern society, bring new and attractive ways of living and working, and further strengthen Europe's democratic and social values.

The association's mission is to provide its members with a European business development platform in order to stimulate development and user acceptance of innovative but practical trustworthy ICT. Guided by its strategic research agenda, TDL acts as an incubator for a portfolio of sprint projects intended to validate new and innovative technology concepts, promotes cross-sector collaboration, and aggregates the results into industry recommendations for policy makers and the European Commission.

trustindigitallife.eu