Data sustainability in the age of Artificial Intelligence

Conference of February 14th 2022, 10h30 -12h.

On 14th February 2022, the Brussels Data Privacy Hub and EDHEC Business School co-hosted the Conference on Data Sustainability, following last year's edition. This year aimed to look at data sustainability through the prism of Artificial Intelligence (AI). Speakers were **Ben Wagner** (TU Delft & Sustainable Media Lab, Inholland), **Christian D'Cunha** (European Commission, DG Connect, Cybersecurity and Digital Privacy Unit) and **Ivana Bartoletti** (Wipro, University of Oxford). **Alessandra Calvi** (Brussels Privacy Hub, VUB, CYU) moderated the discussion.

The conference started with an introduction by **Gianclaudio Malgieri** (Brussels Privacy Hub, VUB, EDHEC). He first addressed the definition of the word "sustainability", whose etymology roots in the Latin word *sustinere*, which means to hold up, to bear, to tolerate, to exist despite something. About data privacy, this brings up the question of data privacy current systems' ability to exist in the future and under what conditions.

He highlighted that, just like environmental matters, data privacy should be addressed at a collective scale, and not individually. "Digital information is the fuel of the new economy. But like the old economy's carbon fuel, it also pollutes" (*Data pollution*, by Omri Ben-Shahar). This marks the introduction of the comparison between data sustainability and environmental sustainability.

Whereas the traditional sustainability scheme builds upon the triangle Economic-Social-Environment, G. Malgieri suggested that "Data" substitutes "Environment".

Going on with the parallel between environment and data, he highlighted that both address a problem of negative externalities (private and collective), and they also use similar tools (impact assessment, certifications…).

There is more to data sustainability than the environmental aspects

B. Wagner stressed that despite data privacy professionals are trying to reflect on the connection between data privacy and sustainability, there are no obvious answers to this issue. Too often, people associate sustainability only with the environment, whereas there is more to it.

C. D'Cunha added that lately discussions on AI have replaced Big Data in the public debate, but these two issues are substantially the same. They are both resource-intensive, have an impact on the environment and also on people.

He questioned to what extent the current path followed by governments is sustainable, considering that a truly sustainable approach would address at the same time social justice, human rights and the environment. He also put forward the need for a sense of scale, as the problems of data processing are not equally distributed on the planet.

I. Bartoletti pointed out that the discussions on sustainability have multiplied both at a policy and business level. Yet, this bears the risk of greenwashing and ethics washing. Avoiding falling into this trap would require a structural shift into our approach to technology, less technocentric and more collective.

Data processing: both a positive and a negative impact on the environment

B. Wagner noted that the simplification of our lives through technologies, in turn, complexifies data processing and requires more energy. Software and data processing are getting all more costly in terms

of energy. He pointed out a tendency to develop media systems that are less sustainable than the previous ones. The only way to tackle this issue seems to be their limitation.

He also suggested regulating at a European Union (EU) level the emissions of data centers. Indeed, whereas the EU appears to favor the implementation of local data centers due to privacy reasons, data storage is particularly pollutant in Europe.

I. Bartoletti noted that even if AI might help solve certain environmental issues, it is not a panacea. AI has a terrible impact on the environment. She suggested that a proportionality framework should be considered in the current AI regulation proposal. Also, the task performed by an AI solution should be coherent with its ecological footprint.

She emphasized that this topic should be addressed at a political scale and the choice of the more ecological options should not solely be left to the consumer. She noted how the lack of information on AI's environmental impacts depends on the lack of incentives for companies to release any information on their global environmental impacts. Thus, a first step would be a clear political recognition of the damage AI can cause to the environment.

There is a public interest in improving the data processing by AI

B. Wagner spotted a difficulty in addressing both collective rights and individual rights under existing legal frameworks. He underlined that the interpretation of rights is rather individualized now, leaving a limited place for collective challenges. He doubted that a market-based approach involving the creation of costs related to unsustainable practices would be sufficient to create an individual incentive for sustainable choices. He underlined that engaging with users is quite difficult because people need concrete tradeoffs to accept more sustainable solutions.

C. D'Cunha called attention to the fact that AI is about inputs and outputs: data is the key element in the use of AI. He thinks such data processing should be used to serve mankind. However, the General Data Protection Regulation (GDPR) is not easily leveraged to assert collective rights.

Whereas the potential for AI would be to track non-personal subjects as pollution, nowadays the tendency is to believe that AI's stake is decoding and recoding the human brains and manipulating them. The fact that facial recognition is a central topic of the AI act proposal demonstrates a lack of accuracy in priorities.

On regulation efficiency, he also highlighted that the data protection authorities want to prohibit some uses of data rather than impose sanctions, especially given that for large companies a sanction does not have an impact significant enough to incent a change of practices. He also pointed out that international action is unlikely to happen because of the fragmentation of the approaches to data around the world.

On this, **I. Bartoletti** stated that the EU has a very strong counter push against the rush to data collection. Indeed, there is no need for a gigantic amount of data but rather for smart-sized data-centric solutions to big issues in AI. She also thinks that the AI process should be challenged more: what does AI do? Does it improve the well-being of the population? What is its impact on social, economic, legal aspects? Can we have a proportionality test?

She suggested that human intelligence and virtual intelligence should be combined more. Human input in the AI process would improve the quality of the data and a more meaningful human intervention could be useful.

Environment and data privacy defenders should realize they should merge their fights

B. Wagner observed that the context in data privacy and environment is similar as both branches suffer the influence of powerful corporations willing to curb change on a policy and governance aspect. He suggested that users question their trust in online platforms in the same way as in the main oil producers to reach a governance model that could be appropriate both from an environmental and a data privacy perspective.

C. D'Cunha's reaction to this was that a way must be found to connect the environmental impact of data practices with individual privacy and suggested that the principle of data minimization could be helpful to do so. He witnessed that there is no real interaction between data protection authorities and environmental groups for now but this is going to be imperative soon.

AI is an integral part of programmatic advertising. The entire ecosystem of online advertising generates substantial amounts of data and distribution of data to intermediaries with a significant cost to privacy. In this regard, the Belgian Data Protection Authority asserted that the existing IAB Cookie consent framework is illegal under GDPR. C. D'Cunha underlined that this system is also very wasteful, considering that this industry is responsible for the emission of 160 million tons of CO2. He raised the question of the acceptability of such a situation, given the environmental and privacy costs.

Concerning the principle of data minimization, the GDPR does not stop the processing of data, but it is supposed to give controllers the reflex of evaluating the impact of their actions before taking them and avoid using personal data when it is not needed. And if that principle was effective, the carbon footprint would be much smaller.