



## **REPORT**

### **High-Level Meeting with philosophical and non-confessional organisations “Artificial Intelligence: addressing ethical and social challenges”**

18 June 2018

The meeting gathered representatives from 12 non-confessional and philosophical organisations with representatives around Vice-President Andrus Ansip, Commissioner for Digital Single Market. The meeting looked at the issue of Artificial Intelligence (AI) and in particular at the ethical and social impact of AI. The aim was to address the potential impact of AI on fundamental rights: privacy, dignity, consumer protection and non-discrimination. It was also to look at the social dimension of AI, the impact on social inclusion and the future of work.

In his welcome address, **Andrus Ansip, Vice President for the Digital Single Market** explained the approach of the Commission which has recently launched its AI strategy. Access to data is at the centre of this strategy and it is necessary to encourage public as well as private sector, and people in general to use more systems based on AI – in particular in view of the competition coming from China and the US. At the same time, we need to have a discussion on ethical principles in relation with artificial intelligence. We also need to talk about the social aspects. As we all know, all human inventions may be used for good as much as bad purposes. This is the reason why need to ensure that new skills to deal with AI are made available to everyone.

Participants in the meeting discussed risks and opportunities with artificial intelligence and the connected issues of access to data and “deep learning” abilities of AI based on algorithms. In particular, the importance to remain alert to the protection of basic rights was mentioned – which means more transparency and training about algorithms.

A number of issues were particularly highlighted:

- ‘Algorithmic governance’: several participants underlined the need for appropriate oversight mechanisms. In view of their impact, algorithms should be decided following a collective deliberation – to which some replied that algorithms being very technical matters it is difficult to communicate them to a wide public (the most important is to ensure that there is pluralism to ensure that one can get access to different sources of knowledge). The mechanisms of control over AI must be permanent in view of the rapid evolution of AI and the risk of unintended consequences resulting from the use of AI. Such control must come from both public institutions and civil society. An ‘ethical audit’ of AI systems was proposed.

- Several participants underlined the need for a global approach to AI, in particular at UN level, as the absence of international consensus would lead to the undermining of ethical principles. VP Ansip mentioned G7 discussions on AI: a global approach is essential to ensure legal certainty. We need a “human-centric” AI – noting the commonalities between the views of religious and non confessional organisations.
- The risk of bias (leading to discrimination for example) in algorithms was mentioned. AI is based on machine learning must reflect social pluralism: the bias noted in AI applications is self-inflicted, as it results from an initial design of algorithms which was not correct. This raises also the issue of responsibility for AI actions – we need to define a framework for ‘responsibility without fault’ (as it already exists for public institutions for example).
- Equal access to AI was also mentioned, underlining the gender dimension in particular when it comes to the bioethical aspects (not necessarily solely linked to AI) and employment (with the risk the ‘gender gap’ increases with changing employment patterns). At the same time, it was noted that (if it is well designed) AI can be blinder to prejudices than humans.
- More generally, the social impact of AI was mentioned with the risk of concentration of wealth based on access to AI (VP Ansip mentioned the need for the EU to step up its investments – giving the example of Microsoft which is already dedicating 12 Bio € per year to research on AI).
- The use of data is unescapable (like the petrol to feed the engine) – but we need also to ensure protection of personal data. One participant asked what the alternative is if one refuses to accept the use of its personal data. Many participants insisted on the need to protect individual autonomy (in particular in the case of health). A distinction was made between ‘thinking’ and ‘doing’: AI is useful to gather data, but they may not be good at understanding the contexts and notion of ‘fairness’. Decision must remain human.
- It was generally agreed that innovation is a positive development, but it must be accompanied by a consensus on values. Value-setting must be accompanied by learning – not only specialised training (which can become obsolete) but also with a citizenship education based on critical thinking.

During **lunch**, the discussion focused on the follow-up and how the Commission is approaching the future discussion around AI. The representative of DG RTD presented the work of the European Group on Ethics in Science and New Technologies (EGE) and their Statement on AI, Robotics and ‘Autonomous’ Systems’ of March 2018 that calls for the launch of a process that would pave the way towards a common, internationally recognised ethical and legal framework for the design, production, use and governance of AI. The statement also proposes a set of fundamental ethical principles, based on the values laid down in the EU Treaties and the EU Charter of Fundamental Rights. The

representative of DG CNECT referred to the launch of the European AI alliance (to which all participants were invited to subscribe) and the drafting process of ethical guidelines which will be overseen by the high-level group of experts on AI.

An interesting discussion followed on the link between values and knowledge and how one cannot be separated from the other. Knowledge is not only a mechanical addition of facts. Ethics act as 'filters' when apprehending reality. The initial design of AI is thus important. As algorithms are very technical, what is essential is not much transparency (which may also have downside as it exposes AI systems to risks of intrusion) but pluralism in the use of data.