Recommendations of the Data Ethics Commission for the Federal Government's Strategy on Artificial Intelligence

9 October 2018

The Data Ethics Commission is pleased that the Federal Government is developing a strategy on artificial intelligence. At its constitutive meeting on 4 and 5 September 2018, the Data Ethics Commission discussed the Federal Government's policy paper for such a strategy. The Commission recommends that the Federal Government should add the following points to its strategy:

(1) the objective "Upholding the ethical and legal principles based on our liberal democracy throughout the entire process of developing and applying artificial intelligence"

(2) the area of action "Promoting the ability of individuals and society as a whole to understand and reflect critically in the information society"

I.

The term "artificial intelligence" (AI) is used in the media and in general discourse to refer to different things. The Federal Government's policy paper does not specify the technologies covered in the paper. This information should be added.

In this context, we understand "artificial intelligence" as a collective term for technologies and their applications which process potentially very large and heterogeneous data sets using complex methods modelled on human intelligence to arrive at a result which may be used in automated applications. The most important building blocks of AI as part of computer science are sub-symbolic pattern recognition, machine learning, computerized knowledge representation and knowledge processing, which encompasses heuristic search, inference and planning.

The range of applications using AI today is already enormous. These applications range from the simple calculation of travel routes to image and language recognition and generation to highly complex environments for making decisions and predictions and for exerting influence. The most important applications involve systems which recognize language and images; collaborative robots and other automated systems (cars, aircraft, trains); multi-agent systems; chatbots; and engineered environments with ambient intelligence. We expect increasingly autonomous and comprehensive applications to be developed which will pervade all areas of life and are capable of automating, (partly) replacing and far outperforming more and more human activity in ever broader fields of action.

Questions with ethical and legal relevance which arise in this context also concern "simple" systems of rules based on algorithms "manually" defined by experts (= rules). These do not constitute AI as it is generally understood. It is important for the Federal Government's strategy on AI to cover these processes as well.

II.

The diversity of possible AI applications and the complexity of the relevant technologies make it especially challenging to design them in compliance with ethics and the law and to regulate this compliance. As more and more decision-making processes are shifting from humans as the subject of action to AI-driven systems, new questions arise as to who is responsible for the development, programming, introduction, use, steering, monitoring, liability and external review of AI and applications based on it. Further, the specific functioning depends on the selection and quality of the data entered and/or used to "train" the application. Simply ignoring certain types of data and using poorly prepared data can have ethical consequences extending to systematic discrimination or results antagonistic to plurality. In this context, more support should be given to research into modern methods of anonymization and into generating synthetic training data, also in order to increase the amount of data that can be processed for AI technologies without threatening fundamental rights.

The data needed for some AI applications are highly concentrated among a small number of companies which also possess a high level of technological expertise. This raises the question as to whether and how access to non-personal data in private hands should be regulated by law.

Finally, with regard to the democratic process, it should be noted that technology which is increasingly able to imitate human behaviour in a remarkably convincing way can also be easily used to influence social trends and political opinions.

Ethical considerations should be addressed throughout the entire process of developing and applying AI, using the approach "ethics by, in and for design" and as the trademark of "AI made in Europe". This includes research, development and production of AI, as well as the use, operation, monitoring and governance of AI-based applications. For the Data Ethics Commission, ethics does not mean primarily the definition of limits; on the contrary, when ethical considerations are addressed from the start of the development process, they can make a powerful contribution to design, supporting advisable and desirable applications.

It is also necessary to consider interactions between technology, users and society ("the AI ecosystem"). Within this ecosystem, it is necessary to ensure sufficient transparency, accountability, freedom from discrimination and the ability to review those automated processes which prepare decisions or draw conclusions which may be carried out without additional human input. This is the only way to generate trust in the use and results of algorithm-driven processes. The policy paper (p. 9) rightly demands these measures for algorithms used in public administration. But the same principles should apply in an appropriate way to private parties as well. Measures for quality assurance are also needed which can be supported in part by independent third parties and in part by automated processes. It is also necessary to ensure that the persons affected and the supervisory authorities have appropriate and effective possibilities to intervene as well as access to effective legal remedies.

The most important standard for dealing responsibly with AI is first of all the Constitution, in particular the fundamental rights and the principles of the rule of law, the welfare system and democracy. This also includes individuals' right to self-determination, including control over their personal data, which also requires companies to inform their customers how they use their data; respect for individual user decisions concerning personal use of an application; protection against unfair discrimination; and the possibility to review machine-made decisions effectively. We also need legal provisions which clearly define the extent of responsibility for developing and applying AI-based technologies

according to ethical, legal and economic principles. This also applies to compensation for damage and the enforcement of public-law obligations with regard to AI.

A wide range of control mechanisms necessary for inserting ethical and legal principles into the process of designing and applying these technologies is conceivable. These mechanisms must be decided at national and European level in a democratic process. The use of AI by government actors must be subject to special oversight. Possibilities for supervision include targeted (material, etc.) promotion of applications which comply with the Constitution, certification and standards, official authorization of supervision and institutions to uphold fundamental rights and ethical rules related to AI and binding law.

With this in mind, the Data Ethics Commission recommends that the Federal Government's strategy on artificial intelligence should promote and demand attention to ethical and legal principles throughout the entire process of developing and applying AI, and that the strategy should include this as an additional objective. The strategy's areas of action should be defined with this objective in mind.

<div align="center">III.</div>

Information and technologies of all kinds pervade every level of society and our lives to a degree never before known. They increasingly influence social interactions and discourse as structurally relevant elements of democracy. The rapid development of new applications for AI also demands a constant process of critical examination. These profound and diverse changes are significant not only for individual expression, but also for our life in society. They make a discourse which reinforces freedom and democracy more necessary now than ever. Among other things, we need a framework in which individuals and institutional actors can acquire sufficient digital and media literacy and the ability to reflect critically on how to deal with technical innovation.

The Federal Government's policy paper already calls for implementing its strategy on artificial intelligence in constant dialogue with representatives of the research community, civil society and business and industry, as well as with policy-makers, in order to establish a culture of AI in Germany which promotes trust. The Data Ethics Commission underscores the importance of these measures. It also recommends adding to the AI strategy a separate area of action: "Promoting the ability of individuals and society as a whole to understand and reflect critically in the information society". This is intended to ensure that individuals and institutional actors acquire sufficient digital and media literacy and the ability to reflect critically on how to deal with AI. Such abilities are essential for society to conduct an objective, informed and nuanced examination which can help promote trust in the use of AI. However, the Data Ethics Commission believes a broader approach will be needed than is currently described in the Federal Government's policy paper.

Ways to promote digital and media literacy and critical reflection range from offering comprehensive, objective information in campaigns (e.g. to explain realistic application scenarios), to teaching media literacy at school and in adult education courses, to using and promoting technologies to enforce the law and uphold ethical principles in the world of technology. The media and institutions of media supervision also have an important role to play in this context: Their role is not only to inform

society about new technologies and examine technological progress critically, but also to provide new forums for debate.

Investment in technology impact assessment must increase to the same extent as technologies such as AI are applied in our society. For example, more research and development should be conducted on data portability, interoperability and consumer enabling technologies; these include AI applications whose primary aim is to help consumers make everyday decisions.

And a balance must be found between the state's responsibility for creating and enforcing framework conditions, which ensures trust, and the freedom, autonomy and responsibility of users and others affected by the new technologies on the one hand, and the forces of the market and competition on the other hand. This balance must be discussed and determined by society in light of these changes. The growing economic strength of those companies which play a major role in the development of AI must not result in research and civil society becoming increasingly dependent on funding from precisely these companies. Government must enable research and civil society to make independent and competence-based contributions to this important societal discussion.

As modern technologies, including AI, evolve and relieve humans of certain chores, we not only gain new skills but also lose existing skills. This demands a discussion of our responsibility to preserve and develop certain skills for the next generation to remain independent. So we also need to discuss the definition of and requirements for sovereignty of the entire society.

The Data Ethics Commission therefore recommends including another area of action in the strategy focused on creating appropriate framework conditions to promote the ability of individuals and society as a whole to understand and reflect critically in the information society.

IV.

Progress and responsible innovation make a major contribution to the prosperity of society. They offer enormous opportunities which we should welcome and promote, but they also come with risks. These opportunities can make a lasting contribution to freedom, justice and prosperity above all when people's individual rights are protected and social cohesion is strengthened. With this in mind, the Data Ethics Commission strongly recommends adding the two items referred to at the beginning of this document to the Federal Government's strategy on artificial intelligence.

Contact persons:

Prof. Dr. Christiane Wendehorst   Prof. Dr. Christiane Woopen

Co-spokespersons of the Data Ethics Commission

Contact: christiane.wendehorst@univie.ac.at; christiane.woopen@uni-koeln.de