Rathenau Instituut

Response of the Rathenau Instituut to the European Commission's White paper on Artificial Intelligence

15 May 2020

Introduction

In February 2020 the European Commission published its *White Paper on Artificial Intelligence: a European approach to excellence and trust*[1]. The Commission called on societal actors to respond to its proposals. In this response the Rathenau Instituut[2] will comment on the ideas brought forward by the Commission. This response is based on the work the institute has done in the fields of ethics and artificial intelligence (AI) and research and innovation ecosystems.

The Rathenau Instituut appreciates the opportunity to comment on the white paper, as it deals with both vital and complex challenges. In dealing with these challenges, the Rathenau Instituut advises the Commission to uphold fundamental rights and public values and recognize that only responsible AI innovations will lead to a sustainable future, in which economic, ecological and societal interests are balanced.

The Rathenau Instituut proposes the following recommendations:

1. Define AI policies in relation to societal challenges.
2. Integrate the ecosystems of excellence and trust.
3. Invest in a distributed but coordinated research capacity.
4. Set up a central agency that coordinates knowledge sharing between regional research centres.
5. Take local contexts into account, especially in the public sector.
6. Take international responsibility.
7. Guard against the pitfalls of a risk-based approach.

1. Define AI innovation policies in relation to societal challenges

We advise the European Commission to define and design its AI policies in relation to societal goals and challenges. The white paper is framed primarily in terms of technological opportunities, positioning AI as a key technology that offers practical

---

[1] European Commision, White Paper on Artificial Intelligence: a European approach to excellence and trust, 19 February 2020. https://ec.europa.eu/info/publications/white-paper-artificial-intelligence-european-approach-excellence-and-trust

[2] The Rathenau Instituut is a Dutch think tank that researches the societal impact of developments in science and technology and facilitates public debate. It supports parliamentary decision-making and publishes in the areas of digital technology, health care, education and the science ecosystem. The Rathenau Instituut is based in The Hague, the Netherlands. It is funded by the Dutch ministry of science and education and has an independent board and free mandate. https://www.rathenau.nl/en

applications in a range of domains, such as healthcare, the energy sector, manufacturing and security. However, it does not define the problems these various domains are currently facing. The white paper takes a technology-driven focus, and thereby misses the opportunity to put societal needs front and centre, and match technological opportunities with clearly defined, domain-specific challenges – as is the ambition of 'third-generation' innovation policies that currently emerge (Rathenau Instituut, 2020a).

As stated by the Commission, the ambition of European research and innovation policies is 'to address global challenges, including climate change and the Sustainable Development Goals'[3]. In order to achieve this goal, we recommend that the European Union takes societal challenges, such as those listed in the Sustainable Development Goals of the United Nations, as the primary starting points for its research and innovation policies – including its AI strategy. This also requires thinking about and developing new policy innovation instruments to realise mission-oriented innovations in practice (Rathenau Instituut, 2020a).

2. Integrate the ecosystems of excellence and trust

The ecosystems of excellence and trust form the two main building blocks of the white paper. The Rathenau Instituut strongly appreciates the combination of these two aspects in the Commission's AI policy, and encourages the Commission to integrate them. Fostering the connection between innovation policy and values is vital to solving societal challenges, and corresponds with the principles of responsible research and innovation (Rathenau Instituut, 2019a; Rathenau Instituut, 2019b).

An innovation can only be successful if it becomes embedded in the societal contexts in which it has to function. That requires early anticipation of societal needs, requirements and expectations. And you have to make sure that the innovations fit legal frameworks and institutional settings, that they are part of production chains and markets, and that they are embraced by users and citizens (Rathenau Instituut, 2019c; Rathenau Instituut, 2020b). It demands expanding the focus of R&D programmes in three ways: in terms of 1) *what* is being researched; 2) *who* is involved; and 3) *how* research and innovation processes are organized. This means, for instance, that the R&D agenda incorporates societal feasibility, that users and civil society organisations are involved in the governance of the programme, or that participatory processes are organized to develop the agenda. These demands are especially relevant to innovation in the public sector (see recommendation 5).

3. Invest in a distributed but coordinated research capacity

A key ambition of the white paper is to focus research and innovation efforts to improve upon the current 'fragmented landscape of centres of competence' that do not reach the scale to compete with the world's leading institutes (section 4B in the white paper).

---

[3]   European Commission 2020, Orientations towards the first Strategic Plan for Horizon Europe, p5, December 2019. https://ec.europa.eu/info/files/orientations-towards-first-strategic-plan-horizon-europe

Therefore, the EU proposes to develop a 'lighthouse centre of research, innovation and expertise that would coordinate these efforts and be a world class reference of excellence in AI'.

The white paper does not specify whether Europe should concentrate substantial research capacity at one physical location, or rather invest in a distributed network of research centres across Europe. At first glance, a physical centre in one central location may seem attractive, because it could benefit from advantages of scale, and develop a global reputation. Geographical proximity would also promote close collaboration across organizational boundaries, including public-private partnerships.

However, the Rathenau Instituut would prefer a more distributed network of research centres. This would fit better with the intended financial model in which European investments should leverage investments from industry and national governments. Moreover, a more distributed research capacity facilitates working towards a robust societal embedding of AI innovations, as facilities are set up in proximity to a more diverse range of local contexts (Rathenau Instituut, 2020c). Such a distributed research capacity can generate robust knowledge about AI's potential to solve societal problems.

A second advantage of geographical distribution is that it enables a larger set of European countries and regions to benefit from the knowledge and spill-overs from the funded AI research and innovation projects. At the moment a number of national governments in central and eastern Europe feel disadvantaged by the current unequal distribution of Research & Innovation funding across member states.

4. Set up a central agency that coordinates knowledge sharing between regional research centres

In a distributed network of research centres, dedicated efforts are required in order to translate and exchange lessons across different contexts (Rathenau Instituut, 2020c). For each research centre should both work in close interaction with its local context, and serve as an effective part of a European research network. In order to secure that local solutions will also work effectively on a larger scale, research centres should exchange their results with the other centres in the network, sharing both technical results and their experiences with embedding technology in a societal context.

To achieve such knowledge sharing, a programmatic and coordinated approach on the European level is needed. In addition to investments in the distributed research centres, a substantial budget should therefore be reserved for a central coordinating agency that oversees the activities across the various centres, and facilitates mutual learning. In this way, EU research can achieve critical mass and compete globally. Without proper coordination, the EU's significant research investments will not lead to the scientific breakthroughs and economic gains the EU is hoping for.

5. Take local contexts into account, especially in the public sector

According to developers, scientists and policy makers, AI has the potential to solve difficult challenges in the public sector, for instance in health care and education. The white paper currently takes a rather narrow perspective, suggesting that companies develop AI applications and the public sector will adopt and implement them. However, research shows that many AI innovations in the health sector fail, as they are often not based on the specific conditions and needs of a particular local situation, and because there is limited interaction with local actors, such as doctors and patients (Rathenau Instituut, 2020d; Rathenau Instituut, 2019d). Therefore, involving stakeholders from local contexts is essential to successful innovation, especially in the public sector.

Innovation in the public sector poses specific challenges. For example, data of sufficient quality to train AI applications is often not available, for instance due to substantial biases in existing medical data that are only available for limited subpopulations. And even when an AI application has been trained with a high quality data set, the application still needs to be validated with local data. Finally, the data required to train algorithms often involves highly sensitive citizen data. Therefore the Rathenau Instituut proposes a decentralized approach to responsible data management in which both the quality of local data sets is promoted and the rights of citizens are safeguarded (Rathenau Instituut, 2019d). In order to achieve this, new AI technology can be employed that minimizes data usage, but still enables algorithms to be trained (such as federated learning).

Because of the high demands of innovations in the public sector, the European Commission should focus on those business models and innovations that are viable and truly contribute to solving to challenges in the public sector (Rathenau Instituut, in preparation)[4]. Often, in innovation policy the uptake of technologies in the public sector takes preference over a careful consideration of the problems at hand. The Rathenau Instituut urges the Commission to facilitate a debate about the question what kind of high quality care, education, public administration and so on, Europe should strive for.

6. Take international responsibility

The white paper makes clear that the European Union should take a leading role in building international alliances concerning ethical AI development (section 4H in the white paper). The Rathenau Instituut encourages the Commission to realise this ambition, and advises the Commission to take three points into account: recognise the importance of global norms, be aware of abuse and practice responsible data management.

Firstly, the importance of global norms is hard to overstate. AI technology is developed in a global market, by a multitude of actors. All these actors have to work in unison to

---

[4] Report in preparation for RECIPES, H2020: Precautionary principle and innovation in Europe: AI for clinical decision making in health care.See also https://www.rathenau.nl/en/terugblik/gesprek-over-voorzorg-veiligheid-en-verantwoordelijkheid-bij-nieuwe-technologie (Dutch)

promote ethical AI applications, and agreeing to global norms makes this cooperation possible. Therefore, the Commission should support that the norm development takes place at forums such as UNESCO and ISO. In addition, it should support the development of codes of corporate responsibility and AI, such as the work done by the OECD on ethical data flows (Rathenau Instituut, 2019a; Rathenau Instituut, 2019b; Rathenau Instituut, 2018).

Secondly, as part of its international responsibility, the Commission should develop policy aimed at ensuring that dangerous AI technology, such as hacking tools and biometric software, does not fall into the wrong hands. Sadly, ethically concerning AI applications have been used by dictators and cybercriminals around the world – and sometimes these applications are, wholly or partly, developed within the European Union. The Commission should evaluate whether the current legal framework, and in particular the *dual use* framework, suffices to prevent the abuse of dangerous AI technology (Rathenau Instituut, 2019e).

Thirdly, international responsibility also includes responsible data management practices within the European Union. That means that European standards and values should be upheld in the collection, processing and application of data. European Union should seize the opportunity to set the tone for responsible, decentralised data management practice worldwide. Moreover, European companies should be facilitated to compete on the global market with AI applications based on European values and standards. This will also strengthen Europe's technological sovereignty.

7. Guard against the pitfalls of a risk-based approach

In the white paper, the Commission sketches the scope of a future EU regulatory framework, based on a risk-based approach. The Rathenau Instituut shares the Commissions' commitment to responsible AI innovation, taking all relevant ethical aspects into account (Rathenau Instituut, 2019a; Rathenau Instituut, 2019b; Rathenau Instituut, 2018). However, regarding a risk-based approach, the Rathenau Instituut advises the Commission to consider the three following points: hold on to the principles of proportionality review, provide clarity about the meaning and requirements of the law, retain the possibility to ban certain AI technologies outright.

Firstly, we believe that the principles of proportionality review, as developed in the judgments of the European Court of Human Rights, are relevant to all AI innovation, whether innovations are categorized as 'high-risk' or 'low-risk'. AI technologies should only be developed and used if they are suitable to achieve a certain goal, if the infringements on fundamental rights are sufficiently balanced by the societal gains they generate, and if there are no available ways to achieve a certain societal goal without infringing on fundamental rights. In practice, actors sometimes do not take these central principles into account, in some cases resulting in Courts or supervisory bodies judging the application as unlawful. Distinguishing between low-risk and high-risk AI-applications should not lead to disregarding these principles in evaluating low-risk technologies, especially since unintended developments might make certain applications riskier in time – for instance because they collect data that can be exploited in new ways.

Secondly, it is crucial that a risk-based approach does not lead to uncertainty about the meaning and requirements of the law. The history of AI technologies, for instance concerning the requirements of consent and transparency, teaches us that when there is legal uncertainty, developers will not necessarily hold themselves to the strictest interpretation of the law. Therefore, it should be very clear what governments expect of AI developers, and the risk-based approach should contribute to this much-needed clarity.

Thirdly, and finally, the risk-based approach should not lead to a refusal to ban certain AI technologies outright. Certain AI applications, for instance concerning facial recognition used in public spaces, raise such grave ethical questions, that the option to forbid these technologies should be on the table. Implementing a risk-based approach should strengthen authorities' ability to take strong measures, if the ethical risks demand it.

In conclusion

Over the past years, the European Union has made significant gains in protecting fundamental rights and realizing public values in the digital world. To mention one example, the GDPR has succeeded in not only protecting the values and rights of citizens in Europe, but also in setting worldwide standards. In the same vein, it is vital that the EU seizes the opportunity to promote responsible AI innovations both in Europe and across the globe.

Relevant publications of the Rathenau Instituut

Rathenau Instituut (2020a). Missiegedreven innovatiebeleid vraagt samenwerking tussen lokaal en nationaal niveau. Website Rathenau Instituut, 5 February 2020. https://www.rathenau.nl/nl/vitale-kennisecosystemen/missiegedreven-innovatiebeleid-vraagt-samenwerking (Dutch).

Rathenau Instituut (2020b). Innovaties laten aansluiten bij de maatschappelijke praktijk: zo doe je dat. Website Rathenau Instituut, 21 January 2020. https://www.rathenau.nl/nl/vitale-kennisecosystemen/innovaties-beter-laten-aansluiten (Dutch). See also https://www.rathenau.nl/en/digital-society/innovating-societal-aims (English)

Rathenau Instituut (2020c). Voorbij lokaal enthousiasme – Lessen voor de opschaling van living labs. Website Rathenau Instituut, 5 February 2020. https://www.rathenau.nl/nl/vitale-kennisecosystemen/voorbij-lokaal-enthousiasme (Dutch)

Rathenau Instituut (2020d). The corona crisis calls for careful action and democratic debate. Website Rathenau Instituut, 23 April 2020. https://www.rathenau.nl/en/digital-society/corona-crisis-calls-careful-action-and-democratic-debate (English)

Rathenau Instituut (in preparation). Precautionary principle and innovation in Europe: AI for clinical decision making in health care. Report prepared for RECIPES, H2020.

Rathenau Instituut (2019a). This is how we put AI into practice based on European Values. Website Rathenau Instituut, 19 March 2019. https://www.rathenau.nl/en/digital-society/how-we-put-ai-practice-based-european-values (English)

Rathenau Instituut (2019b). Directed digitalisation – Working towards a digital transition focused on people and values – The Dutch approach. Website Rathenau Instituut, 20 November 2019. https://www.rathenau.nl/en/digitale-samenleving/directed-digitalisation (English)

Rathenau Instituut (2019c). Voorbereid op de praktijk – Anticiperen op de maatschappelijke inbedding van innovatie bij onderzoeks- en ontwikkelprogramma's. Website Rathenau Instituut, 9 April 2019. https://www.rathenau.nl/nl/vitale-kennisecosystemen/voorbereid-op-de-praktijk (Dutch)

Rathenau Instituut (2019d). Health at the centre – Responsible data sharing in the digital society. Website Rathenau Instituut, 6 June 2019. https://www.rathenau.nl/en/making-perfect-lives/health-centre (English)

Rathenau Instituut (2019e). Cyberspace without conflict – The search for de-escalation of the international information conflict. Website Rathenau Instituut, 27 August 2019. https://www.rathenau.nl/en/digital-society/cyberspace-without-conflict (English)

Rathenau Instituut (2018). Artificial Intelligence, what's New? AI What's new? Website Rathenau Instituut, 29 November 2018. https://www.rathenau.nl/en/digital-society/artificial-intelligence-whats-new (English)