The EU White Paper on Artificial Intelligence: the five requirements

April 15, 2020

Artificial intelligence (AI) remains one of the main features of most European countries' strategies, even during these times of the COVID-19 emergency. AI can in fact not only improve health care systems but also be a fundamental tool to analyze data to fight and prevent pandemics.

While there is little doubt about the benefits that that can be drawn, there are also increasing concerns about how to effectively address the risks associated with the usage of AI systems. Such concerns include, among others, data privacy risks – AI may easily be used to de-anonymize individuals' data etc… (see this previous bite on this point) – and also potential breaches of other fundamental rights, including freedom of expression, non-discrimination, human dignity, etc.

There has been a demand for a common approach to address such concerns, in order to give citizens and corporations enough trust in using (and investing in) AI systems, while also avoiding the market fragmentation that would limit the scale of development throughout Europe.

With this in mind, the European Commission recently published its White Paper on Artificial Intelligence, which is aligned with the key principles set out in the Guidelines on Trustworthy AI published by the EU High-Level Expert Group, namely human agency and oversight, technical robustness and safety, privacy and data governance, transparency and accountability, diversity, non-discrimination and fairness, societal and environmental wellbeing.

In addition to some improvements to the liability regime (such improvements are separately addressed in our TMT Bites), the EU Commission proposes to opt for a risk-based approach, to make proportional regulatory intervention in order to address mainly "high-risk" AI applications. Such high risks are identified where both the relevant sector (e.g. health care) **and** the intended use involve significant risks.

According to the EU Commission, AI regulations should be based on the following main requirements:

1. **Training data** – Datasets and their usage should meet the standards set out in the applicable EU safety rules, in addition to the existing provisions set out in the GDPR and in the Law Enforcement Directive. There should also be ad hoc AI training data provisions. For instance, AI systems should operate with sufficiently broad data sets to cover all scenarios needed to avoid dangerous situations, thus avoiding unnecessary risks. This includes also taking reasonable measures to avoid discrimination, e.g. where applicable, adequate gender and ethnical coverage.
2. **Record-keeping** – Adequate measures should be taken to avoid the so-called "black box effect". Accordingly, records should be kept on the data sets used to train and test the AI systems, as well as the main characteristics. There should also be clear documentation on the programming, training and processes used to build and validate the AI systems. In certain cases, the data themselves should also be kept, although this may entail additional storage costs.
3. **Information** – AI systems should be transparent. Information on the use of AI systems should be provided, also including information on the capabilities, limitations and expected level of accuracy. This also implies a proactive

approach, e.g. informing individuals when they are interacting with AI systems, while all information needs to be concise and understandable.

4. **Robustness** - Many AI technologies are unpredictable and difficult to control, even ex-post. There should be an ex-ante assessment of risks, as well as assessments to check that AI systems operate accurately during the whole of their life-cycle phase, with reproducible outcomes. Furthermore, AI systems should also adequately deal with errors, with processes to handle and correct them. Additional regulations should also be drawn up to ensure resilience against attacks and attempts to manipulate the data or the algorithms.
5. **Human oversight** – There should be adequate involvement by human beings, in addition to what is already established by the GDPR for automated decision making. Depending upon the circumstances, the human oversight should intervene prior to the output to be produced, or afterwards and/or throughout the whole learning and output process. This will depend upon the type of system and its usage: for instance, an automated driverless car should have a safety button or similar device in order to allow a human to take control under certain circumstances; it should also provide for an interruption of operations when certain sensors are not operating in a reliable way.

Other requirements may be set for other specific systems, including remote biometric identification, which allows identification at a distance and in a public space of individuals through a set of biometric identifiers (e.g. fingerprints, facial image, etc.) which are compared to other data stored in database(s). Additional requirements may be set, whatever sector is involved, in order to ensure that any such processing is justified, proportionate and subject to adequate safeguards.

The Commission further highlighted that, in order to make future regulations effective, there should be a level playing field, and accordingly any such requirement should be applied to all those that provide AI products or services in the EU, thus including non-EU companies.

The detailed implementation of the above requirements is yet to be determined, including the frameworks for testing and certification.

Do you agree with the above requirements? We would be interested in hearing your views.

This bite has been co-authored by Chiara Bocchi and Giangiacomo Olivi. For any clarification, contact our TMT Bites team!

Your Key Contacts

Giangiacomo Olivi
Partner, Milan
D +39 02 726 268 00
M +39 344 27 62 550
giangiacomo.olivi@dentons.com

Chiara Bocchi
Senior Associate, Milan
D +39 02 726 268 00
M +39 348 548 71 94
chiara.bocchi@dentons.com