Sella

Biella, August 5, 2021
Piazza Gaudenzio Sella 1

Feedback on the proposal of "*Regulation laying down harmonised rules on artificial intelligence (artificial intelligence act) and amending certain union legislative acts*" .

Introduction

Sella Group thanks for the opportunity to provide contributions about a notable regulatory project, which responds to ethical needs and respect for values and fundamental rights in the EU. The technological evolution should not affect health, safety and protection of fundamental rights, nor lead to discrimination (new or already existent forms).

On the other hand, Artificial Intelligence is one of the key factors that could foster social inclusion. As the Proposal of Regulation states: "*Artificial intelligence […] can contribute to a wide array of economic and societal benefits across the entire spectrum of industries and social activities. By improving prediction, optimising operations and resource allocation, and personalising digital solutions available for individuals and organisations, the use of artificial intelligence can provide key competitive advantages to companies and support socially and environmentally beneficial outcomes, for example in healthcare, farming, education and training, infrastructure management, energy, transport and logistics, public services, security, justice, resource and energy efficiency, and climate change mitigation and adaptation*" (see Recital 3).

It is essential to set the conditions to open the innovation process to the entire entrepreneurial system by fostering the research and experimentation not only by BigTechs, but also by start-ups and SMEs. In such context, the Regulation should be inspired by proportionality and its compliance costs should be sustainable even for the smallest innovators.

Moreover, the Regulation should not penalize AI systems themselves, but it should mitigate the AI-related risks which occur in some specific contexts/uses and that would not subsist - or would have a minimum impact - without the use of AI. Regulating AI if it is the mere executive mean of human choices would break the principle of technological neutrality and, at the same time, divert the attention from the effective causes of potential harmful and non-ethical behaviors.

For these reasons, Sella Group believes that a risk-based approach is a proper solution. The provision of minimum requirements for the medium-risk systems (those that determine a direct interaction of the final user with the output, without human intermediation – article 52) and the choice not to regulate low-risk systems, would limit the regulatory pressure to those uses of technology that could bring negative externalities to society.

At the same time, the residual character of the low-risk systems – opposite to the mandatory list of high-risk systems – theoretically allows to regulate only risky use cases.

However, the adoption of an approach actually proportioned to the risk is hindered; indeed:

- The definition of AI appears too wide: even systems which do not present "adaptivity" and "autonomy" fall in regulatory scope;
- Annex III provides a mandatory list of high-risk systems contradicted by an open definition approach, leading to a wide application of most restrictive requirements without a specific and precise identification of risky use cases.

Therefore, we suggest the following amendments and integrations to the Regulation.

Wide definition of AI

It is not our intention questioning about the definition of Artificial Intelligence. There have been plenty of debates about AI technical and technological features without getting to a unique and shared definition.

Considering literature and scientific debates, we believe that a wide definition of Artificial Intelligence is acceptable.

That said, it is important to highlight the risks that the Regulation wants to prevent and the role that AI plays as the source of these risks.

As the Regulation states, it "*pursues a number of overriding reasons of public interest, such as a high level of protection of health, safety and fundamental rights*" (see Whereas 1).

It is clear that the Regulation should intervene in AI systems that introduce or aggravate the encountered risks: the use of AI (so other technologies) should be specifically regulated only when risks and potential negative externalities arise from the technology itself and not from the type of the service where technology is used. In most cases, the output generated from the AI system is caused by the choices of enterprises/business organizations, sometimes outlined in their internal rules and approved by corporate bodies(and not from the learning ability and adaptivity of the system). In such cases, AI is just the technological mean used to apply predetermined rules.

In order to identify which AI systems should be regulated as a potential source of further discriminatory conducts - other than those determined by human choices -, it is important to distinguish between:

- Systems in which the programmer defines the algorithm and the machine executes it, without "adaptivity" and "autonomy" features;
- Systems in which the programmer describes the context, the objectives, the instruments (through a formal language or examples), asking the system the definition of the algorithm needed to solve the problem, and which are adaptable and autonomous.

In the first case, the cause of the potential harmful conducts is in the decisions that led to the programming of the algorithm. Regulating not-adaptable and not-autonomous AI systems means that technology itself is going to be regulated. This could determine, from one side, to lose sight of the sources of the risk by concentrating efforts and compliance resources on the mere executive mean,

and from the other side, it could lead to an unjustified violation of the technology neutrality principle. Moreover, higher compliance costs could end up hindering the capacity of start-ups and SMEs to invest in research and innovation, thus limiting their ability to propose innovative AI systems.

For these reasons, we believe that the Regulation should not refer to every AI system, but only to the ones in which the model independently defines the algorithm based on the context, the defined objective and the used data. Cases in which programmers, on the base of defined human decisions, create the model and the algorithm should be exempted because of their nature of mere executive technological means.

Otherwise, the paradoxical conclusion could be that the same conduct/decision would be considered as potentially harmful only if executed using AI systems.

Therefore, we propose the following reformulation of AI (art. 3 par. 1 n. 1):

"software with "adaptivity" and "autonomy" features that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate models and algorithms that are not defined when programmed, and generates outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with".

Moreover, the use of declarative techniques establishes a condition that is necessary but not sufficient to identify further risk profiles other than the ones generated by human choices in the provision of services. At the same time, it is necessary to distinguish the systems that could take decisions and generate outputs without human intervention from those used as a mere support. The proposed Regulation should be limited to the use of decision-making AI systems and exclude those systems which are a mere support to human decisions, provided they are explainable or interpretable.

Here is a brief summary of the suggested amendments to limit the objective application of the Regulation to systems and uses of technology that request a further mitigation of the risks other than the provision of the services without the use of the AI:

- The definition in article 3 par. 1 n. 1 should exclusively refer to "adaptable" and "autonomous" systems: *"software with "adaptivity" and "autonomy" features that is developed with one or more of the techniques and approaches listed in Annex I and can, for a given set of human-defined objectives, generate models and algorithms that are not defined when programmed, and generates outputs such as content, predictions, recommendations, or decisions influencing the environments they interact with".*

- In article 2, an exemption to systems that are used as a mere support to the final decisions should be provided.

These considerations only refer to high-risk AI systems, which are defined based on the context of use of the system and not referring to the produced effect. Otherwise, we understand the need to prohibit manipulative, exploitation and social control uses. In this case, as AI is the manipulation, exploitation and control instrument, we consider that the proposed definition is appropriate.

Following the different type of use, scale of risk and objective application, there could be four different regulations:

- Prohibited systems and uses: they could be based on a wide definition of AI (art. 3 par. 1 n. 1) and the list could remain the proposed one;
- High risk systems: limited to the "adaptable", "autonomous" and decision-making ones (also excluding the mere support systems);
- Medium-risk systems (art. 52): the provision of mere transparency obligations considering the direct interaction of the system to the final user;
- Low-risk systems: as proposed, no regulation should be provided.

List of high-risk systems

In the previous paragraph, we highlighted that the definition of AI used to identify the objective scope of the Regulation (referring to high-risk systems) should be limited to "autonomous" and "adaptable" systems.

About the list of high-risk systems in Annex III, the intention to provide a mandatory list of risky uses is contradicted by their descriptions, wide and open to extensive interpretations. On one hand, this situation undermines legal certainty and, on the other, it opens the high-risk class to some AI systems that do not bring harmful risks to society.

Based on the regulatory objectives, a precise description of the high-risk uses is expected (including specific cases – please see recital n. 32: "*As regards stand-alone AI systems, meaning high-risk AI systems other than those that are safety components of products, or which are themselves products, it is appropriate to classify them as high-risk if, in the light of their intended purpose, they pose a high risk of harm to the health and safety or the fundamental rights of persons, taking into account both the severity of the possible harm and its probability of occurrence and they are used in a number of specifically pre-defined areas specified in the Regulation*").

Despite high-risk class should be residual and referred not only to single "sectors", but also to the concrete probability that they could be harmful, it seems to be excessively extended and the requirements appear in some cases disproportionate to the concrete risks. A risk-based approach requires a more accurate descriptive specification thus limiting extensive interpretations or gold-plating.

In particular, an overly extensive approach was found in:

- "***AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score***" (Annex III, point 5, let. B)

  As noticed in the Regulation, "*AI systems used to evaluate the credit score or creditworthiness of natural persons should be classified as high-risk AI systems, since they determine those persons' access to financial resources or essential services such as housing, electricity, and telecommunication services. AI systems used for this purpose may lead to discrimination of persons or groups and perpetuate historical patterns of discrimination, for example based on racial or ethnic origins, disabilities, age, sexual orientation, or create new forms of discriminatory impacts*" (Recital 37).

The regulatory *ratio* – based on the needs to prevent discriminations in granting loans – is transposed in Annex III with a wider approach ("*AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score*"; Annex III, point 5 let. B) and ends up regulating in the same risk class even systems that are not used to grant credit, but also any other algorithm that can be used in the credit processes (e.g. systems which continuously verify and calculate the risk of the position). This is an obstacle to the ability of Banks and intermediaries (in particular the small and medium-sized ones) to improve their credit monitoring systems.

For these reasons, we suggest the following re-definition of the case in Annex III, point 5, let. B: "*AI systems intended to be used to evaluate the creditworthiness of natural persons or establish their credit score in the phase and with the aim of disbursement and resolution of the credit, with the exception of AI systems put into service by small scale providers for their own use*".

Further uses in the credit processes – that do not affect the access to the credit itself – could be included in a medium-risk class with simplified requirements.

- "***AI systems intended to be used for recruitment or selection of natural persons, notably for advertising vacancies, screening or filtering applications, evaluating candidates in the course of interviews or tests***" (Annex III, point 4, let. A) and "***AI intended to be used for making decisions on promotion and termination of work-related contractual relationships, for task allocation and for monitoring and evaluating performance and behavior of persons in such relationships***" (Annex III, point 5, let. B)

  In our opinion, systems that could be decisive for access to employment (e.g. "*screening or filtering applications, evaluating candidates in the course of interviews or tests*") should be distinguished by those about the management of the employment relationship ("*task allocation, monitoring and evaluating performance and behavior of persons in such relationships*"). In the second case, it is crucial to observe labor law, since it already provides the necessary rules, prohibitions, obligations and conditions (e.g. about demotions and transfers).

  Therefore we suggest that those uses which do not affect the access to employment should not be included in the high-risk systems but, if strictly necessary, in a medium-risk class with simplified requirements.

  Similarly, the generic indication of "*monitoring […] performance and behavior of persons in such relationships*" could lead some interpreters to include in this group some AI systems used by Banks and supervised entities inside their internal control framework, which purpose is different from employment decisions or task allocation. For instance, AI systems used in compliance processes to verify that staff is observing the prescribed obligations. Hence, we suggest the following rephrasing of point 4 let. B:"*AI intended to be used for making decisions on promotion and termination of work-related contractual relationships, for task allocation and for monitoring and evaluating performance and behavior of persons in such relationships, excluding AI systems used by regulated entities in their internal control framework*".

- "**AI systems intended to be used for the 'real-time' and 'post' remote biometric identification of natural persons**" (Annex III, point 1 let. A)

About this point, the needs of biometric data protection are already regulated by GDPR.

Further protection issues deriving from the use of AI technologies should refer to specific uses and not to the generic biometric identification. It is difficult to identify potential "*discriminatory effects*" (see Whereas 33) in the use of systems for the mere identification/authentication of already acquired customers (e.g. biometric identification to log in to banking apps).

We suggest specifying on point 1 let. A Annex III: "*AI systems intended to be used for the 'real-time' and 'post' remote biometric identification of natural persons, excluding those used for the mere identification/authentication of the customers*".

Requirements for high-risk systems, competitive disadvantages and transitory dispositions.

The biggest regulatory effort and most of the compliance costs are focused on high-risk systems, which are excessively extended and include cases in which the proposed regulatory provisions appear disproportionate.

Proposed regulation also presents organizational and governance restraints that require a stronger regulatory effort inside the firms and the outlining of a control framework that start-ups and SMEs could not arrange. This setting ends up favoring big-sized enterprises or entities, which already have a complex system of internal regulation and controls that are difficult for smaller enterprises to arrange, since they already struggle to fundraise and invest to compete with the bigger technological players.

Moreover, some provided requirements lead to benefit BigTechs, which can use their privileged position on data to fulfill the requirement on art. 10 about completeness of data sets.

Therefore, we believe that this requirement – if maintained – could be applied without affecting the level playing field only when easily accessible data sets and an adequate regulatory framework about open-finance will be set up.

Similarly, the obligation of compliance certification of AI systems (above all when an external certification is required) involves high costs for those companies in early stages.

In this context, disadvantages are not totally compensated by the provisions in art. 55.

For these reasons, we believe that the suggested amendments about a more detailed definition of AI subject to Regulation, a more accurate detailing of high-risk uses and, if strictly necessary, the provision of a medium-risk class with simplified requirements, are appropriate.

Similarly, in order to foster the competition of small and medium-sized enterprises in this market development phase, we suggest to provide a transitory framework in which limiting requirements to a voluntary and experimental context, also encouraging the adherence to the proposed regulatory scheme. The definitive legislation could be adopted with some amendments/different identification of the high-risk uses, based on the acquired experience. Moreover, small-sized enterprises could invest in research and programming, entering the market and avoiding concentration risks which benefit larger corporations and can affect level playing field and the whole ecosystem.

Sella

Regulatory sandbox

New Technologies – in particular AI – lead to new service models, which present regulatory challenges. In this context, regulatory sandbox fosters the dialogue between market, Regulators and Supervisory Authorities: market can quickly adopt new business models in a context of properly-identified risk, while Regulators can have prior knowledge of new risks in order to evaluate timely and proportionate regulatory interventions.

The identification of new regulatory, interpretative and enforcement solutions through regulatory sandboxes supports innovation and the maintenance of a regulatory level playing field.

For these reasons, we believe that the provision of European regulatory sandbox mechanisms is crucial while the only provision of national ones could hinder the level playing field; indeed:

- The mechanisms and conditions of access to regulatory sandboxes would be regulated and applied differently in the EU;
- The powers of experimentation of national Authorities would be limited – because of the application of supranational legislation – and applied in different ways in Member States;
- Undertakings in Member States that have not adopted regulatory sandboxes yet would have competitive disadvantages.