Ref. Ares(2021)4995045 - 06/08/2021

Comments on the EU project of regulation on AI.

Date: 05 August 2021
Author: Thibault Helleputte

Foreword.

This note provides feedback on the project of regulation on AI, currently being built by EU.

For more information on this project, see here:

https://ec.europa.eu/commission/presscorner/detail/en/ip_21_1682

as well as the text itself here :

https://eur-lex.europa.eu/legal-content/EN/TXT/?qid=1623335154975&uri=CELEX%3A52021PC0206

The present note has been written by Thibault Helleputte, computing sciences engineer, PhD in Machine Learning, MBA, with 15+ years of application of data sciences to healthcare, MIT Innovator under 35, CEO at DNAlytics, board member at Biowin the health cluster of Wallonia, Belgium. The note is a personnal exercice but has been built after discussions with and consultation of representatives from companies and institutions in Belgium, EU, and in the USA, small, medium and multinational companies, in different economic areas: health care, metallurgy, finance, automotive, ... (>20 institutions).

The following persons also wish to formally support this text (for themselves or for their company) :

- Jérôme Callut (Finance, Switzerland)
- Sébastien Deletaille (Telecom and Healthcare, Belgium)
- David Frenay (Healthcare, Belgium)
- Dirk Loeckx (Healthcare, Belgium)
- Laurent Paul / 3D-Side (Healthcare, Belgium)
- Thomas Quinet / PSI Metals GmbH (Metallurgy, Germany)
- Olivier Staquet / GabiSmartCare (Healthcare, Belgium)
- Tanguy Swinnen / SynNeo (IT Consulting, Belgium)
- Michel Verleysen / Dean of Engineering school at UCLouvain (Belgium)
- [... this list is regularly updated with new names ...]

Most interactions we could have on the proposal with political representatives, or some federations at the Belgian level, have been constrained to the possibility of expressing small adaptations on the current proposal. However, despite the fact that all consulted stakeholders

Copyright @T.Helleputte – 2021

are not against regulation in general, we feel the current work on AI regulation as a whole is unsuited to the rest of the EU regulatory landscape, for several reasons.

Consequently, the text that follows makes comments at two different levels. The first level is about why the initiative seems quite wrong as a whole, and the current work should be stopped and completely reworked. The second level is about what specific points within the proposal should be modified, should the project remain on the table.

As the profesionnal experience of the author is in healthcare, most examples are in this field. Nevertheless, we received feedback from our network in other industries pointing at the fact that these examples would find their counterpart in other industries as well.

The last section of the note proposes how to move forward with regulating on AI, while keeping more consistency with existing legislation, guaranteeing more practicability for companies while being at the same time more future-proof.

Copyright @T.Helleputte – 2021

I. A misfit regulation project

There are two major issues why a regulation on AI is questionable in the first place: One, there is no such thing as "AI", which makes the regulation's scope very difficult to define (as the legislators have probably discovered by themselves, trying to define the concept), and the regulation's actual goal is as a consequence uneasy to grasp; Two, EU has built over the years a set of "vertical" regulations (per applicative domain), this regulation project is orthogonal to that mindset. This "horizontal" regulation will cause double administrative and financial burden, as well as conflicting scenarios.

Let's discuss those two points.

I.1 Global issue #1. The very principle of regulating "AI" seems a fragile endeavor. There is no such thing as a strict definition of AI, which is rather a collection of techniques of various natures. Most people with even a small education in the field know that. We'll get back to this later on, namely discussing the definition that is used for AI in the text. AI is more and more a kind of myth in citizen's mind. Everybody uses it, no one knows what it is. Thus, using this term in the very title of the foreseen regulation hints at the fact that the fundamental societal objective is likely ill-defined, in terms of concrete impact for EU citizens and companies. A regulation which true objective cannot be understood even by profesionnals in the field will have a hard time to get their adoption, not to mention the one of the regular EU citizen.

During discussions with political representatives and in some publicly available documents, we see that one of the main goals of the regulation is to ensure that AI systems placed on the Union market and used are safe and respect existing law on fundamental rights and Union values. But how could it be the case that anything put on the EU market, involving AI or not, would be allowed not to respect other existing laws? The absence of a regulation on AI does not involve that a product or service put on the market would be allowed to infringe, for example, on GDPR, anti-discrimination laws, or any other law. This objective can basically be rephrased as: we need a law to ensure AI-products comply to the law. This is seems highly questionable (because useless, redundant).

We also have heard and read that regulating the use of AI applications will also lead to a competitive advantage in the global AI market because it will foster innovation and increase users' confidence and promote its implementation and use by citizens and businesses. Because of the poor definition of AI, we believe it might just do the opposite: make people believe there is some kind of (black) magic in many applications that are totally common or clearly not based on actual AI (as perceived by profesionnals of the field), and draw suspicion on many products and services, just because this regulation will encompass it. As for competitiveness, EU is already far behind USA and China in terms of products and services involving AI despite it does not lack experts in the matter. This regulation will further distanciate EU economy with respect to the integration of these techniques.

Copyright @T.Helleputte – 2021

Going through the text of the regulation, it may be that what is actually targeted by this regulation would be clearer if it were coined as "regulation on decision making". That would clearly state what kind of application we want to regulate. But in such a case, then again, it should regulate **all** decision making processes, even human-based decision making, not restricting itself to mathematical/computer-based approaches. As far as the risk of biased decisions or the fear of "blackbox" decisions is considered, people, too, make decisions based on partial, approximately correct information, and are rarely forced to express their complete reasoning. They must nevertheless comply to all laws and regulations in place: see medical diagnosis example below, but also anti-bribery or anti-discrimination laws and others that already tackle a good portion of how decision should and should not be made.

To broaden the perspective, with more and more systems passing the Turing Test[1], the enforcement of the currently foreseen regulation will not be even feasible at some point.

I.2 Global issue #2. EU has already put in place many regulations for services and products put on its internal market, and we are definitely in favor of that. But these texts regulate applications (and rightly so), not methods. AI is a set of methods. It is a mean, not an objective. Adding regulations on means/methods to regulations on objectives/applications will inevitably represent a double regulatory burden on entities willing to put new products or services on the EU market. It will on top of that introduce doubts on which regulation has priority on the other.

Examples:

- The regulation on medical devices and in vitro diagnostics (MDR and IVDR) state the conditions to which these categories of products can enter the EU market. These regulations specify in detail how risk management must be covered and how performances must be assessed and demonstrated, before putting new products/services on the market. But MDR and IVDR do not say anything (and that is their strength) on technologies, or methods involved in these products/services (but for clarity, they mention that, yes, software elements also fall into their scope). So, if one wants to put a new diagnostic tool on the market, and that tool is based, among other things, on a model derived from machine learning, he must comply with IVDR. And thus not only having conducted a risk analysis, but also having performed successfully a clinical trial/study to validate performances, *regardless* of the internal details and technologies of the product. When a patient benefits from this new diagnostic tool in EU, he can rely on the health product strict validation and does not have to worry about the methods involved. The general principle is the same for drugs and patients rarely ask what mechanism of action is involved by this or that drug. And if they do, they would realize that for many everyday drugs (such as many pain killers), the mechanism of action is even not known. The efficacy and the safety are what matter. Hence, the objective that is (probably) pursued by this new regulationon IA is completely redundant with IVDR and MDR, for example, for that applicative filed.

1A system passing the turing test is an artificial system (a computer, a machine, an artifact,...) for which an actual human cannot tell from the outside if it is human or not.

Copyright @T.Helleputte – 2021

- The automotive industry is currently consulted on how automated driving should be regulated. The work is currently oriented, just like IVDR and MDR in the medical field, on what guarantees of efficacy, safety, should be required by this field. But it is definitely not entering into a discussion on **how** the sector should make this automated-driving work (with AI or something else), because there is no point.
- The financial field is highly regulated with respect to public stock market transactions (among others). These regulations are very strict, but do not at all forbid (nor position themselves on) automated, computer-base ("AI-based") transcactions.

Otherwise said, this endeavor of regulating a set of technologies is not suited to the EU regulatory landscape. As an analogy, just replace "AI" by "electricity". There is no such thing as a need for a global regulation on electricity. Rather, there are regulations on how toys (that might involve electricity), cars (that might be electric), medical devices (that might be electrically powered), etc, can or cannot be put on the market. EU has not so far adopted a position of regulating a set of means or technologies rather than applications in a field-specific manner. Are there other comparable regulations on global sets of techniques? Electricity *per se*? Chemistry *per se*? Biology *per se*?

Copyright @T.Helleputte – 2021

II. Specific comments.

As explained in the foreword and in the previous section, the complete initiative would benefit from a total reset. Nevertheless, should the project continue on the current basis, several specific elements would benefit from modifications. Below are these suggestions of modification. This list of observations is clearly non exhaustive.

II.1 Title I : general provisions.

Article 2, point (3) and (4) provides exemption of compliance for military and govermental (in short) applications. The said aim of the regulation being to make citizens more confident in AI applications, these exceptions are difficult ot accept, even more to understand for the layman.

Definition of AI in Article 3 and Annex I. The definition that is used has two weaknesses.

First, the definition is clearly too broad, and that is specifically related to point (c) of Annex I. In fact it comes to a regulation on mathematics, period. Think about mere Excell spreadsheets. Or the basic Google or Bing web search engines, which use search and optimization methods. Or much more basically: traffic lights at crossroads (which are using optimization methods for traffic fluidity). Or solar pannels which need to optimize the current that goes through them to deliver an optimal energy. And what if some decisions in government, or company boards, require computations with a pocket calculator, computing a simple mean for example (which is clearly a statistical approach)… and so on and so forth.

Second, the definition is stated in a positive elicitation of a limited set of methods, which will be outdated the day of its release. We agree, and we would go further: it is doomed to be outdated the very day it is released. The legislation timeframes are huge with respect to the speed of innovation. That logic is exactly what undermined for example the former directive on in vitro diagnostic, which is one of the reasons the newer IVDR has been written: it listed a set of biology techniques that were known at the time for being involvd in in vitro diagnostic tools **by then**. As a result, most of the in vitro diagnostic tools getting on the market did not fall under that directive.

This definition also raises a question: will electronic hardware-encoded AI be also encompassed? A software can be, if needed, translated into a purely hardware system, without the need for software. Would such a system fall under the scope of the regulation? To my understanding, it would not, though it would accomplish exactly the same outcome / conduct the same tasks.

II.2 Title II: prohibited practices.

Article 5(1)(a), on the prohibition of the placing on the market, putting into service or use of an AI system that deploys subliminal techniques beyond a person's consciousness in order to

Copyright @T.Helleputte – 2021

materially distort a person's behaviour in a manner that causes or is likely to cause that person or another person physical or psychological harm (Article 5, (1)(a)). Again, regardless of the set of techniques used, from the field of AI or not, why would a system, a service or a product that causes harm (in short) not already fall under laws against physical damage or psychological harasment, for example? What is the need for an AI-specific regulation on that matter? This article might be cancelled.

The same goes for (1)(c) on the prohibition of the placing on the market, putting into service or use of AI systems by public authorities or on their behalf for the evaluation or classification of the trustworthiness of natural persons over a certain period of time based on their social behaviour or known or predicted personal or personality characteristics. Regardless of the set of techniques used, from the field of AI or not, why would a system, a service or a product that clearly discriminate between people not already fall under laws against discrimination, for example? What is the need for an AI-specific regulation on that matter? This article might be cancelled.

Same holds for (1)(d).

II.3 Title III on High-risk AI systems.

The **requirements** for high-risk AI systems (Chapter 2 of Title III), are sometimes vague. This limitation is actually inherent to the fact that the project of regulation regulates a set of methods, not applications. Thus, being specific then on applications classification is not feasible, because the whole text is decorrelated of any application.

The **Standardisation** and **Technical Documentation** that will come with the regulation will enter into conflict with standardization specified by the many regulations on applications that are already in place according to the more "vertical" mindset which prevails. And on top of incompatibility, the time and money required to satisfy yet another regulation with respect to standardization and documentation will bear upon competitivity of EU companies (excepted consultants in regulatory matters, obviously). We have difficulties imagining complete fields of products/services that would not already be covered by a regulation already stating standardization and technical documentation prescriptions. So it will be a double burden, administratively, time-wise and financialy.

II.4 Title IV on transparency of certain AI systems.

This title covers the obligation of providers of AI systems that interact with natural persons, to ensure that these systems are designed and developed in such a way that natural persons are **informed transparently** about the fact that they are interacting with an AI system. Given the scope of the AI definition, is it possible to imagine a single piece of software on the market that will not fall under the scope of the regulation? Imagine the day of any office-worker, sitting in front of a computer all day-long, having to acknowledge the fact that he/she is using a piece of

Copyright @T.Helleputte – 2021

software falling under the scope of this regulation. The same applies to smartphones. Such information will have to be provided for a lot of everyday life's objects. A car, a traffic light, an electronic watch, a television, maybe some modern kitchen devices such as an oven with self-regulated temperature, etc etc. But this merely translate the ill-defined purpose of the regulation in the first place: what do we intend to achieve?

This title also stresses the need for AI systems to respect fundamental rights, but seems namely to link to that question a requirement of explainability/traceability of AI systems. Enforcing a complete explainability of decision making of these systems may be in some cases unethical, and against some fundamental rights. Here is an illustration of that statement. A system that is able to beat human at chess is quite simple, and we can follow its decision process. To beat human at the game of GO, on the other hand, other techniques are needed where the reasoning of the machine is impossible to grasp for a human being. This is important: explainability may in some case imply poorer performances. Now, if this is applied to medical diagnosis or say automated driving. What is preferable? A system which decisions are completely traceable but that will achieve poorer medical performances (which has consequences for the patient, but also globally for our social security systems in EU) and comit more traffic accidents? Or more powerfull results at the price of more complexity and less transparency? We do not understand how a deep learning system beats human at the game of GO. But we know in the end, it wins. So precisely for ethical and fundamental rights concerns, several dispositions stating how AI should be implemented should be withdrawn.

Copyright @T.Helleputte – 2021

III. Recommendation to move forward

For all the reasons exposed above, we recommend to withdraw this project of regulation on AI, and rather proceed to something that would be at the same time more effective, easier and future-proof.

We argue that AI as such should not be regulated in an independent ("horizontal") manner, but rather, EU should make sure that laws and regulations on what actually matters, i.e. applications, are robust enough to encompass those applications that embed data-driven and/or "AI" components. For the EU citizen, a "safe AI" does not mean much. To trust cars, toys, medical devices, to be ensured that decisions are made in a fair way, without discrimination, etc, that is what likely matters for the citizens.

The good news is that this exercice is easier, and even more: it has already been made in practice. Let us take the example of EU regulation 2017/746 on In Vitro Diagnostics medical devices (https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32017R0746&qid=1627907459494). In the introduction, we find paragraph (17) :

It is necessary to clarify that software in its own right, when specifically intended by the manufacturer to be used for one or more of the medical purposes set out in the definition of an in vitro diagnostic medical device, qualifies as an in vitro diagnostic medical device, while software for general purposes, even when used in a healthcare setting, or software intended for well-being purposes is not an in vitro diagnostic medical device. The qualification of software, either as a device or an accessory, is independent of the software's location or the type of interconnection between the software and a device.

That makes the context and the spirit clear: an application for medical diagnosis integrating software components will not elude the regulation, and a solution for in vitro diagnostics based solely on software will also be considered as a medical device for diagnostics. This is later on formalized in, Article 2 (2), with the definiton of "in vitro diagnostic medical device":

'in vitro diagnostic medical device' means any medical device which is a reagent, reagent product, calibrator, control material, kit, instrument, apparatus, piece of equipment, software or system, whether used alone or in combination, intended by the manufacturer to be used in vitro for the examination of specimens, including blood and tissue donations, derived from the human body, solely or principally for the purpose of providing information on one or more of the following:

(a) concerning a physiological or pathological process or state;

(b) concerning congenital physical or mental impairments;

(c) concerning the predisposition to a medical condition or a disease;

Copyright @T.Helleputte – 2021

(d) to determine the safety and compatibility with potential recipients;

(e) to predict treatment response or reactions;

(f) to define or monitoring therapeutic measures.

Later on, in Chapter II of that regulation, EU regulates more specifically the requirements regarding performance, design and manufacture. Here, all articles apply also to what would be called "AI" systems, given the previous definition. But some extra specificities are added, in point 16:

16. Electronic programmable systems – devices that incorporate electronic programmable systems and software that are devices in themselves

16.1. Devices that incorporate electronic programmable systems, including software, or software that are devices in themselves, shall be designed to ensure repeatability, reliability and performance in line with their intended use. In the event of a single fault condition, appropriate means shall be adopted to eliminate or reduce as far as possible consequent risks or impairment of performance.

16.2. For devices that incorporate software or for software that are devices in themselves, the software shall be developed and manufactured in accordance with the state of the art taking into account the principles of development life cycle, risk management, including information security, verification and validation.

16.3. Software referred to in this Section that is intended to be used in combination with mobile computing platforms shall be designed and manufactured taking into account the specific features of the mobile platform (e.g. size and contrast ratio of the screen) and the external factors related to their use (varying environment as regards level of light or noise).

16.4. Manufacturers shall set out minimum requirements concerning hardware, IT networks characteristics and IT security measures, including protection against unauthorised access, necessary to run the software as intended.

We see here that the article regulated important aspects also pursued by the proposal of regulation on AI: security, availability, performance qualification, … It is not the aim of this note to dive more into IVD regulation, but the interested reader can verify that many other aspects of interest in the project of regulation on AI are properly dealt within that regulation, without once entering into the details of the technologies (which makes it stronger, more future-proof).

Proceeding this way is also a guarantee that any manufacturer of a specific kind of product, service, application will manage properly with the necessary requirements. It will also guarantee there is no conflict between an application-specific regulation and a technology-specific regulation.

Copyright @T.Helleputte – 2021

EU may thus review existing regulations on most important fields, and consider them as to determine if, as for now, there might be gaps or "holes" in them consisting in the fact that AI-based applications would elude important aspects that actually matter. If such gaps are identified, proper remediation should be taken.

Copyright @T.Helleputte – 2021