CEA welcomes the new approach of the European Commission in the "White Paper on Artificial Intelligence", and stresses the importance of secure, safe, explainable, reliable, and ethical AI systems. It is no longer necessary to prove that AI has become essential for innovation in industrial and societal applications. However, it should not be overlooked that AI is also a very important tool for advancing life and earth scientific fields.

Ecosystem of excellence

CEA would like to underline the important role of RTOs in the partnership between the private and the public sector to achieve an 'AI ecosystem of excellence'. RTOs are the important links in the entire value chain from research and innovation to the adoption of solutions based on AI, including by SMEs.

CEA ponders the action "Focusing the efforts of the R&I community" as very vital to build an ecosystem of excellence that can support the development and uptake of AI across the EU economy. Note also, in the current context, companies are making less profit and then are less investing in R&D. Therefore, the Commission financial support of world reference testing centres in Europe is even more than very important.

Furthermore, it is important to note the ecosystem of excellence requires the development of cooperation with regulatory and standardization bodies. Indeed regulation is necessary in many sectors where AI can be used such as occupational safety and health, vehicle manufacturing, transport traffic, finance or cyber-security regulations.

The promotion of the uptake of AI by business and the public sector is of course important. Nevertheless, publicity has been already made for the promotion of AI. It is more important to focus the efforts and therefore the budget towards the design, development, and deployment of AI solutions.

In addition, it is important to point out that building up the European data space is imperative for global issues such as health, climate and safety. The European Data Space is also of great importance for industrial data, even if manufacturers do not share the data that helps them stay ahead in highly competitive environments.

The lighthouse research centre could be very important. However, its concept needs to be clarified in the white paper. Actually, AI is a tool and must have a strong link with the fields of use, as well as with the data producers. We have to be careful not to create a new off-ground structure. The lighthouse research centre needs to be a hub relying on RTOs as RTOs are the point of articulation between all stakeholders. It may also rely on a PPP that is an association representative of most stakeholders. In the same way, the network of existing AI research excellence centres should have the obligation to communicate with RTOs and networks of industrial partners.

Regarding SMEs, it is crucial to let them have access to AI technologies. RTOs are actually the catalysts for the support partnerships between SMEs, larger enterprises and academia around AI projects. It is obviously a priority to support SMEs, particularly with ERDF-type tools or hubs. Though, raising SME's awareness about AI benefits is today less important, as AI became the buzz word everywhere, and promotion has already been done.

Ecosystem of trust

Building an AI ecosystem of trust is without any doubt a policy objective in itself. It provides citizens with confidence to take up AI applications. It also gives companies and public organisations the legal conviction to innovate using AI.

It is worth noting AI must be secure, robust, reliable, and accurate in addition to safe. These features should not be confused or be used for the same semantics. In particular, one should not mix up safety with explicability. The explanation of taken actions is important if confidence can be established in the explanation provided.

Note also all digital techniques may lead to discriminatory outcomes, which rather derive from the use cases. Digital discriminatory issues are not specific to AI.

Regarding compensation, the real subject is the responsibility and accountability. The AI disciplinary itself is not responsible. The manufacturers or operators/users are the only accountable. They are also responsible of the learning process. It becomes very unfair to conclude, after an accident, that it is the fault of AI and nobody is accountable. Therefore, it is necessary to build, as it has already been done for personal data, a legal sphere on this subject that defines the responsibilities of the stakeholders: manufacturers, operators, or users.

EU legislation is necessary to address such concerns related to the AI use. Current legislation may have some gaps to be filled. Protecting people and assets needs to be the main priority while filling these gaps. An entirely new legislation would not be more protective

Regarding the safety of AI system, “high-risk” AI applications should not be binary; there must be a gradation of risks as in all sectors with a safety or security aspect. Different rules adapted to the application risk levels (like those in the IEC 61508 standard) are needed. The applications and the systems that are highly concerned with safety and security are those like cyber physical systems with real-time applications, autonomous systems, decision systems (with and without human verification) because human beings tend to over-trust the machine suggestions. Safety and accuracy are also very important in sectors like health or justice sectors.

The AI security relies on both the data security and the software security. Suppliers of AI systems must deal with these two aspects in parallel.

Particular attention should be paid to the concept of labelling that may lead to dangerous issues. Labelling must always be carried out against a charter verifiable by a third party. Self-labelling involves too much risk with regard to its verification, in particular vis-à-vis non-European players subject to other legal and commercial rules.

Although auto validation of a system is a very ambitious feature, it corresponds to a real need. A system needs to be able to validate itself during its lifetime, because AI systems are subject to frequent evolution of dates, or use context. For example, in automatic driving, the images captured from the environment 10 years ago are not the same as those of today, so the system has to adapt. The auto validation feature can be performed by installing embedded software in the system that maintains the system's validation guarantee despite the evolution of the environment or the system itself. Such validation software must be third-party certified.

Finally yet importantly, the white paper does not mention at all the “System” vision and the need/obligation to define system engineering processes for AI-based systems (and not only for machine learning). The use of system engineering processes should be made mandatory. Needless to say, the “System” must treat the “Green” and “Sustainability” concerns.

Safety and liability implications of AI, IoT and robotics

New important cyber risks stemming from the use of AI may occur if application owners do not control and protect the data used. There is a danger if a cybercriminal finds data used for learning or even data not used during learning because it constitutes weaknesses or even safety and security breaches. Thus attacking the data is like attacking the AI. AI application providers should have the obligation to protect such information and are accountable of the application safety and security.

Legislation should consider an AI application like any software application. We can't blame the AI. Following the use case, the application providers, issuers or operators are responsible for the malfunctioning of the applications.

Last but not least, changes in the system may occur several times a month, thus the safety requirements will be too light without a continuous risk assessment. There is a real need for more R&D in the risk area in order to conceive and create continuous risk assessment processes.