White Paper on Artificial Intelligence: a European approach to excellence and trust
Additional Upload Document attached to Siemens' online response

General remarks / concerns / recommendations:

- The paper is comprehensive and addresses many needed aspects of AI such as R&D landscape issues, skills, regulations.

- The matter of trade agreements and undue / unilateral nation's behavior is touched with respect to the data aspect (where the EU has a strong foot).
  However, the trade and specifically tech war between the US and China escalates these days. USA's intends to mark 14 emerging technologies as "dual use" and therefore under export control.
  - It is appreciated that the EU's AI initiative itself is a response to this situation.
  - The EU bodies seem to lack a structure to efficiently discuss and address international Technology Transfer issues for emerging technologies.

  Consequently, the paper (and approach) can do better in **Embedding the Access to AI technology into International Trade- and Export Control Agreements for the interest of** EU member states.

- **Definition & Scope:** If the European Commission is considering new or adapted legislation/rules/guidelines for AI the question arises what exactly is an AI system: what kind of products, systems and services will be considered by this initiative?
  An exact definition of AI would be needed.
  Furthermore many of the issues raised in the document are not specific for AI. Coverage of standalone SW, changing functionality of products due to SW updates, uncertainty on the allocation of responsibilities and changes to the concept of safety that shall be addressed by the legislative framework are not limited to AI, but apply to SW/ICT in general. Specific requirements on robustness and accuracy, keeping of records and data, information provisioning and human oversight should apply independent of the underlying technologies.
  The Commission should therefore carefully consider how the scope of the future legislative and regulatory framework is defined.
  We **advise to focus on today's narrow AI applications and not speculate about General AI** (like e.g. the paperclip maximizer example: https://wiki.lesswrong.com/wiki/Paperclip_maximizer) or Super AI and try to cover them. Regulation should also focus on current and emerging AI applications and not try to consider speculative General AI. Even with constant re-training during use, an AI system will not automatically have new functionality; to use an example, an autonomous vehicle cannot be re-training during use to improve from autonomy level 4 to 5

- **(Re-)Certification issue:** Most AI solutions are trained during development and the trained model is not modified during use, except for dedicated software updates. This is covered by traditional development approaches where verification, validation and if needed certification is done before new or updated software is deployed. However, some AI applications might be re-trained during use with an immediate update of the model. How verification, validation and certification is performed in that case is an open issue which requires further research activities.
  Autonomously and constantly learning AI today is very rare, but we suppose in the future we want to take care of this; then the certification has to take a holistic approach and look at the whole system and not necessarily only at the current model, what provides only a

snapshot and changes constantly with the update frequency.
To have a valid certification, the whole system has to be taken into account, e.g. how the process of model update works, how is it guaranteed that the necessary requirements are met (minimal accuracy, data quality, etc.) and what happens if these guarantees are not met (e.g. stop the system). Then the model is embedded in its certified environment and does not need to be re-certified each time it is updated.
To keep it practicable we must **avoid periodic re-certifications due to model updates (independent of how frequently)**. In our opinion this is just a matter of how the model is embedded in the whole system and what "watchdogs" (surveillance bodies) are available. For example, does this require constant re-certification (at which intervals, how is it performed even in off-line conditions)? Should the system itself ensure that regulatory requirements are always full filled?
In the future, AI will also be trained more and more on the devices itself as the computational power constantly increases. This is a trend reflected in technologies as AI-on-the-edge or federated learning. But this does not mean that the models cannot be evaluated before deployment
Is the EC planning to form an independent 3rd party AI certification body, like e.g. ENISA for cybersecurity?

- **High Risk Sectors / Applications:** The Commission proposes to identify high risks sectors and within these sectors high risk applications. While some sectors may have higher numbers of high-risk applications, it cannot be excluded that some high-risk applications also appear in other sectors, especially as beside safety also privacy and fairness / bias risks are covered by the frameworks. In this context it is also important to clearly define the criteria for a high-risk rating of application and who will do the rating.
  From Siemens' perspective AI Applications used in critical infrastructures (transport, energy, water supply, electricity grids, hospitals,...) and AI in Industrial applications where personal privacy or where life & limbs are at stake must be critically evaluated (case-by-case approach). The risk level must be objectively determined by the criticality of the application itself.

Specific additional and complete remarks (due to 500 Character restriction)

Section 1 – Revising the Coordinated plan on AI

Question: Are there other areas that that should be considered?

Siemens Answer (complete):

Activities to build European data spaces should take sector specific industrial use cases as a starting point and should also be well aligned with national activities.

Data ecosystems, based on contracts between companies of how to access and use data, will be a vital part of a future European Industrial and Services Ecosystem, enabling a European Cloud Service and Data Economy and supporting the mass adoption of AI.

At the same time, it might be advisable to systematically drive best practice exchange, e.g. by formulating use case types with reference examples in the industrial domain that may serve as blueprint for a future productization of AI solutions for industrial data. This should be driven by the industry, but potentially with project support from the EU.

A fail-fast innovation culture can foster the rapid generation of PoCs for all types of use cases and explore the feasibility and profitability of emerging use cases at an early stage. This will allow member

states to further tailor policy measures according to most promising application fields and thereby unlock a competitive advantage for European AI.

Section 1 – Question: Are there any other actions to strengthen the research and innovation community that should be given a priority?

Siemens Answer (complete):

Yes: Europe needs to make focused investments in industrial AI, based on a combination of technologies such as machine learning, semantics, NLP (Natural Language Processing), vision, combined with domain know-how, in domains where Europe plays a leading role. Building on the idea of "Lighthouse Research Centre**s**", industry-led AI R&I super clusters should be established that can generate global leading innovations, enabling European AI talents and stakeholders to bundle forces for fast innovation and avoid dispersed efforts. A strong collaboration between industry and academia will allow for the development of practicable AI solutions within those superclusters, so that supply and demand will be intertwined.

Section 1 – Question: Are there any other tasks that you consider important for specialized Digital Innovations Hubs?

Siemens Answer (complete):

The specialized DIHs where SME's are enabled to test their use cases, should focus on domains where Europe plays a leading role: combining hardware solutions, automation, semantics, edge computing, (data) analytics, explainable and data scarce AI, with the goal to take the efficiency of industrial infrastructures (factories, power, transportation, etc.) to the next level, such as through the Digital Europe Funding Program.

Industrial companies with expertise in automation technology can play a crucial role to support SMEs on the applicability of such technologies in their specific verticals

Section 2 – Question: Do you have any other concerns about AI that are not mentioned above?

Siemens Answer (complete):

"not always accurate" is not inherent to AI but depends strongly on the application where AI is used, on the used data set and on the input order of data. Explainability will always be difficult ("black box") for complex, deep learning algorithms, but it should be possible to at least have AI algorithms advertise their level of confidence in their recommendations. A major risk related to AI applications are cyber attacks; hence the need to ensure the cybersecurity of AI applications (depending on their criticality).

Technical approaches like "trustworthy AI", "safe AI", "robust AI", "federated learning" and especially "AI-on-the-edge" could contribute to solve some of the most urgent trade-offs between the potential benefits of AI solutions for the collective good and the inherent intrusion in personal privacy rights by acting as trustee.

In addition, we recommend considering the potential misuse of initially well-intended research areas like GANs (videos / pictures / text) to manipulate content and undermine trust (e.g. by generating "fake news"), while keeping in mind that in such cases AI technologies are amplifiers, but clearly not the root cause.

Section 2 -Question: Do you think that the concerns expressed above can be addressed by applicable EU legislation? If not, do you think that there should be specific new rules for AI systems?

Siemens Answer: OTHER

Most industrial applications (80-90%) do not need new regulation, because they are already sufficiently covered by existing ones (GDPR, Mach. Directive). No “one-size-fits all” approach for AI regulation is possible: Europe must address each vertical market separately.

Use sandboxing to test new concepts (e.g. autonomous driving) within a delimited, regulation-flexible and innovation-friendly space.

Free flow of data and support for contract-based solutions for sharing B2B (private) data as a general principle.

A separation between sectors and even specific application fields might be advisable in the future, in order to tailor potential regulation to higher risk applications. However, the administrative burden for companies needs to be minimized in order to avoid inhibiting innovation capacity in Europe

Section 2 - Question: Do you have any further suggestion on the assessment of compliance?

Siemens Answer (Complete):

Extensive exchanges with all stakeholders across the AI chain will be necessary to discuss further the appropriate compliance and enforcement mechanisms. In our view, a combination of ex-ante assessment & ex-post surveillance and enforcement mechanisms would be purposeful, in particular in light of the positive experiences with the EU product safety legislation following the so-called “New Legislative Framework / NLF” (Decision 768/2008/EU, Regulation 765/2008/EU), so-called “CE-Directives”.

For the purpose of AI, reference should be made to the “internal production control” procedure (Module A) under the NLF, which provides for the conformity (ex-ante) assessment under the sole responsibility of the manufacturer (self-declaration), without the mandatory involvement of a third party body. This implies the use of effective post-market surveillance and enforcement, as foreseen in the accompanying Regulation (EU) 765/2008/EU resp. Regulation (EU) 2019/2010. It will be important to consider how these mechanisms would apply to high-risk applications that are already regulated in this respect.

Section 3 - question: In your opinion, are there any further risks to be expanded on to provide more legal certainty?

Siemens Answer (Complete):

In our view, the current EU product safety legislation (e.g. Low-voltage Directive, Machinery Directive, General Product Safety Directive,) is so-called “total safety” legislation and covers all risks that arise or can arise from covered products and related technology. There is no need to expand on further risks in the text of the Directives. Clarification of aspects, such as relating to risk coverage or safety concepts, should be done through guidelines, where necessary.

As for “Cyber risks” and “personal security risks”, we consider cybersecurity and the related risks (including stemming from AI applications) as a “horizontal”, i.e. non product-specific, issue which should be regulated on the basis of a common, horizontal approach.

Section 3 - Question: Do you have any further considerations regarding risk assessment procedures?

Siemens Answer (Complete):

The risk assessment process and the principles of safety integration for risk mitigation according to the CE Directives have proven their effectiveness and have been successfully implemented. Today, almost all industrial sectors carry out risk assessment and implement risk reduction measures

according to processes that are required by the safety legislation. Therefore, the iterative process of risk assessment and risk reduction measures, as it is currently defined, does not need further considerations for AI.

Only if the scope of the AI application changes (other purpose or environmental context using the same data) the risk assessment must be re-initiated. By considering the life cycle of AI applications in the risk assessment one can ensure that AI application in the very immature stage (testing, research, exploration, etc.) are out-of-scope of the risk assessment procedure.