AI HLEG piloting phase survey on assessment list (general user) – techUK response

---

G1. After reviewing the draft Assessment List in its entirety, are there any important elements that could contribute to “Trustworthy” AI that were not covered by the Assessment items?

1 – yes
2 – no

G2. Please list the aspects that you found to be missing:

techUK welcomes the opportunity to respond to the European Commission’s Survey on the Trustworthy AI assessment list.

techUK represents the companies and technologies that are defining today the world that we will live in tomorrow. Around 850 companies are members of techUK. Collectively they employ approximately 700,000 people, about half of all tech sector jobs in the UK. These companies range from leading FTSE 100 companies to new innovative start-ups. The majority of our members are small and medium-sized businesses.

techUK would like to encourage the European Commission to consider improving the Trustworthy AI assessment list as follows.

1. Ensure that the AI assessment list can be operationalised

In its current form, the AI assessment list is lengthy and would be difficult for many companies to operationalise. AI start-ups and SMEs in particular have limited resource and there is a real concern that the Assessment List would introduce unnecessary burden to these businesses investing in AI-based technologies.

Many of the proposed questions lack the level of detail necessary and in some sections, such as Respect for Privacy, many questions appear to repeat themselves or overlap. To make the assessment list more manageable, techUK would recommend stripping out these redundant questions. Where possible, the assessment list should prompt action.

Many of these current questions invite a simple yes/no response, which may not always get to the heart of the issue. For example, a company may consider Q47 and decide not to involve different stakeholders in the development of the AI system. When in fact what this question is trying to do is help a company work out whether its training is likely to result in unfair bias. There is a risk that companies engage in some discussions internally but have no decision-making framework or guidance to help them make the right decision at the right time.

The questions should be specifically related to the adoption and deployment of AI systems, such as issues of explainability, data governance, fairness and bias. Many of the themes covered by the assessment list are related to good product design principles and are covered by existing horizontal legislation. To effectively
streamline the assessment process and avoid overlaps with other compliance efforts, we would encourage the revising and, in some cases, even removing questions, related to GDPR compliance.

The relevance of some questions will vary greatly depending on the sector or application. For the list to be widely applicable there will need to be a degree of flexibility on how these questions are answered. This allows teams to draw on questions where helpful from the assessment list, even if not to apply the list in its entirety. For certain sectors, for example financial services, more questions would be needed around the different risks the solution pose to the bank. In cases where automated decisions will have a significant impact on consumers’ livelihoods, assessment will need to be more stringent.

The assessment list would also benefit from being more practical and flexible. To begin with, the order of the questions in the assessment list should be re-ordered to reflect the way a typical AI product is developed and deployed to the market.

To help with the practical, operationalisation of the assessment list, the European Commission needs to include a selection of real-life case studies or examples. These case studies will help to contextualize the questions and will be instrumental as to how the guidance is interpreted. techUK would therefore recommend that these use cases are reviewed as part of a second consultation. The AI HLEG have suggested that the use cases should represent four sectors- healthcare diagnosis and treatment, autonomous driving, insurance premiums and profiling, and law enforcement. We’d encourage this list to be extended to represent a wider range of sectors, such as recruitment and financial services.

Another observation is that there are no procedural recommendations in the current assessment list, such as on training, composition on ethics boards or how regularly a product or service should be audited. There is also little reference to the practice of documenting the discussions and outcomes in order to track the thinking throughout the course of the AI systems’ lifecycle and use it as a governance tool to keep honest to the desired outcome of the AI system. A proposal on how to document and audit the decision-making process and use the assessment list could bring it to life, particularly with case studies.

Some further specific feedback on questions within the list include the following:

- Question 7 might be better worded as ‘Did you perform a formal threat modelling exercise against the AI system and identify points of risk to the overall system (e.g. cyber-attacks, physical infrastructure, data leakage, data pollution, etc)?’ Additionally, Q7 and Q7.1 are very similar and with this amendment to Q7, Q7.1 becomes less necessary.

- For Q8 it would be helpful to include examples, such as “For example, how does the system behave with extraneous input? With boundary conditions? And invalid or impossible data?”

- In relation to Q10, sufficient should be replaced with fail-safe.
• Within the resilience section, we would expect questions to feature penetration testing against AI systems to ensure there are no cyber vulnerabilities.

• On technical robustness and safety, there should be clarification on what is meant by ‘unacceptable accuracy/results’.

• The transparency section of the assessment list is currently quite vague and has some overlap with earlier sections on human agency and explainability. More precise language will be important particularly when addressed in more socially sensitive applications of AI.

• In the explainability section, this should take into account the different types of audiences (other businesses, developers, deployers or end-users).

• For Q36 it would be more helpful to ask, ‘Have you addressed explainability? How do you envisage the user experience of it?’.

• For Q44, determining an adequate working definition of “fairness” is difficult and may be undesirable. Defining fairness very difficult and is typically subjective and context specific. An automated decision can be fair for an individual and unfair for another simultaneously.

• Q45 under the accessible and universal design section, uses the term ‘preferences’ which is too vague and does not make the necessary distinction between a preference and a genuine need.

2. Clarify the target audience

The assessment list in its current form would entail an extensive cross-functional effort, spanning from lawyers, human and social rights experts, product managers, engineers, etc. and it’s very likely that a single person in a company could not meaningfully and independently answer all the questions.

Further clarity is needed on who the current assessment list of questions is aimed at specifically. The document states that the assessment list is for “AI developers, deployers and users to operationalise Trustworthy AI” but it is currently unclear which stakeholder groups should be considering, and answering, each of the questions. The AI life cycle development process involves a range of people that have different roles and responsibilities, which is not reflected in the current guidance. It would be helpful for the assessment list to include a taxonomy of AI experts, and introduce different questions for each category of these AI experts.

As the questions are currently drafted, we would anticipate that these questions should be answered by those developing the AI systems rather than the users of AI systems, however this isn’t clear. If the responsibility is placed on developers, this document doesn’t recognise the responsibility of the wider ecosystem, for example the buyers, suppliers, etc.
G3. How would you rate the usefulness of the Trustworthy AI Assessment instrument (assuming that its current smaller defects will be corrected after the pilot stage) in assessing the AI systems and products?

1 – not useful at all
2 – 
3
4 – extremely useful
5 – DK/NA

G4. Should you have any other remark or comment, please use the text box below to share those with us.

For the assessment list to remain relevant as technology evolves, it will need to be dynamic list rather than a static one. New use cases/scenarios will also need to be included as they emerge. Further thought is also needed as to how this list will cohabit with any future EU regulation related to AI.

The questions in the assessment list are based on the past- they are all about ‘did you…’ whereas the purpose of the assessment list should be to help business and technical teams address these considerations before, or during the design, development and then implementation and monitoring. techUK would suggest this assessment list should be addressing questions to help teams frame how they will identify issues or how they are considering identification of issues.

We’d strongly encourage the Commission to listen to the feedback of industry players that have attempted to use these guidelines in real-life scenarios and use the lessons learnt to define the checklist.