Position paper by EUROLAB

In response to

EC report COM(2020) 65 final:

"On Artificial Intelligence - A European approach to excellence and trust"

EUROLAB Secretariat
Rue du Commerce 20-22
B-1000, Brussels
Belgium

Tel: + 32 2 511 50 65
Email: info@eurolab.org
Website: www.eurolab.org

Content

Summary

EUROLAB aisbl thanks the EC for delivering a white paper[1] on artificial intelligence, in which risks and opportunities of AI are clearly highlighted. The white paper mentions pathways where AI can be used to help societies take advantage of such a technology and to mitigate associated risks.

The EC white paper explicitly clarifies that AI is made up of data and algorithms, which are amenable to be incorporated into hardware. Such option will increase the impact and contribution of AI on how conformity assessment bodies (CABs) perform their task.

It is impossible to conceive a safely functioning society without conformity assessment of products and services, to a large extent carried out by the private sector. Despite the enormous importance of conformity assessment, this area is generally not well known beyond the relatively small world of its direct stakeholders. At the same time, everybody will underscore the importance of having access to some form of information about the safety or quality of products and services. With the advent of powerful AI, the impression might gain ground, that AI based recommendations on safety and quality would suffice to maintain and build societal trust in all the products and services which a society needs in order to prosper.

This position paper is short and focused. It aims to address the question on how to deal with AI in the context of conformity assessment. It voices EUROLAB's position on the role AI can play in accredited conformity assessment.

This paper clearly points out that AI requires certified calibration against standards, in order to be able to judge the quality of its deductive or reasoning performance. In case such AI-standards do not yet exist, their provisioning should build upon the quality infrastructure that has proven its value in Europe. We advise to appoint a European working group made up of manufacturers, software developers, conformity assessment bodies, and research and development institutions.

It concludes that - while AI is a tool that will (and already does) enhance and improve the service of conformity assessment - it should never replace conformity assessment by an accredited CAB.

The EUROLAB Board of Administrators

July, 2020

[1] EC report COM(2020) 65 final

Acknowledgements

The Board of Administrators are thankful for the contributions to this position paper by:

- Dr. Bernd Kroon (Fenelab, Netherlands),
- Dr.-Ing. Michael Nitsche (BAM, Germany),
- Mr. Gerben Visser (Kalibra, Netherlands),
- Mr. Luca Salini (ALPI, Italy),
- and Dr. Guillaume Avrin (LNE, France).

Considerations

EUROLAB aisbl is the **European Federation of National Associations** of Measurement, Testing and Analytical Laboratories, composed of 26 national associations from Europe and beyond. Grouping more than 3000 members (each of them being an accredited conformity assessment body, or CAB), in over 9000 accredited laboratories and representing over 150,000 professional experts. Seated in Brussels, EUROLAB provides a platform for their members to discuss issues of concern and to exchange information and viewpoints. Additionally, EUROLAB acts as a facilitator of their members' interests to the EU institutions.

CABs assess whether products and services comply with standards. Accreditation safeguards the trustworthiness of CABs. Accreditation is given to a CAB if they prove to be governed by a professional management system. The management system must maintain an adequate knowledge level of their personnel, must provide adequate instruments and workflows and must yield a governance system to protect impartiality. In doing so, CABs execute an important part of a multiparty, interconnected process that allows societies and citizens to rely on quality and safety of products and services which are required to live and interact with a minimal (or at least acceptable) level of risk.

With the advent and maturation of Artificial Intelligence (AI) the performance of products and processes will increasingly rely on AI. This holds true for the products and services which are assessed by a CAB, as well as the service of the CAB itself. While not part of this position paper, we wish to underscore that the increasing digitalization will require serious attention for security policies which serve as the foundation for the triad of confidentiality, integrity and availability of information. We limit our discussion here to the application of AI on the products and services that are assessed by a CAB.

The power of AI to leverage the value of analytical services is extensive and fundamental. Think about the diagnostic interpretation of an X-ray image. The input is the original X-ray image, the output a measured response in the form of a qualitative statement on health. And AI is the transfer function that translates input to output. Yet even though the exact pathway is unknown of how such AI algorithm works to arrive at a certain conclusion, it is, in a sense, as magical as any other complex non-linear computational procedure that yields an output based on an input.

The real difference with classical methods is that AI is preferably used to translate data into information. Typically, the amount of data is large, the location where the data originate or processed unknown. The size and nature of the body of knowledge against which the data are referenced is not well defined and may change over time, the behavior of the driving algorithms fuzzy rather than exact. Therefore, a deterministic approach where one would use existing, non-AI based conformity assessment tools in order to validate all elements of AI based evaluations poses an unsurmountable problem. It would be like trying to understand the behaviour of all atoms in a ruler in order to evaluate the correctness of measuring a distance.

When classical technologies are used in accredited conformity assessment to yield a quantitative or qualitative statement about an object, they must be calibrated. Calibration is, simply put, the act of establishing the reproducible accuracy (or 'measurement uncertainty') of the ability an instrument to provide a number on the known value of a 'standard' object. Calibrated instruments do not necessarily yield the true value of an object. But the users know at what accuracy a calibrated instrument performs a measurement. As it is the case with classical instruments, AI based instruments should be calibrated.

It is quintessential that standards shall be developed for AI based conformity assessment processes (existing and emerging). Developing standards is sometimes straightforward, many times it is more an art, requiring a blend of knowledge of practical applications, error-proofed procedures, deep insight in risk and failures characteristics of products, desired levels of safety, and highest possible scientific R&D knowledge. The choice of standards drives the ultimate level of attainable trust in all products and processes where those standards were applied. EUROLAB advises to build upon the European quality infrastructure that has proven its value and develop it further within a realm of an urgently needed "Digital Quality Infrastructure".

At this point we have reached the general insight that all AI-based conformity assessment processes should be calibrated. We further propose to divide the application of AI in conformity assessment into two classes:

1. *AI as an indirect conformity assessment tool (ICA).*

If the analyses of an X-ray image would be part of an accredited service by a radiologist, then the previous mentioned X-ray example would classify as an ICA. At the end of the day, it is the hospital, where the radiologist works, that would be accredited to carry out the diagnostic service. Many more examples of such ICA's are found in laboratories carrying out complex chemical or biological analyses. Most of the state-of-the-art instruments that are used for such services use AI modules to interpret the raw analytical data and to make information out of that raw data.

2. *AI as a direct conformity assessment tool (DCA)*

The AI associated interpretative power (and magic) easily creates an impression of truth onto the information that it generates. This aspect of AI is amenable to be used as an authoritative entity in itself, one that could be assigned responsibility.

We further propose to put relevance to the place where data are processed through AI algorithms: it can be either internally (within the device which generates the raw data, including its peripheral on-site computing power), or externally (cloud computing).

For our evaluation, we can now evaluate four quadrants to place AI aided conformity assessment: (A) ICA Internal, (B) ICA External, (D) DCA Internal and (D) DCA External. We shall assume that AI will be calibrated in all four scenarios, as explained above.

Evaluation

Using the previously created quadrants it is possible to evaluate the consequence and advise on the different ways AI may impact conformity assessment.

A. ICA-internal
With ICA and data internally: this, in fact, is business as usual. Fundamentally, it does not matter at all, what kind of algorithm is used to transform data to information. If, like with the Turing test[2], one cannot infer from the result who carried out the calculation, it is not a factor of relevance, as long as it is possible to assign responsibility for adding credibility to the results. With ICA-internal, the accreditation is given to the organization for a defined test, inspection or certification result. The system of accreditation, not the fact that AI played a role, yield trustworthiness to the results.

B. ICA-external
In case of ICA and data externally (i.e. cloud computing): this scenario is almost the same as scenario A. With one difference: data are transferred outside the region of control of the CAB, using AI to return information on which to base a conformity assessment. This scenario is acceptable if the data-processing entity complies with all data protection laws as they are in place for the region where the conformity assessment is carried out. Furthermore, the importance of calibrated AI performance increases significantly when AI resides external to the CAB.

C. & D. DCA-internal and external
In principle, DCA could yield the same reliability of assessment as ICA. But this is an utterly wrong perception. There are many examples showing that AI based conclusions can be biased, or flawed, often unintentionally. Citing the EC's white paper "*…their [AI based products] behaviour is largely defined and constrained by its developers. Humans determine and programme the goals, which an AI system should optimise for.* ", it is clear that it should not be the AI algorithm who carries responsibility for its behaviour, but the instance who created it.

Conformity assessment under accreditation goes one step further: not the creator or owner of an instrument, but the organization who uses it carries the responsibility for the validity of the assessment. Therefore, under the current definition of accreditation DCA may yield an assessment of a condition, but it will not be possible to offer accredited conformity assessment based on DCA-internal, or external.

Yet with increasing power of AI, DCA may become an acceptable proposition for accredited conformity assessment. The fundamental problem then is: who carries responsibility? Should the AI be switched off if it does not meet the accreditation requirements? Given the situation where a car can drive autonomously: who is responsible in case of an accident? The car manufacturer, the AI algorithm, or the driver?

We would like to suggest that such a form of conformity assessment should only be possible for very low-risk (impact x occurrence) applications. An example of such a low risk application would be the assessment if music has a good quality or not. If one would allow DCA for any other application, it will then become beyond manageable control to assign responsibility in case such DCA system turns out to be in error, impartial, biased or a combination thereof.

[2] The Turing Test, defined by Alan Turing in 1950 as the foundation of the philosophy of artificial intelligence.

Conclusion

It is a fact of a progressive human development that AI will claim more space and impact in the way it will act in our societies. Rather than fighting AI, EUROLAB embraces it and welcomes the opportunities it offers to provide reliable conformity assessment in increasingly complex areas.

EUROLAB strongly advises to create a novel European standardisation body, with representatives of all relevant knowledge stakeholders, to create applicable and relevant standards for AI procedures and processes.

EUROLAB recognizes the value of human ability to take responsibility for the application of AI in complex analytical procedures. EUROLAB sees no threat - nor need to mitigate risks - of AI per se in the field of accredited conformity assessment. Our position assumes that associated Cybersecurity risks are mitigated by recognition and application of existing and emerging European legislation.

EUROLAB does not endorse any application of AI as stand-alone, accredited conformity assessment but for the most trivial and very low risk, assessment tasks.

Used abbreviations

AI	Artificial intelligence
CA	Conformity Assessment
CAB	Conformity assessment body
DCA	Direct conformity assessment
aisbl	Association internationale sans but lucratif (International Non-Profit Organization)
EC	European Commission
ICA	Indirect conformity assessment