Commentary on the EU Commission White Paper on Artificial Intelligence, BMW Group

The following commentary relates to documents:

- European Commission COM(2020) 65 final https://ec.europa.eu/info/sites/info/files/commission-white-paper-artificial-intelligence-feb2020_en.pdf

Regarding and with reference to

- European Commission COM(2020) 64 final https://ec.europa.eu/info/sites/info/files/report-safety-liability-artificial-intelligence-feb2020_en_1.pdf

both in the English language version.

I. General assessment:

The White Paper rightly emphasizes the significance of location when it comes to AI technology. The importance of fundamental research, development and industrialization of AI is reflected in the need for coordinated European approaches, a single economic area for data and AI, dedicated “lighthouse centres of research”, as well as “excellence and testing centres”. These intentions must be supported. We also agree that AI systems have considerable benefits for the whole of society beyond the individual perspective and can make a substantial contribution to the Sustainable Development Goals.

A positive aspect to underscore is the plan to invest around € 20 billion per year in Europe over the next ten years, particularly in light of rapidly accelerating international competition. A differentiated approach to initial and advanced training should not only focus on educating and retaining talented AI researchers and developers in Europe, but also, more generally, on providing citizens and the workforce with the qualifications necessary to allow them to apply AI and deal with the associated changes. The significance of a European data strategy as an essential foundation for AI was recognized and emphasized, along with the important role of international cooperation, standardization, harmonization and mutual recognition of standards and a regulatory structure.

Dangers and risks are associated with AI, and these are amplified by how AI is portrayed in the media. These possible threats do not correspond to the actual applications and real capabilities of AI technologies currently available. In relation to this, the White Paper highlights that AI will only be accepted by the general public if people feel they can trust the technologies and if there is a human-centric approach, in line with fundamental rights and data privacy.
Unfortunately, instead of putting the focus on the potential for greater productivity, product safety, performance and objectivity, the White Paper overemphasizes the concerns and fears regarding AI and, based on this approach, derives a very broad risk filter and, ultimately, a need for strict regulation. Differentiating and defining AI poses a significant challenge in this regard, because the White Paper falls back on a phenomenological description, which remains too vague and would need to be clarified. The development of uniform and clear rules for handling this technology across the whole of Europe is very welcome in order to avoid a fragmented landscape of national structures. Support should be given in particular to the development of clear rules on product safety and liability matters, which ensure sufficient legal certainty for the development, market launch, operation and application/use of products and services. Even if these are conceived in a technologically neutral manner, they must, if necessary, be adapted to the specific properties of the AI technology. However, it must be ensured that this development does not lead to overregulation, which would inhibit innovation and stand in the way of the intentions for the future development of AI in Europe, described in the introduction of the White Paper.

II. Commentary on specific passages selected from the White Paper on Artificial Intelligence:

- **Section 5, p.9 et seq.: "An Ecosystem of Trust: Regulatory Framework for AI":**

This section gives the impression that trust in the technology can be achieved through regulation and certification only; the term "trust" is used unspecifically for numerous aspects, from user acceptance to legal conformity, compliance with ethical standards and respect for (European) values. The entailed "Problem Statement" is similarly unspecific, but significant regulatory demands are derived from it. On the one hand, the seven key requirements for trustworthy AI formulated by the EU High-Level Expert Group on Artificial Intelligence (AI HLEG) are reasonable and should be supported. On the other hand, these need to be conveyed in specific legal and regulatory terms in such a way that the existing legislation on basic and human rights, consumer protection, product safety and liability is both reinforced in terms of its applicability and only expanded and supplemented where new technologies such as AI explicitly create new risks or obstacles.

Here, differentiating and defining AI poses a significant challenge. The White Paper falls back on a phenomenological description (p.16 et seq., with reference to COM(2018) 237 and AI HLEG): It describes the perception of the environment through data acquisition, the interpretation of this data, conclusions drawn based on knowledge or information gleaned from the data, and decision-making in relation to a complex objective as key elements of an AI system. This depiction carries the risk of introducing regulations for traditional software systems under the umbrella of the AI regulatory framework. In this context, autonomous vehicles are listed as an example of such AI systems. However, key features of highly automated vehicles can already be implemented without the use of AI: Traditional driving dynamics control and advanced driver assistance systems already assess the vehicle's surroundings via distinctive sensor technology, use this information to regulate the driving and braking forces acting on individual wheels to ensure superior driving stability and prevent the vehicle from swerving, employ radar to maintain distance from the vehicle in front, and warn the driver if they are deviating from the road lane or help them stay in lane in certain traffic situations. All these functions, which are already available on the market, are implemented without trained AI algorithms, but display a certain degree of environmental perception, autonomy and adaptivity which is in this White Paper being associated with AI. If AI is not clearly defined and differentiated, the risk arises that existing functions, services and products will be subsumed under the umbrella term of AI. Putting highly automated driving on a par with AI and directly deriving from this a need for additional regulation that goes beyond previous rules seems problematic, especially where extensive, directly applicable regulations already exist for product safety and vehicle homologation.

- **Section A, p.10: "Problem Definition":**

The risks are described as risks caused by properties that are supposedly specific to AI technology, such as "opacity", "complexity", "autonomy" and "updatability". It is undisputed that increased reliance on these properties by systems can carry risks, for instance, because existing legal standards could be more difficult to be applied or enforced or rendered completely unenforceable, the regulatory scope may no longer be appropriate to actual circumstances or adequately cover these, or the responsibilities of economic operators may no longer be adequately defined from a legal standpoint.

However, the properties described do not only apply to AI systems but also occur when using other technologies. Consequently, the greatest possible degree of **technology neutrality** should be respected in the legal standards. Furthermore, existing systems and technologies should not be subject to unnecessary tightening of requirements.

- **P.10 et seq.: “Risks for fundamental rights, including personal data and privacy protection and non-discrimination”:**

AI algorithms and the underlying data are attributed with a possible discriminatory bias. A focus on human oversight of the automation of discriminatory processes by AI is recommended where that automation is more difficult to understand. Depending on the risk assessment and the type of algorithmic decision-making, mechanisms may be conceived that either allow all algorithmic decision-making to be confirmed by humans or at least require ongoing monitoring of automated algorithmic decisions.

In addition, although the data on which an algorithm is based still plays a significant role, potential discrimination only occurs when the trained algorithm is applied. Any potential regulation should take this into account and not create major additional obstacles for the recording and quality of training data in itself, but should word the requirements in such a way that any potential discrimination in the selection of data and its use in training algorithms is sufficiently considered.

Furthermore, the extent to which earmarking and provenance of training data can in itself be problematic with regard to a possible bias should also be examined. As a minimum, international exchange and clearly regulated acceptance of data must make it possible to select representative training data sets that cannot be generated per se in a self-contained economic area such as the EU because the groups in question are not sufficiently represented in the average population. Damage frequency and risk must not be confused here. For example, individuals’ clothing, skin color, ethnic background or culturally determined individual behavior in road traffic should not lead to a higher risk of harm for that specific individual because they only rarely travel within the EU and are underrepresented in the training data gathered there.

- **P.12, Section 3, “Risks for safety and the effective functioning of the liability regime”:**

It is true that a lack of insight/controllability relating to AI (especially with third parties as intellectual property) in autonomous driving can bring about new issues regarding how to allocate product liability or enforce legal claims. However, the fact that autonomous driving leads to a significant increase in safety (no tiredness, drunk-driving, etc.) is not mentioned. Here, it should be clarified that AI has no value in itself and the use of AI is generally associated with better performance and also an increase in product quality and safety. Therefore, AI should not unnecessarily be associated with new, additional or increased risks, which are based on speculation. Instead, the focus should be on the fact that the risks involved in using AI may be difficult to describe and quantify and that the comprehensibility of AI systems and the decisions they make or assist in making should be given particular consideration.

- **P.13: “Product Liability Directive” and Section B.: “Possible adjustments to existing EU legislative framework relating to AI”:**

This addresses the issue that data subjects will demand access to relevant data as evidence in the event of a claim. In this context, it must also be considered that the rights of a data subject as stipulated in the GDPR will also be invoked in order to request the relevant data (particularly the right to access in accordance with the GDPR). The fact that reviews and adjustments to the legal provisions are being considered in order to address the actual risks of AI systems is appreciated.

- **P.14, Section 4: “Limitations of scope of existing EU legislation”:**

If AI from third parties is integrated into vehicles, the software and services should be understood as products and the associated risks and responsibilities be passed on via relevant

contracts. The existing laws are functional and effective. As a rule, every company has a significant interest in bringing correct and, above all, safe software onto the market.
It is not clear what the specific characteristics of AI software are, when compared to traditional software, for which new legislation would be necessary. Why are the laws on product safety not sufficient and why do new regulations need to be devised because of AI software?

- **P.14, Section 5: “Changing functionality of AI systems”:**

The changing functionality of existing software in the automotive sector is already regulated and controlled. As AI systems are also approved under the same standards, it is unclear why further regulation is needed. Existing software also poses a risk, and this is already controlled. What changes so significantly through AI that a new regulatory framework is required?
In general, new functions already need to be approved, homologated, monitored and observed in operation. If updates become more frequent in future, a product – whether based on AI or not – may change over its life cycle in such a way that it no longer resembles its originally approved state. It is true that certain uses of AI speed up such updates or allow them to be carried out autonomously, which may entail a need to at least clarify existing regulations. Particularly for safety-critical functions in the automotive sector, far-reaching approval regulations, the General Product Safety Directive and the like, which define the procedure for determining the approved status of a product and any deviations thereof, apply in all these cases. There is no discernible need for new regulation, especially with specific relation to AI.

- **P.16, Section 5: “For the purposes of this White Paper (...)”:**

In contrast to the usual “algorithm development”, this section refers to “AI algorithm training”. It should be emphasized and ensured that the resulting regulations only apply to trained software, not traditionally developed software. In turn, this “trained” software only relates to a specific sub-category of AI. This again raises the question regarding the extent to which the regulations can be kept technologically neutral. In general, however, the product liability laws and product monitoring obligations, which are already extensive, apply (see the remarks above on the definition of AI, and the comments on trust and the “problem definition”).

- **P.16, Section 5: “Algorithms may continue to learn when in use (…)”:**

Re-training may be carried out, but certainly not “when in use” for safety-critical functions. For these, particularly in the automotive sector, the approval and homologation requirements continue to apply, such as those stipulated in the General Product Safety Directive, the approval and operational safety regulations, as well as the product monitoring obligations and technical monitoring requirements. From a technical standpoint, autonomous learning of individual vehicles is, in most cases, not reasonably feasible, because although on-board computers have the capacity to run trained algorithms (inference mode), they are not capable of running deep neural networks in training mode at the same time. In our opinion, the expression “While AI-based products can act autonomously by perceiving their environment and without following a pre-determined set of instructions...” is misleading because even a trained algorithm at its core contains instructions defined during the training process.

- **P.17, Sections 3–5: “The determination of what is a high-risk AI application (…) should consider the sector and the intended use involved (…) applied in a sector where significant risks can be expected to occur (…)”:**

It is important to prevent entire sectors from being placed under general suspicion, as each sector involves applications, products and services with different risk requirements. In every sector, risks must be assessed on a case basis. Even though there are already sector-specific rules in the transport sector, the following is worth considering: a list of “high-risk sectors” is very difficult to manage and could significantly impair the use of AI for non-critical functions/services in all the sectors listed. On the other hand, it is unclear why a risk classification based solely on the particular AI application would not be sufficient. Every AI system must be considered separately, because otherwise the training outlay for the vast majority of uses that pose no ethical risk (e.g. in production) would be greatly increased as a result of general suspicion. This would unnecessarily complicate and impede the use of AI altogether and would significantly harm the competitiveness and employment situation in Europe / Germany / Bavaria. A classification or segmentation by "high-risk sectors" carries no benefit for any party, but does impede the location's innovative power and must therefore be rejected. The focus should be on the AI applications and on a differentiation based on how critical its function is ("safety-relevant applications" vs. "non-safety-relevant applications").

- **P.18, Section D: "Types of Requirements":**

The categorization of requirements is concise and useful. A clearly differentiated view would be welcome on where it is about clarification of the application of existing rules for data-driven AI systems and where about new, AI-specific draft regulations.

It can be particularly difficult to formulate requirements for data whose suitability depends significantly upon its use in a specific product or service. Legally binding regulated guidelines on data quality requirements would need to be described, if at all, in the context of their real-world application. As data can generally be used in or to train multiple different applications, it will be difficult to stipulate absolute data requirements. It is important to state clear purpose limitations and meaningful metadata to describe the data sets, as a basis for making and documenting a choice for product development. However, this is already covered in the extensive regulation of product liability, quality, safety and reliability.

While particularly the requirements for "robustness and safety", "record-keeping" and "information to be provided" are already well adapted to the established processes in the automotive industry, it is not clear, especially in the discussion of "human oversight", how the current duties of care regarding development, the use of the latest technology and product monitoring can be redefined or improved upon. Again, great care should be taken to address only **new risks actually arising** from AI, but these are not yet covered clearly enough in the White Paper.

- **P.18-19: Section a) – "Training Data":**

Just like specifications or software, data is fundamentally a development artefact, especially when it comes to data used to train algorithms or test functions, services or other products. They are subject to the same requirements for product liability, safety and reliability and should therefore always be considered in connection with their real-world use. The consistent application of these rules on data and algorithms created (trained) by this data must be ensured and may require clarification, particularly with regard to quality and verification mechanisms. New, additional requirements specifically for data would entail a considerable increase in complexity. In the White Paper, it remains unclear what the aim of this new regulation would be beyond that of existing legislation.

- **P.19, Section b) – "Keeping of records and data":**

We welcome a clarification regarding the documentation and retention obligation for development documentation including – in legitimate cases – the training data sets and training methods. However, it is not clear to what extent there is really a need to act regarding the automotive industry beyond the existing approval regulations and the areas already addressed in the relevant working groups (WP29, VMAD (Validation Methods for Automated Driving)) for specific areas of application such as autonomous driving.

- **P.19: Section b) "Keeping of records and data", bullet point 3 – "avoiding bias that could lead to prohibited discrimination (…)":**

This would mean that a professional data discrimination management process would need to be established. However, it is not apparent why a new, additional anti-discrimination regulation specifically for data and AI algorithms is necessary. In fact, discrimination has been covered by law for a long time; there is no convincing argument in the White Paper that AI brings a new quality to this.

- **P.20: Section c) – "Information provision":**

The White Paper emphasizes the significance of adequate information for the "affected parties", but also sees the necessity to not create unnecessary obstacles. When legislation is drafted, it should consider including exceptions to the obligation to provide information under Article 13 GDPR when processing personal data by means of AI, similar to those already provided for under Article 14 (5) b GDPR, e.g. if and to the extent that providing the information proves impossible or would involve a disproportionate effort. Furthermore, it would also be useful to establish a certain standard or guideline as to what the information obligation would look like in AI systems.

- **P.21: Section f) "Specific requirements for remote biometric identification"**

The White Paper references Article 9 GDPR for the processing of personal data and presents the prospect of a European discussion about specific situations in which the processing of biometric data could be justified. Article 9 GDPR (in conjunction with Article 6 (1) GDPR) provides for increased requirements for processing special categories of personal data, which includes biometric data. Including specific grounds for allowing the processing of biometric data in an AI context should be considered when developing a potential legislative framework.

It would also be useful for this legal framework to clarify that not only research in the public interest is covered by the **research privileges** when processing personal data (cf. Article 5 (1) b (2nd half sentence) GDPR, Article 9 (2) j GDPR, Article 89 (2) GDPR) in the case of research involving AI. Recital 159 GDPR already stipulates that privately funded research also has recourse to these privileges. However, privileges for research interests in the private sector are sometimes denied when it comes to discussions on data protection law. For example, it is sometimes stated that privileges arising from Article 9 (1) j GDPR only apply to research in the public interest (Schiff, in: Ehrmann/Selmayr, General Data Protection Regulation, 2nd ed. 2018, Article 9, Recital 63), although this is not how the standard is grammatically worded.

- **P.23: Section F, bullet point 4 – "… re-training the system in the EU (…)":**

It is not clear how the location of the (re)training should influence the result. Regulation methods must ensure they are valid regardless of the location in which the work is carried out. On the other hand, the quality of a training algorithm is highly dependent on the data entered. We consider the training of algorithms with meaningful European data that meet high standards to be unequivocally beneficial for use in products that are to be approved in Europe.