



Brussels, 19 September 2018

Dear Members of the High-Level Expert Group on Artificial Intelligence,

On behalf of BSA | The Software Alliance, I congratulate you on your appointment to the European Commission's High-Level Expert Group on Artificial Intelligence ("HLEG"). The formation of the HLEG is a unique opportunity for Europe's leading experts from industry, academia, and civil society to help the European Commission develop a "coordinated approach to make the most of the opportunities offered by AI and to address the new challenges that it brings."¹ We agree with the Commission that the success of such a framework will turn in large part on whether it fosters an "environment of trust and accountability around the development and use of AI."² We write to offer initial views to help inform the work that lies ahead, particularly as it relates to the immediate task of developing draft ethics guidelines for AI.

BSA is the leading advocate for the global software industry before governments and in the international marketplace. Our members are at the forefront of software-enabled innovation that is fueling global economic growth, including cloud computing and AI products and services.³ BSA members include many of the world's leading suppliers of software, hardware, and online services to organisations of all sizes and across all industries and sectors. BSA members have made significant investments in developing innovative AI solutions for use across a range of applications. As leaders in AI development, BSA members have unique insights into both the tremendous potential that AI holds to address a variety of social challenges, and the governmental policies that can best support the responsible use of AI and ensure continued innovation.

The smart use of AI has the potential to spur tremendous economic growth across every industry sector, improve human decision-making in ways that will make the world more inclusive, and enable cutting-edge breakthroughs on vexing social challenges such as climate change and cancer research. At the same time, AI services that are not developed and used responsibly, like other groundbreaking innovations, creates risk of unintended consequences or malicious uses. Governments around the world are therefore rightly focused on developing policy frameworks to address both the opportunities and possible risks associated with certain applications of AI. BSA members are equally attuned to these challenges and are committed to developing AI in a manner that will maximize the benefits and minimize the risks associated with the deployment of the technology.

¹ European Commission, *Communication on Artificial Intelligence for Europe*, at p. 3.

² *Id.*

³ BSA's members include: Adobe, ANSYS, Apple, Autodesk, Bentley Systems, Box, CA Technologies, Cadence, CG Tech, CNC/Mastercam, DataStax, DocuSign, i&k software, IBM, Informatica, Intel, Intuit, MathWorks, McAfee, Microsoft, NetCad Ulusal CAD, Noctranet, Okta, Oracle, PTC, Salesforce, SAS Institute, Siemens PLM Software, Splunk, Symantec, Trend Micro, Trimble Solutions Corporation, and Workday.

Artificial Intelligence – a Multi-Sectoral Technology

The HLEG guidelines are poised to become the point of reference for AI ethics in Europe, not only supporting the work of EU Institutions, but most importantly strengthening trust in the general public towards such revolutionary technologies. In addition, the HLEG guidelines will constitute an important benchmark for AI development throughout the Union, thus reducing fragmentation at the Member State level.

As the HLEG works to develop its AI ethics guidelines, it is important to bear in mind the incredibly diverse set of underlying technologies and use cases that are comprised in the AI ecosystem. Indeed, AI systems can be used in an almost unimaginably wide variety of different contexts, and to improve an exceedingly diverse array of business and consumer experiences across a range of applications and devices. Moreover, exciting and beneficial new applications of AI are constantly being developed, making it difficult to predict the full range of potential uses cases when considering ethical use considerations at a single point in time.

In view of the complexity and evolving nature of the ecosystem, endeavoring to create a single set of ethical “requirements” for all forms of AI cannot realistically account for all potential consequences running the risk of being over-prescriptive, chilling or even prohibiting beneficial uses of AI. Because the salience of any single ethical principle will vary depending on context – context which depends on uses as well as different cultural perspectives – the HLEG should not endeavor to create bright line rules. Instead, the HLEG is right to remain focused on identifying a high-level framework of ethical principles that is adaptable enough to account for the unique concerns that might be implicated by a particular use case. A risk-based, adaptable framework is critical for ensuring both the continuous innovation of this promising technology, while also providing developers and users of AI with a tool that will help ensure it is designed and deployed in a manner that is consistent with European values.

Promoting Confidence and Trust in AI Systems – Guiding Principles

Ultimately, the HLEG’s goal should be to create a framework of ethical principles that developers and entities deploying AI can use to ensure that the public has confidence and trust that AI systems are operating in a manner that is consistent with their expectations. Facilitating increased confidence and promoting trust in the use of AI technologies is an important priority, and BSA has highlighted five key principles that could aid industry in pursuing it:

- (1) **Fairness**—considering measures to evaluate AI systems to help recognize improper or unconscious bias;
- (2) **Accuracy**—acknowledging the importance of data quality and, where feasible, identifying sources of error in data inputs and system outputs;
- (3) **Data Provenance**—considering measures that could facilitate evaluation and documentation of data used to train AI systems, how those data are collected, and how data is used over time within AI systems, consistent with any other data retention obligations;
- (4) **Explainability**—exploring how to provide reasonable explanations of how AI systems operate; and
- (5) **Responsibility**—considering whether processes are available to address unexpected issues that may arise after AI products and services are deployed.

Operationalizing Ethical Principles – The Need for Flexible Approaches

The appropriate measures for ensuring the public has trust and confidence in any particular AI system is inevitably a context-specific exercise. There are many forms of AI, which vary in complexity, and are used to enable different services. As a result, there is no one-size-fits-all approach. Instead, considering the variation in technologies and applications of AI, and ongoing research efforts, there are many possible ways to build trust and confidence in AI systems, and in many cases they may overlap. For example, such practices could include:

- Conducting in-house testing and evaluation of AI systems to ensure they meet their specified goals;
- Developing guidelines and providing necessary resources to developers to help evaluate fairness and guard against improper bias;
- Identifying persons with relevant expertise who are responsible for addressing significant problems identified with operating AI systems;
- Ensuring subject matter experts, especially those with knowledge of the policy landscape in which the AI system will be deployed, are available to assist computational scientists in the design and implementation phases;
- Providing descriptions of procedures used to assess the quality of data inputs and address errors identified in outputs;
- Providing general descriptions, where appropriate, of training datasets that AI systems use to learn;
- Developing mechanisms for consumers to request information, obtain guidance and address potential concerns;
- Continuing monitoring after product release to detect and address unintended outcomes;
- Providing visual aids and/or plain language explanations that communicate important facts about AI systems and their operation; and/or
- Supporting continued research and analysis of transparent modeling.

The degree to which any specific ethical principle or implementing measures will be relevant to the objective of promoting trust in an AI system will naturally vary. For instance, public trust in an AI system that provides restaurant recommendations based upon user-inputted criteria is unlikely to require an intricate level of explainability. In contrast, when an AI system is deployed in a context that affects consumers' eligibility in consequential areas, such as access to credit or housing, the public will rightly have far greater expectations about measures the system's developer has taken to ensure it is consistent with shared expectations of fairness, accuracy, data provenance, explainability and responsibility. In developing the AI ethics framework, the HLEG should ensure that it is flexible enough to account for such context-specific nuances, including considering the differences between uses that are consumer-facing and those that are developed as enterprise solutions. Given the context-specific nature of an AI risk analysis, the private sector has a particularly vital role to play in crafting use case-specific principles, best practices, and guidelines to promote trust and confidence in their AI systems.

It is also important to recognize that there may be contexts in which efforts to safeguard one ethical principle might come at the expense of others. For instance, there may be circumstances in which design choices aimed at preserving the fairness and accuracy of an AI system could involve trade-offs with the degree to which the underlying model can be made explainable to the public. For example, in designing a fraud detection system, the need to ensure that it is operating in an accurate and unbiased manner may necessitate limited disclosures to the public about how it operates. Indeed, there are instances where too much transparency can have the unintended consequence of making an algorithm vulnerable to gaming. Moreover, research has shown that disclosing the algorithms, source code, or associated data sets is ineffective in helping to provide explanations, in part because they cannot be meaningfully understood in isolation.⁴

⁴ See Kartik Hosanagar & Vivian Jair, *We Need Transparency in Algorithms, But Too Much Can Backfire*, Harvard. Bus. Review, July 23, 2018, available at <https://hbr.org/2018/07/we-need-transparency-in-algorithms-but-too-much-can-backfire>.

AI Challenges and Opportunities

As the HLEG works to develop an ethical framework for AI, it is critical to bear in mind the significant positive social impacts that the technology can have. Notably, AI can be used to reduce bias in decision-making and promote social inclusion. For example, AI systems, powered by data analytics, are at the heart of new devices and applications that can improve the lives of people with disabilities. For instance, AI is helping people with vision-related impairments interpret and understand visual content, such as photos and their physical surroundings. This technology opens new possibilities for people with vision impairments to navigate the world, giving them increased independence and greater ability to engage with their communities. Thus, while it is appropriate to ensure that AI is not designed in ways that imperil important ethical considerations, it is likewise important not to impose a set of stringent requirements that might undermine potential uses of AI that advance social good.

A static set of ethical requirements is unlikely to help AI developers or users address the unique challenges they may confront. Overly-prescriptive requirements based on today's technological landscape are also likely to become quickly outdated. We therefore recommend that the HLEG pursue an ethical framework that is based on realistic risk and proportionate to the concrete harms that might be implicated by a specific use case. Moreover, such a framework will be most effective if it is technologically-neutral and outcome-oriented to ensure that it remains enduring as the state of the art evolves. Rather than creating any technological red lines, an effective ethical framework should provide developers and users of AI with a toolset to perform a risk analysis aimed at identifying and mitigating risks of unintended consequences and ultimately to ensure that AI is designed and deployed in a manner that is consistent with European values.

While the guidelines need to reflect European values, we must also keep in mind that AI will be developed and deployed in an international context. If European good practice or European guidelines are too draconian, prescriptive or overly rigid, AI will be developed elsewhere and other geographies will reap the benefits of AI deployment while Europe is left behind. The international standards community is beginning to address many of the issues raised in this paper. BSA recommends that European authorities and industry fully engage in these international efforts.

We remain at your disposal, should you need additional information or have any inquiries about the work of BSA | The Software Alliance. We look forward to working with you on the development of Artificial Intelligence in the European Union.

Yours faithfully,



Thomas Boué
Director General, Policy – EMEA
BSA | The Software Alliance