Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

CIPL's Response to the EU Commission White Paper "On Artificial Intelligence – A European approach to excellence and trust"

How to Leverage the GDPR, Accountability and Regulatory Innovation in AI Development, Deployment, and Uptake

Executive Summary

CIPL recommends a minimal, risk-based and layered approach to regulating AI, relying on existing laws and standards and building on accountable practices of organisations. This approach should be backed by innovative regulatory oversight and co-regulatory instruments. To properly maximise AI benefits while minimising risks, the EU AI regulatory approach should:

(1) Rely on impact assessments performed by organisations to trigger the application of the law that would take into account the context and impact of a proposed use of AI, rather than the sector it is utilised in or its type. The regulatory framework would provide illustrations of rebuttable presumptions of high-risk, rather than rigid pre-defined classifications. Organisations would assess the overall output and impact of the AI application, including its benefits and potential reticence risk, rather than focusing on risk only. Prior consultation with regulators should be limited to the most risky AI uses where risk cannot be mitigated. This would also enable an iterative approach to compliance that takes into account the dynamic character of AI. Similarly, the actual AI requirements would be calibrated by organisations and implemented depending on the risk of the AI application.

(2) Foster innovation through accountable practices of organisations. Rather than imposing prescriptive and indiscriminate requirements, the regulatory approach should set forth a general risk-based accountability requirement and outcomes that organisations should achieve through concrete, demonstrable and verifiable risk-based accountability measures. Organisations would define and tailor these measures on the basis of their own assessment, industry best practices, regulatory guidance, or external technical standards. The EU AI approach should also set up the appropriate "incentives" to stimulate and accelerate accountable practices of organisations and encourage co-regulatory tools such as self-assessments, voluntary labels, certification and codes of conducts. Finally, any AI rules must not duplicate, nor conflict with, the requirements in the GDPR, to avoid legal uncertainty that could have a chilling effect on development and deployment of innovative AI technologies in the EU.

(3) Enable consistent and modern approaches to regulatory oversight on the basis of the current ecosystem of regulators. This means a) keeping the competence of data protection authorities intact when AI involves the processing of personal data and b) setting up an EU governance structure of regulatory hubs composed of AI experts from different regulators to address inconsistencies and deal with cases where the use of AI has a cross-border and/or cross-sectoral impact. This approach should be complemented by a consistent EU level scheme of voluntary codes of conduct, certification and labeling, which should be designed through consultation with stakeholders. Innovative regulatory oversight and constructive engagement among stakeholders should be enabled and promoted, in particular, through data review boards (to consider the impacts of a particular high risk AI use prior to deployment), or regulatory sandboxes (to provide supervised "safe spaces" for building and piloting innovative AI uses in a reiterative manner and with regulatory feedback, to ensure responsible and accountable innovation).

The EU Commission is seeking to propose an EU regulatory approach to artificial intelligence ("**AI**") in early 2021, recognising that data analytics and AI will be instrumental in driving economic growth and addressing a wide range of societal challenges, while at the same time creating new societal risks.

Building on its prior work on organisational accountability and accountable AI in data protection, the Centre for Information Policy Leadership (CIPL)[1] has convened a group of EU and multinational companies, leaders in AI, to provide thought-leadership in defining an appropriate regulatory path forward on AI for the EU.[2] This discussion paper (the "**Paper**") presents CIPL's response to the consultation on the EU Commission's 19 February, 2020, White Paper "On Artificial Intelligence – A European approach to excellence and trust", which sets out several policy options.[3]

The Paper complements CIPL's previous paper on **How the General Data Protection Regulation (GDPR) Regulates AI,**[4] which examines the GDPR provisions that (1) apply to AI systems in the same manner as any other processing of personal data; (2) are particularly relevant in the context of AI; (3) specifically regulate AI; and (4) overlap with the EU High Level Expert Group's ("**HLEG**") principles for trustworthy AI.

Preliminary Section: CIPL Recommendations for Regulating AI

Just like any other EU regulative initiative, the EU approach to regulating AI should enable consistency and harmonisation across the EU, including at the sector level, be drafted with a technology-neutral and future-proof mindset, consider the existing legal framework, avoid creating duplicative obligations, provide for clear rules to resolve potential conflicts of laws, not recreate barriers within the internal market, ensure a level playing field, and lastly take into account the new challenges posed by the COVID-19 crisis and the need for an agile regulatory framework.

Due to the anticipated lengthy nature of the legislative process, it is key that the Commission also considers that by the time the EU AI legislation is effective, AI-related technologies and practices will have evolved and new challenges will have appeared, as well as solutions and new standards.

Finally, to meet the EU Commission's objective of maximising the opportunities and benefits of AI while addressing its risks, and to adapt to the variety and rapidly evolving nature of AI, CIPL believes that the following overarching principles should also apply:

- **Build on the existing legal frameworks and adopt a minimalist approach to regulating AI:** many risks are not AI-specific (AI may simply amplify the risk) and are already addressed by existing laws which provide for baseline structures, requirements, tools and remedies;

- Ensure timely **involvement of all EU stakeholders in discussions around AI regulation**, including ethicists, lawyers, data scientists, engineers, privacy and security experts, computer scientists, epistemologists, statisticians, AI researchers, business leaders and public sector representatives;

- Benchmark **approaches outside the EU, for example the OECD Principles on AI,** to promote globally harmonised and interoperable guidelines on AI where possible;

- Adopt a **principles and outcome-based** regulatory approach that enables **organisational accountability**, leaving it to organisations to operationalise outcomes and principles through risk-based, verifiable and enforceable practices;

- Align with the Commission's approach to **defining a framework for data access and use** by favoring an "*an agile approach to governance that favours experimentation"* over *"heavy-handed ex ante regulation;"* [5]

- Promote **"AI risks/benefits balancing tests" and organisations' contextual impact self-assessments** and limit prior regulator review to the few very high risk cases ;

- Promote and "incentivise" the **development and implementation of best practices** by all stakeholders in the AI ecosystem, with the objective of **continuous improvement** and **risk mitigation**.

- Employ a **co-regulatory approach** to develop the applicable assessment and testing standards for AI systems;

- Encourage innovative, risk-based and collaborative regulatory oversight, by **providing regulators with modern and agile oversight tools.**

On this basis, CIPL recommends a **layered approach to regulating AI (see Appendix 1)**, relying on existing laws and standards, in particular the GDPR, and building on existing accountable practices of organisations, backed by innovative regulatory oversight and instruments. Moreover, CIPL demonstrates in this Paper that in order to properly maximise benefits and minimise risks, the EU AI regulatory approach should **(1)** be grounded in **impact assessment**; **(2)** seek to foster **innovation** through **accountable** practices; and **(3)** enable **consistent and modern** approaches to oversight.

1. A risk based approach grounded in impact assessment

The EU Commission is currently looking to regulate only "high risk" AI applications. High risk AI applications are determined on the basis of several pre-defined criteria: a) a specified sector in combination with the significant impact of the AI, or b) pre-determined, by default high risk AI applications. The AI applications in scope would have to comply with a set of specified requirements and be subject to a prior conformity assessment.

CIPL **welcomes a risk-based approach to regulating AI** – both in defining the applicability of the regulation and in establishing the applicable requirements. However, CIPL believes that the proposed approach is too prescriptive, potentially too wide-ranging and not well-adapted to the variety and rapidly evolving nature of AI-related technologies. In the context of the fight against COVID-19, for example, AI can be extremely useful (for the purposes of contact tracing, understanding and predicting the spread of the virus, allocating spaces in hospitals, etc.). In these kinds of extreme circumstances (which require even more rapid and expeditious measures), an agile framework is indispensable for the development of a European AI market. Prior conformity assessments would be so burdensome as to prevent realisation of the full potential of AI in the EU.

CIPL recommends a similar approach to that of the GDPR, i.e. replacing indiscriminate prior reviews by regulators and burdensome, inefficient and lengthy administrative procedures[6] with a *modus operandi* based on organisational accountability that includes impact assessments by the organisation. This approach would be applicable to AI systems using both personal and non-personal data. This would not prevent the application of current EU law if specific regulatory obligations are already applicable.

The current legal framework for data regulation is divided between **personal data and non-personal data**. Yet, this distinction may not be relevant in the context of risk assessment of AI, as there are situations where personal and non-personal data coexist. There are instances where AI does not process personal data[7] (such as when AI is used for weather forecasting, industrial processes or agricultural planning) and therefore does not trigger application of the GDPR. It should not be assumed, however, that because an AI only uses non-personal data, it cannot have an impact on individual or collective rights, or society at large. For example, using AI to improve the efficiency of industrial processes may have an economic impact on employment and the right to work, as well as on environmental protection. Also, AI systems may be trained with non-personal data and, when deployed, they may use or create personal data in order to make decisions about individuals. Finally, the determination of what is not personal data or when information may be considered anonymous is still an unresolved question, as demonstrated by the current debate over contact tracing apps in the COVID-19 context. Organisations should have the option to adopt a holistic approach to risk assessment in the context of AI, regardless of the type of data used.

Instead of relying on pre-defined criteria for high risk AI applications and prior conformity assessments, CIPL proposes a more flexible approach, grounded in impact assessment of AI applications that would (1) be based on context and real impact of a proposed use of AI, rather than the sector it is utilised in or its type; (2) rely on illustrations of rebuttable presumptions of high-risk, rather than rigid classifications; (3) assess the overall impact of the AI application, rather than focusing only on risk; (4) enable an iterative approach to compliance in AI; and (5) allow for the identification of appropriate accountability measures depending on risk. This approach would be applicable to all potential risks of AI (and not limited to risks relating to privacy and security).

1.1 Rely on an assessment of the AI's context and impact rather than classifying sectors or types of AI as inherently high risk

CIPL believes that the **sector-based** approach to defining high-risk AI applications is too rigid[8] and may be discriminatory - any sector may use AI applications with varying degrees of risk. Also, creating blacklists and highlighting certain sectors as more "risky" implies that AI has an inherent negative impact, which is not accurate and not helpful when the EU has to further build its AI capabilities and trust in AI technology. Furthermore, in the digital economy, sectors are not static or clearly delineated. They continuously evolve with new business models and technologies disrupting traditional sectors. In addition, data, products and technology move across sectors, and the same algorithm can be used in different sectors with different impacts depending on context. For example, an AI-based application in an autonomous vehicle that measures the behavior and health of a driver (temperature, focus, alertness) is produced by a manufacturing and software company, used by the automotive industry, and can also be used by the health industry or public health authorities to fight a pandemic. Natural Language Processing algorithms can be used to develop a chatbot for 24/7 customer support or detect hate speech against minorities on the Internet, but also to survey populations and take actions against political dissidents. Another example is when a company in a certain sector may have in its portfolio products originating from another sector:

an energy company or a travel company may sell insurance products (home, payment protection, travel insurance) or a telecommunications company may sell entertainment services (film rental, audio books). If AI applications are embedded in these products or services, why would insurance products sold by an energy company be more risky than insurance products sold by an insurance company? Similarly, why would entertainment services sold by a telecommunications company be more risky than the same products sold by an entertainment company?

Furthermore, a focus on sector or types of AI does not sufficiently appreciate the **continuous evolving nature** of AI applications. For example, an AI application that may initially be considered a low risk may, through interaction with other applications with new risks, transform into an unanticipated higher risk application. For instance, a computer vision algorithm initially designed to classify documents for their processing by a human, could be installed in an autonomous vehicle to "read" traffic signals and cause accidents if not sufficiently accurate. Conversely, as the technology improves, an AI application that is high risk might become less risky if e.g. accuracy is improved. In sum, the sector-based approach to defining risk will be both over-inclusive and under-inclusive at the same time, capturing risks where they don't exist and omitting others that do exist, thereby defeating its purpose while not adding any specific value to the risk analysis.

Lastly, classifying certain sectors and categories of AI applications as **automatically high risk** may undermine the overall objective of maximising benefits of AI while reducing risks. This approach does not allow for weighing of benefits of AI versus risks and may result in eliminating certain data uses at the outset that carry significantly higher benefits than the perceived high risk (see Section 1.3).

Therefore, the Commission should rely on the **impact of an AI application in a given context**, on a case-by-case basis, rather than focusing on a specific sector or pre-defined use. Certain sectors are already highly regulated given their level of risk (healthcare, energy, etc.), and it would be worth combining this vertical approach with the impact of an AI application in a particular use context. In practical terms this means examining the decision the AI is making or assisting humans with (output), rather than looking only at the underlying algorithm (technology) and training data (input). In other words, the underlying issue is more the impact of the decision rather than the input or the technology itself. An AI system could be trained to comply with all requirements, but if it is used for nefarious purposes, or if the output is used for purposes other than those for which the AI was trained, this may pose risk to individuals and society. This requires moving the focus from the **"AI application" to the "AI decision or AI use",[9] its individual or societal benefits and the likelihood and severity of individual or societal risks**. A focus only on regulating the input of an AI system risks missing vital aspects of an AI's use that could prove far more damaging. For example, if a desired outcome is for an AI to take a non-discriminatory decision, this outcome depends both on the dataset used to train the AI, and how its output is used. As another example, in the energy sector, AI is used to automate energy trading decisions. Millions of transactions take place in seconds. In this situation, the focus should be on building control layers around the AI or in the marketplace that would actually prevent execution of decisions made by the AI that are not in accordance with the trading or marketplace rules.

It should be noted that the **GDPR already requires organisations using AI to undertake a Data Protection Impact Assessment (DPIA)** to identify and address processing of data that is "likely to result in a high risk to the rights and freedoms of natural persons".[10] This exercise has a broad scope. It extends beyond assessing the impact of the processing on data protection and privacy rights to examining the AI's

potential impact on all the rights and freedoms of natural persons as provided for by the EU Charter of Fundamental Rights.[11] This may even include potential broader societal impacts.[12] This type of impact assessment can therefore be easily **expanded to encompass an appraisal of the potential impact of the use of an AI application**, including when non-personal data is involved.

In addition, Article 22 of the GDPR already deals with the notion of impact and high risk in the specific context of AI and Automated Decision Making (ADM). It applies a stricter legal regime for solely automated processing, but only where such processing results in a **decision that produces legal or similarly significant effects for an individual.**[13] This may be a useful criterion in assessing the type and level of risk in an AI impact assessment and may be easier for organisations to operationalise as they are already required to apply it under the GDPR. When personal data is involved, this appears to already meet the Commission's proposed threshold on high risk AI to "produce legal or similarly significant effects for the rights of an individual".

Finally, an analysis of the impact of an AI application should balance the human or societal impact of the AI's use (whether positive or negative) with the AI user's interest, taking into account the **severity of the potential individual/societal harm, the likelihood** that the harm will occur, and the **actions that have been taken to mitigate the severity of harm** to what was deemed an acceptable level by the AI user. It is important that in all cases, the analysis takes into account what is probable (i.e. a risk's likelihood of materialising), as opposed to what is merely possible, or theoretical. Any actions taken to mitigate the potential harm should be proportional to the likelihood of the event happening and the harm that may occur. For instance, if the anticipated individual harm is low, the likelihood of the harm happening is high, and the societal benefit is high, it may be reasonable for a company to perform a given AI use-case without taking mitigating actions. However, if there is a high and likely risk of harm to individuals, mitigating actions may become necessary.

For example, in the COVID-19 response, using aggregated location data and AI technologies to formulate government responses can be extremely useful to fight the pandemic. Such use may be deemed acceptable despite individuals' privacy fears where the net benefit exceeds the potential privacy risks, and these risks are mitigated by the use of aggregated data. The situation would differ if instead of aggregated data, non-aggregated data were used, as that would require different risk mitigations to ensure proper balance between the intrusion into data privacy and the benefits to public health and potential for saving lives.

1.2 Provide rebuttable presumptions rather than rigid criteria for identification of high risk

CIPL has previously recommended, with regard to assessment of high risk under the GDPR, that a good approach would be to provide illustrative presumptions of high risks, allowing organisations to rebut such presumptions through risk assessments and/or appropriate mitigations.[14] CIPL believes that such an approach is equally appropriate in the context of AI. Providing suggestive criteria, examples or presumptions of high risk would be of more practical use to those developing and using AI – including SMEs - than rigid ex-ante lists of high risk applications, as this is more suited to the highly contextual and evolving character of AI. Indicative lists **of presumptive high risk AI** use could then be rebutted on the basis of the specific context in which the AI is used, or on the basis of its counterbalancing benefits (see section 1.3). The input of industry, academia and civil society would be key in determining these criteria or presumptions and updating them on a regular basis. The "rebuttable presumption approach" permits

organisations to undertake their own "risk triage" and risk assessments (whether on the basis of their own methodologies and tools or on the basis of public resources – see Section 2.5), and make their own judgements regarding such risks based on the knowledge of how they intend to use the AI, the safeguards that are in place, and the potential benefits of its use.

For example, AI may be used for simple automation purposes, i.e. to streamline an already existing process, which is not high risk. In that case, organisations would be able to act more strategically and efficiently by **easily identifying low risk AI** uses without undertaking time-consuming and costly risk assessments. This could take the form of a "quick scan risk assessment" or pre-screening to determine whether a full scale impact assessment is necessary. This would allow organisations to better allocate their resources to AI applications that may carry a high risk and find ways to mitigate that risk. This would prevent organisations from undertaking assessments of AI use in contexts where it is clear from the outset that there is no high risk involved. Organisations should not have to go through the process of proving that the AI use is low risk in cases where that is obvious.

Finally, inspiration may be taken from the GDPR's approach with regard to **prior consultation**, which is used as a last resort and for a minimal number of cases. Indeed, the GDPR consciously abandoned the ex-ante external "prior checking" system that existed under Directive 1995/46/EC and replaced it with the duty to implement accountable data protection programmes, which includes DPIA for high risk situations. Under the GDPR, organisations are now required in the first instance to conduct a DPIA where the processing activity is likely to result in a high risk to the rights and freedoms of individuals. This assessment identifies the relevant risks and the measures necessary to mitigate those risks. Only where organisations cannot mitigate a high risk to individuals revealed by a DPIA, must they then consult with and seek the approval of the relevant data protection authority (DPA) before the processing activity in question can commence. The same approach may be considered with regard to AI, where prior consultation would be required only where a presumption of high risk cannot be rebutted and mitigated. In practice, such prior consultations may apply only to very high risk situations such as the use of facial recognition for unique identification purposes, or public sector uses of AI in policing or in the justice system. CIPL believes this is a reasonable approach that would also alleviate the pressure on regulators, who are already struggling for resources and time with regard to performing their statutory roles in the modern digital economy.

1.3 Holistic Impact Assessments: Assess risks as well as benefits and reticence risks

Risk assessments must reflect the overall impacts and objectives of the use of AI and consider other relevant factors beyond just whether an application creates risks and harms for individuals and society. CIPL recommends that any "AI impact assessment" must include **weighing the concrete benefits** and the stated objectives of deploying the AI against the risks it creates. High risks related to an AI system may be overridden by compelling benefits to individuals, organisations and society at large. Examples of such compelling benefits include the use of predictive machine learning in the health sector that enables more effective detection, accessibility to medical services, diagnosis and provision of health services or delivery of tailored, precision medicine. It also includes using AI to monitor online content in order to prevent terrorism, child abuse or other illegal behavior, hate speech or the spread of harmful content, which could outweigh the risk associated with the system's use. This approach is already included in the GDPR, which provides for a weighing of legitimate interests and benefits against intrusion on the rights and interests of data subjects as part of the legitimate interest balancing test.[15]

Similarly, the assessment must also include consideration of the relevant **"reticence risks**" or the risk of lost opportunity, i.e. what benefit would we lose out on in the absence of the proposed processing or what would be the consequences for individuals and society of not using an AI system because of its potential risks?[16]

CIPL believes that the GDPR can and should be leveraged as a basis for designing **a broader form of AI impact and risk assessment** that can encompass the specific risks and benefits created by AI and take social context into account. Some organisations are already creating these broader assessments (see **Appendix 3**). In any case, the Commission should encourage the design of such AI impact assessments and facilitate the gathering of views and best practices of various stakeholders.

For this approach to work in practice, CIPL reiterates the importance of not classifying certain sectors and categories of AI applications as automatically high risk, as this may eliminate certain beneficial data uses at the outset (see Section 1.1).

1.4 Facilitate an iterative approach to AI compliance

AI is amorphous and constantly changing, with new risks and benefits appearing all the time. Some of these risks and benefits may be impossible to predict at the outset. Any regulatory approach must capture this dynamic character and **allow for constant adaptation, evolution and improvement**, fixing issues as they appear, rather than demanding perfection from organisations at the outset. More focus should therefore be put on organisations' ability to identify and address risks continuously, as they appear and change, for instance, by updating remediation processes, determining and adjusting the need for and extent of human oversight safeguards, performing regular timely checks in the post-deployment phase, implementing specific technical safeguards or providing redress to potentially affected individuals. It enables responsible organisations to experiment, learn and grow as they develop best practices for implementing innovative AI technologies, rather than focus on whether or not that use of AI falls within a specific pre-defined category.

This approach also **avoids unnecessary risk aversion**, i.e., not moving forward for fear of not having a perfect system, where all possible risks have been identified and are fully mitigated before the product launch, from the start.

It also **encourages organisations to be accountable** - to be even more vigilant with regard to the actual use and functioning of the AI system, to remain alert to new risks that may appear, and to require updated or new controls and mitigations to be implemented at appropriate intervals, including in a time of crisis. Extending the organisation's accountability throughout the life cycle of an AI application increases business sustainability, responsible innovation and overall corporate digital responsibility.

This approach also reflects the way organisations are already operating with regard to compliance with the GDPR (and other corporate compliance areas, such as anti-bribery, anti-money laundering, and export controls). As further demonstrated in CIPL's recent report "**What Good and Effective Data Privacy Accountability Looks Like: Mapping Organisations' Practices to the CIPL Accountability Framework**",[17] internal **accountability programs are ongoing endeavors driven by continuous risk assessment and improvement**, to correct errors, improve and adapt to external factors such as legal changes, evolving

technology and risks. Applying the same compliance approach to AI systems is a natural next step for many organisations.

Finally, as we stated in CIPL's paper "Regulating for Results – Strategies and Priorities for Leadership and Engagement",[18] we believe that regulators should also approach compliance, oversight and even enforcement as **a dynamic process and journey, allowing bona fide trial, honest error and constant improvement**. This is necessary in the modern digital economy, in order to align with changing business practices and the reality of fast evolving technologies, such as AI, that are often in tension with static legal norms.

1.5 Enable organisations to tailor requirements and accountability measures to the risks identified

The Commission's current approach provides for several requirements to apply to high risk AI applications relating to training data, record-keeping, provision of information, robustness and accuracy and human oversight. The White Paper provides for "necessary measures", covering "all relevant scenarios", and avoiding "dangerous situations". CIPL believes that this approach is too prescriptive, as these requirements may not be relevant in all circumstances. For instance, the record keeping obligation on the data set used to train and test the AI systems may be challenging to address. Because of the dynamic nature of AI, this would entail having to perform system "snapshots" in order understand how and why AI made certain decisions. In addition, in some cases, even lower risk AI applications have to apply some of these best practices and requirements to build trust with customers and ensure fairness or safety of an AI application.

CIPL suggests a more practical approach that allows organisations to **determine, apply and calibrate the requirements that should apply to each specific AI system.** The approach should be based on the system itself and on the results of the above-mentioned risk assessment (as the risks that are identified and their potential impacts will differ from one AI application to the next). As a general rule, the fewer risks that are identified (or the less severe or likely such risks are, or the better they have been mitigated), the fewer the number of requirements that should be imposed on the system, or the less onerous those requirements should be.

For instance, the degree of **human oversight** should depend on the impact on individuals and the outcome that is sought. The timing of human oversight should depend on the specific AI system and the phase (e.g., research or deployment). In certain situations, it may be perfectly reasonable to have a machine with relatively limited oversight operate where impacts are low or where testing has revealed high rates of certainty in outcomes (e.g. an AI automating a routine manual operation).

Similarly, aiming for maximum **accuracy** of an AI system is not always appropriate. In the financial sector, for instance, reducing accuracy on purpose in order to increase false positives, and thus flag more payment transactions as potentially risky, allows for extra due diligence steps in assessing fraud. In other words, taking specific risks (i.e. by reducing accuracy) can be a conscious trade-off by the organisation to achieve an important benefit (combatting fraud). Consequently, in the case of trade-offs, requirements should be adapted.

The same might apply with **transparency**. Higher risk AI systems might require enhanced and user-centric transparency, as well as additional transparency to internal or external oversight committees and/or

regulators. The Project expl*AI*n conducted by the Alan Turing Institute and the ICO,[19] illustrates well that the explanations of AI decisions should be tailored to the specific audience based on relevant contextual factors of the AI decision at hand, such as: the urgency of the decision, the impact of the decision, the ability to change the factors influencing the decision, the scope for bias in the decision, the scope for interpretation in the decision-making process and the type of data used in the decision-making process. For instance, a survey of citizen juries empirically demonstrated that individuals facing AI healthcare scenarios cared more about accuracy than transparency, while transparency expectations were heightened for the use of AI in job recruitment and criminal justice scenarios. In addition, there are certain benefits deriving from highly accurate AI systems that, in particular contexts, may justify different levels of explainability. This is also discussed in CIPL's **Second AI Report – Hard Issues and Practical Solutions.**[20] For example, the resources that would be required to interrogate AI models for the purposes of transparency (i.e. to understand how they work and thereby make them explainable to data subjects) may be disproportionate to the benefit of doing so, since by its nature AI is often designed to deal with tasks too complex for individuals to complete.

It should also be noted that the requirements outlined in the White Paper should not be taken as necessarily having a remediating effect on AI risks in all instances – there may be some AI systems, such as those conducting particularly complex mathematical processes, for which increased human oversight would be more of a hindrance than a help. Organisations have to **consider such trade-offs on a case-by-case basis**, as the right combination of human oversight and trust in the algorithm of the machine will vary dramatically based on the particular application.

This **sliding scale approach of having AI requirements apply differently depending on context** is more flexible and more reflective of the nature of AI systems than a rigid no requirement/all requirements dichotomy. It also allows for adaptation of the requirements that are applied as the AI application itself evolves. Further, it permits organisations to prioritise and allocate their resources to those activities that present the highest risk and necessitate the most requirements, rather than being burdened with the same cumbersome and ill-adapted set of requirements in relation to each AI application.

Summary of CIPL Recommendations

- Rely on **impact assessments performed by organisations** to (1) **trigger the application of the law** and (2) allow for the **adaptation of accountability measures** depending on risk.
- AI impact assessments should:
  - Be based on **context and real impact** of a proposed use of AI, rather than the sector it is utilised in or its type;
  - Rely on illustrations of **rebuttable presumptions of high-risk**, rather than rigid classifications;
  - Assess the **overall impact of the AI application,** including benefits and reticence risks, rather than focusing only on risk;
  - Keep prior consultation with regulators limited **to the most risky AI uses** where risk cannot be mitigated; and
  - Enable an **iterative approach to compliance in AI**.

2. An AI Approach that fosters innovation through accountable practices

2.1 Avoid duplicating GDPR provisions for personal data processing in AI

As already explained in the paper "How the GDPR Regulates AI", when AI relies on the processing of personal data, it is already regulated by both the general and AI-specific provisions of the GDPR.[21] Therefore, the AI approach described in the White Paper should seek to **regulate only those areas where the GDPR is silent** (such as, for instance, the processing of non-personal data that do not coexist with personal data), if relevant and appropriate.

Thus, the AI approach should clarify that **any processing of personal data in the AI context remains fully subject to the GDPR** to avoid confusion with respect to the relationship between the GDPR and any AI regulation. It is important to avoid anything similar to the current overlap (and inconsistencies) between the GDPR and the e-Privacy Directive that is creating huge confusion in the market. A similar lack of legal certainty in the context of AI would not be sustainable and would most likely have a chilling effect on new and innovative AI-related EU projects.

The table in **Appendix 4** shows the overlap between the White Paper's suggested requirements for high risk AI and the existing requirements of the GDPR.

2.2 Consider outcomes organisations should seek to achieve rather than prescriptive requirements

CIPL's view is that the proposed requirements for high risk AI are drafted **too prescriptively** to enable flexible and future-proof regulation. They should instead be recast as outcomes for organisations, leaving it to them to define **concrete and verifiable measures** and processes to put in place to achieve these outcomes. This allows for more contextual adaptation, taking into account the AI system itself as well as the specifics of the organisation. CIPL believes that this approach can provide more effective protections for individuals and society in the context of AI, while driving the benefits of AI technologies and building EU AI capabilities, especially with SMEs.

Again, we can look to the GDPR for inspiration, since organisations are required to comply with several data protection principles, such as fairness, transparency and data security under article 5(1). For the most part, the GDPR is not prescriptive as to how to comply with these principles, or how organisations should demonstrate compliance as required under Article 5(2). The accountability principle set out under Article 24(1) requires controllers to *"[…] implement appropriate technical and organisational measures to ensure and to be able to demonstrate that processing is performed in accordance with this Regulation. Those measures shall be reviewed and updated where necessary*".

If these outcomes are not achieved (e.g. the processing leads to unfairness) individuals may, under the GDPR, raise a complaint and demand redress before DPAs and the national courts.

The table below demonstrates how the suggested requirements of the AI White Paper can be turned into outcomes that organisations would be required to achieve based on context and effective risk assessments.

Organisations can convert outcomes into requirements based on their own assessment, industry best practices, regulatory guidance, or external technical standards (such as ISO norms). In the context of AI, organisations could rely on the HLEG Trustworthy AI self-assessment list and the upcoming self-assessment tool to convert the outcomes into requirements.

This combined approach of providing outcomes in the law and having organisations operationalise them on the basis of external standards that become de facto mandatory within the organisation (and as the case may be through the adoption of delegated and implementation acts), that can be adapted if necessary, is very well adapted to the fast evolving nature of AI.

AI White Paper Requirements	Outcomes to be implemented by organisations on the basis of risk
Requirements applicable to training data	Organisations shall ensure that AI is developed according to the principle of **fairness to protect against prohibited discrimination**
Requirements for keeping records and data	Organisations shall implement technical and organisational measures to **demonstrable compliance with the regulation**
Requirements for information provision	Organisations shall implement technical and organisational measures for ensuring actionable and targeted **transparency of AI** systems relevant to the intended recipients.
Requirements for robustness and accuracy	Organisations shall implement technical and organisational measures to ensure the **robustness, accuracy and security** of the AI system based on the risk they present.
Requirements for human oversight	Organisations shall implement technical and organisational measures to provide for **human oversight of AI** where appropriate, including human review of automated AI decisions with legal effects.

2.3 Position accountability as a key element of the AI regulatory framework

Accountability requires organisations to be thoughtful about the risks and impacts of their processing activities on individuals and establish processes and controls to anticipate and address these in compliance with the law. Typically, accountability is implemented through **comprehensive data protection compliance and management programs**. The concept of organisational accountability has become a common feature of privacy and data protection regulation in the EU and globally. It has also been deployed in many other compliance areas, such as anti-bribery, anti-money laundering, export control, medicine and food regulation.[22] Accountability-based compliance and governance programs enable organisations to operationalise principles-based laws into risk-based, verifiable, demonstrable and enforceable corporate practices and controls, supported by technology tools. This enables organisations to be responsible data stewards in the AI context by assessing the potential impacts of a given application, implementing procedures to ensure accountability and to continuously improve and adapt to change.

Organisational practices rooted in such accountability-based compliance programs benefit individuals by delivering real, relevant and effective privacy protections on the basis of legal requirements. They also help organisations demonstrate legal compliance to EU regulators, business partners and individuals. This results in increased trust by these constituencies in organisations' development, deployment and use of

technologies. This also enables the development and use of "accountable AI". Indeed, because accountability requires that organisations document their risk assessments and decision-making processes and be able to demonstrate them to a regulator on request, **they are beneficial to the regulator** in reducing the burden on their already-stretched resources. CIPL has developed an accountability framework to help organisations design, structure, build and implement their data protection management programs based on the key elements of accountability, which also align with all relevant GDPR requirements (Leadership and Oversight, Risk Assessment, Policies and Procedures, Transparency, Training and Awareness, Audit and Monitoring and Response and Enforcement). This framework is available in **Appendix 2**.

CIPL believes that an "accountability approach" is well suited to the context of AI because of its effectiveness and flexibility. As CIPL noted in its second report on **Delivering Sustainable AI Accountability in Practice**,[23] many organisations are proactively starting to use accountability frameworks to address the risks, challenges and tensions presented by the use of AI and to comply with relevant laws, including data protection and anti-discrimination laws, as well as proactively considering social expectations. AI products are already being designed or deployed with these considerations in mind -- not only to comply with legal requirements but also because organisations want to ensure customer trust.

These **emerging best practices** are starting to take shape in the form of coherent and comprehensive accountable AI frameworks, including those based on the CIPL Accountability Wheel as demonstrated in the table in **Appendix 3**. It outlines examples of accountable AI activities undertaken by selected organisations from different sectors, locations and sizes that CIPL has collected. Most importantly, these measures are not exhaustive or stated in order of importance – they are simply examples of appropriate measures that organisations of all sectors and sizes might consider within the context of their specific AI applications. In addition, organisations should be able to integrate their AI accountability programs into their existing compliance and accountability programs, rather than being required to create AI-specific standalone programs.

CIPL's recent report "What Good and Effective Data Privacy Accountability Looks Like: Mapping Organisations' Practices to the CIPL Accountability Framework"[24] confirms that this approach is also **scalable to SMEs**. In practice, small organisations tend to calibrate accountability measures differently than larger, multinational organisations, and are often able to do so with more agility. It is important that in the AI context, SMEs put in place necessary processes and controls as the market, business partners, and users more generally, will expect that they do so and will ask for proper assurances.

In line with this and with the HLEG recommendations,[25] CIPL believes that the EU approach on AI should **explicitly incorporate risk-based accountability as a core principle** and proposes the following wording:

> *"Taking into account the nature, scope, context, purposes, impact and benefits of an AI application, the organisation shall implement, and be able to demonstrate that it has implemented, appropriate organisational and technical measures to enable that the AI application complies with the principles in the AI Regulation. Such measures shall be reviewed and updated where necessary."*

2.4 Encourage and accelerate AI accountable practices

In order to further promote accountable practices in the AI context and to allow best practices to continue to emerge and adapt across the digital economy, the EU AI approach should set up a framework that encourages and accelerates accountable practices of organisations. The regulation **should give tangible "incentives", either directly or through regulators, to organisations that develop and implement accountable AI best practices and internal responsible AI frameworks**. Such encouragement should be available both in the public and private sector and could include for instance:[26]

- Promoting organisational accountability though Digital Innovation Hubs;[27]
- Using demonstrated accountability as a form of "licence to operate", by giving accountable organisations **greater freedom to use and share data responsibly**, which could facilitate further growth in responsible AI, machine learning, and ADM;
- Allowing broader use of data in **AI for social good** projects, public health and research, secured by relevant risk assessment, mitigations, oversight, and controls in accountability programs;
- Publicly recognising **best-in-class AI organisations** in the EU and showcasing accountable "best practices" in the AI field (such as AI design rules that have assertion-evidence and can be audited);
- Offering accountable organisations the opportunity to **play a proactive advisory role** for organisations seeking to implement accountable AI practices, especially SMEs (e.g. through sharing of competence through Digital Innovation Hubs);
- Making available and encouraging participation in relevant **codes of practice, codes of conduct and certifications**;
- Using demonstrated AI accountability as a criteria for **public procurement projects**; and
- Using demonstrated AI accountability as **a mitigating factor or as a liability reduction or exclusion** factor in the enforcement context.

2.5 Encourage co-regulatory tools: self-assessments, voluntary labels, certification and codes of conducts

CIPL suggests that **self-assessments, including potential review by internal or external AI/data advisory boards,** would be better suited in the context of high risk AI than rigid and time-consuming prior conformity assessments.

The HLEG **Trustworthy AI self-assessment list** (currently under review) and the Canadian Algorithmic Impact Assessment tool,[28] as well as the CIPL accountability framework, could serve as good starting points for these self-assessments. Additional assessment toolkits and schemes could be set up through a co-regulatory approach and be updated regularly based on technological developments and new practices. These lists and toolkits should provide suggested considerations and control points to reach outcomes on the basis of risk, such as traceability, explainability and resilience.[29] They would be

particularly useful to SMEs by helping them assess and mitigate the risks associated with the deployment and use of AI. Enabling organisations to rely on such tools would also help address the concern of legal uncertainty in the AI legal framework (due to AI's highly contextual character, as explained above) in particular for **SMEs**.

They could be **complemented by voluntary labelling or certification schemes,** as suggested by the White Paper, to enable organisations to better signal that an AI application **meets certain criteria that have been assessed and confirmed by an independent body**, that would make them binding on the organisation.

CIPL also suggests that **codes of practice and codes of conduct** should be further considered as possible co-regulatory tools that would enable organisations to define the relevant processes and demonstrate, on a voluntary basis, that their AI product, uses or related activities comply with certain standards. Smaller organisations may benefit from joining pre-existing codes of conduct that already specify the measures to be implemented by code members. This would help accelerate compliance of their AI products or uses. This would also help organisations demonstrate accountability and would contribute to building trust in the AI ecosystem.

All these schemes may be especially relevant for **consumer facing AI applications** as they would enable individuals to easily identify certain AI features and have trust when using AI applications. They would also be very welcome in B2B ecosystems, to **enable quicker due diligence**, faster speed to market and drive trust in business partners. They may become a market standard or a licence to operate in the AI ecosystem. Market and peer competitive pressure may end up acting as a catalyst for proactive adoption of labels or certifications (instead of imposing burdensome, one-size-fits-all prior conformity assessments). These schemes would also allow organisations to certify to a set of requirements that are specific to their sector, technologies or to requirements that go beyond what is strictly required by law.

The EU commission has an important role to play in promoting and facilitating these initiatives by arranging for a **publicly available repository of labels, certifications, codes of practice and codes of conduct related to AI**. It could also encourage best-in-class organisations to make their AI self-assessments publicly available on a voluntary basis to help the entire ecosystem.

Summary of CIPL Recommendations

- Regulate only those areas where **the GDPR is silent**;
- Clarify that **any processing of personal data in the AI context remains fully subject to the GDPR;**
- Replace requirements **with outcomes,** leaving it to organisations to define **concrete and verifiable measures** to achieve these outcomes;
- Explicitly incorporate **risk-based accountability as a core principle** of the EU AI approach;
- Provide **incentives to organisations that develop and implement accountable AI best practices** and internal responsible AI frameworks;
- Rely on **self-assessments** rather than prior conformity assessments;
- Encourage **voluntary labels, certification and codes of conducts**;
- Provide for a **public repository** of AI labels, certifications and codes of conducts.

3. An AI approach that enables consistent and modern regulatory oversight

The EU Commission White Paper's objective is to avoid fragmentation while relying on competent national authorities. It proposes a European governance structure that would have a variety of tasks (forum, facilitator, exchange of best practices, issuing guidance, advising on standardisation and certification). A network of national, European and sectorial authorities, involving participation from all stakeholders (consumer, civil society, industry, academic, research) would be set up to complement existing expertise and assist in monitoring activities involving AI (without affecting the powers and responsibilities of these authorities in their specific sectors). Testing and conformity assessments would be entrusted to notified bodies designated by Member States.

3.1 Enable effective and agile cooperation among regulators

AI is often at the intersection of several disciplines governed by different regulations and overseen by separate regulators. DPAs have general competence over the processing of personal data using AI, while other regulators have a more sector-specific remit (healthcare, banking and financial, telecoms, consumer, pharmaceuticals, etc.) regardless of whether personal data is used.

CIPL welcomes the Commission's proposed approach of **relying as much as possible on existing authorities and not creating an additional layer of AI-specific regulators** or agencies. This should also include relying on the current avenues for **redress and legal recourse at the national level** in case

individuals are negatively impacted by an AI decision, and not creating new schemes on top of existing ones.

CIPL agrees that **expertise on AI within existing authorities must be expanded, rather than new structures being created.** CIPL also cautions against the possible creation of a highly fragmented approach at the member state level and at the sector level, if effective consistency mechanisms are not set up and respected.

This may be particularly **impactful for organisations** that look for uniform guidance and a level playing field across the EU to invest in AI technologies. Facing a multitude of national approaches as well as a number of different oversight and enforcement bodies will heavily impact the way organisations create policies and compliance for AI technologies. It will make it more difficult for EU organisations, including SMEs, to implement European-wide programmes, launch products, attract new pan-European customers, and apply consistent risk assessments of AI on individuals or society, or work with business partners. It would result in the recreation of barriers to the internal market and negatively affect the EU AI and data strategies. It could also create incentives for regulatory arbitrage as, depending on the national and/or sectoral authority/ies that supervise(s) a given activity, the regulatory requirements could differ widely.

In this context, much can be **learned from the GDPR** which, although originally created as one law for the EU, has resulted in a fragmented set of rules during implementation. First, the GDPR contains too many **open clauses** for Members States to deviate from key provisions which has resulted in "legitimised" differentiated implementation across the EU. Second, national regulators have been issuing too much **duplicative and conflicting guidance** on key GDPR concepts, leaving organisations faced with huge compliance and operationalisation hurdles. At the same time, the European Data Protection Board (EDPB) **has not stepped up sufficiently to standardise approaches**. For instance, the 28 different lists of high risk processing requiring a DPIA that have been produced by EU DPAs render the obligation to conduct a DPIA very cumbersome, if not impossible for organisations with cross-border operations. As a result some organisations tend to avoid having to run DPIAs, sometimes even by abandoning complex cross-border business projects rather than taking on a burdensome and lengthy process with unpredictable costs and uncertain outcomes. Unfortunately, the EDPB has not produced a single authoritative list but limited itself to reviewing each national list **without ensuring full harmonisation** between the different national approaches. A similar situation would not be sustainable when it comes to identifying potentially high risk AI. Finally, **the one-stop-shop (OSS) mechanism**, originally intended to foster consistency and coordination between different DPAs and to enable organisations to work with one single regulator/interlocutor for their cross-border cases, does not function optimally and ends up creating more bureaucracy at the expense of efficiency and legal certainty.

At the same time, CIPL believes that the current regulatory framework of the GDPR should remain fully applicable **whenever AI processes personal data,** and that DPAs (or the Lead DPA, where applicable) and the EDPB should continue to be competent in these cases.[30]

In all other cases with an EU cross-border impact, where a particular AI application also has an impact at the sectoral level, or where AI does not process personal data at all, CIPL would recommend the **setting up of an EU governance structure in charge of dealing with these cases.** This would not affect (1) the competence of local regulators; (2) the existing distribution of competencies in different sectors between the local and national levels (see for instance in the antitrust, telecom, healthcare or banking sectors); or

(3) the existing collaboration schemes via memoranda of understanding that may already be in place in some countries to address potential cases where several regulators may be competent over a specific AI application.

CIPL does not believe that the EU AI governance structure should consist of the creation of an overarching AI regulator. Instead, CIPL recommends that this structure be made up of **AI experts sitting with the different regulators** that would cooperate through **regular or ad hoc regulatory hubs for AI**. Each regulator would keep its own competence, but could exchange views and knowledge, align interpretation, resolve any conflicts of law and participate in joint initiatives, such as joint sandboxes together with notified bodies and centres of excellence (see section 3.3).

The conditions for the operational success of the AI governance structure are as follows:

- Creating **clear and consistent definitions** of the different situations and scenarios where questions have to be escalated to the EU AI governance structure;

- Establishing **regulatory hubs** composed of AI experts from different regulators (upskilling of the current workforce and recruitment of AI experts within the different regulators must be promoted) as well as creating centres of excellence;

- Setting up a simple and transparent system for organisations, providing for a **clear allocation of roles** and ensuring that organisations can benefit from interacting with a single interlocutor (CIPL would propose that the DPA/EDPB remain the sole interlocutor when the AI application processes personal data);

- Implementing an **efficient decision-making process** that is **time effective**;

- Clarifying that guidance issued by the EU AI governance structure should govern, and **discouraging the issuance of conflicting guidance** at the local or sectoral level (or making sure they are reviewed at the EU level);

- Instilling a collaborative approach between regulators and stakeholders that are developing, deploying or using AI.

3.2 Build a robust and consistent voluntary certification and labeling model

As noted, CIPL believes that any certification or labelling model put in place should not be compulsory and should be primarily based on self-assessment, which could lead to a voluntary but enforceable certification (see section 2).To avoid fragmentation of the EU market, certification or labelling schemes should be designed at the EU level. CIPL proposes that the implementation of such EU certification and labeling schemes be entrusted to national notified bodies designated by Member States in accordance with Regulation 765/2008. In addition, CIPL highlights that successfully setting up a **certification/labelling framework** in the AI context requires taking into account the following considerations:

- Certification schemes, referentials, standards, testing methodologies, and criteria for testing must be **defined in a uniform manner at the European level** and have **EU-wide validity** to ensure consistency and not recreate barriers to the internal market;

- The certification framework must be based on **a co-regulatory model**, with certification, labelling schemes and standards defined with input from industry;

- Certifications must be pragmatic and **leverage existing certifications** and frameworks, such as ISO standards in the AI field;[31]

- Certifications should focus on **AI controls and safeguards**, and certifications of AI Accountability frameworks, rather than certification of specific AI tools and features;

- Certifications and the certification ecosystem (i.e. certification bodies and national accreditation bodies) must enable **maximum interoperability** with non-EU certification schemes and labels;

- The certification framework must enable national certification bodies/national notified bodies to **offer certification services across the EU and allow cross-recognition of certifications** across the EU regardless of the country of certification;

- The certification framework must ensure that any certification/label issued in the field of AI is **achievable in a reasonable timeframe and in a realistic and pragmatic manner**; and

- There should **not be specific enforcement and sanctions for certification bodies related to AI certifications**, as this approach appears to have discouraged uptake of certification business models under GDPR (certification bodies are liable for breaches of GDPR by the company they have certified with the same high sanctions).

A code of practice or a code of conduct may also be relevant when coexisting and potentially conflicting regulations require a forward-thinking or holistic approach. They should be reviewed at the level of the EU AI governance structure and in agile and reiterative procedures. Codes of practice and codes of conduct could also be designed through co-regulation, and would have the advantage of being quicker to establish than certifications. In fact, they could even be the basis for certifications, helping expedite the latter.

3.3 Promote innovative regulatory oversight

Finally, in order to support the development and uptake of AI in the EU, while ensuring that individuals and society benefit from the potential of AI and addressing risks, policy and lawmakers need to encourage new and agile approaches to regulatory oversight. These include encouraging voluntary certification and establishing regulatory hubs, as explored above, as well as other tools that promote constructive engagement and collaboration among stakeholders, such as data review boards and regulatory sandboxes.

Data review boards (DRBs) are helpful for both public and private-sector organisations as they consider the impacts of a particular use of data prior to development, deployment or use.[32] These standing

committees convene when triggered by a DPIA, or other indicators, to promote a thoughtful dialogue and consideration of risks and benefits. Regulators have an opportunity to define the characteristics of DRBs that would be most instructive to them and encourage DRBs that foster trust and accountability. For example, DRBs with robust external membership and formal documentation may demonstrate an organisation's commitment to protecting individuals' rights and help focus any regulatory inquiry. Organisations using DRBs with these pre-determined qualities may thereby have reduced penalties.

In addition, in line with the recommendations made by the HLEG,[33] CIPL has suggested the adoption of **regulatory sandboxes** in the area of data protection and broader digital oversight, based on successful models deployed in the context of regulatory oversight over the financial services sector.[34] Also, the experiences of the UK ICO, which has launched regulatory sandboxes in a pilot phase with 10 organisations, should be illustrative for other regulators and encourage them to embark on these innovative regulatory methods.[35] The need to reconcile the use of data (including health data) to save lives with other individual and societal interests as recently highlighted by the COVID-19 crisis, as well as the expected increase of AI uses in the post-COVID world (such as for personalised medicine, online recruitment, airport biometric checks, etc.), would provide the perfect subject matter for an agile regulatory sandbox.[36] Out of the 20 EU countries that published their AI national strategies, 15 have already recognised the relevance of regulatory sandboxes in the AI field.[37]

Regulatory sandboxes provide a supervised "safe space" for piloting and testing innovative products, services, business models or delivery mechanisms of participating organisations in the real market, using the personal data of real individuals. They provide for **open and constructive collaboration between organisations and regulators**, giving the former assurance that their products meet regulators' expectations while helping the latter better understand the technologies they are regulating. In particular, sandboxes have the following benefits in the context of development and deployment of AI technology:

- They provide organisations with **reduced time-to-market** for new products and services, combined with **assurances** that they have built in appropriate safeguards;

- They have the potential to **address and resolve some of the more challenging aspects** of developing and deploying AI against the backdrop of less agile and more traditional regulatory frameworks;

- They provide an opportunity for accountable **organisations and innovative regulators to work and learn together in a collaborative fashion** to enable all the benefits of AI and the protection of individuals' rights;

- They **reduce regulatory uncertainty** in getting new ideas to market in the field of AI and **create incentives—especially for SMEs—to innovate** with more confidence in the regulatory environment;

- They enable **frank and confidential discussions** about the implications and acceptability of groundbreaking AI-based products or services, while providing **assurance that innovation is taking place in a responsible and accountable manner**, which ultimately contributes to enhancing individuals' and society's trust;

- They are very well adapted to **AI's rapid technological developments and the "data-fication**" of our societies and economies that requires innovative regulatory approaches;

- They enable the application of legislation **to new and emerging situations**, contribute to developing regulators' understanding of new technologies and inform future regulator guidance and regulatory approaches;

- They trigger initiatives to **boost regulators' AI capabilities** by launching, for instance, data scientist secondment programs in partnership with academia and industry; and

- **Joint sandboxes** involving several regulators can enable resolution of inconsistencies in laws or between regulators' interpretations. They also enable regulators with less AI experience to learn from regulators with more AI and general regulatory experience.

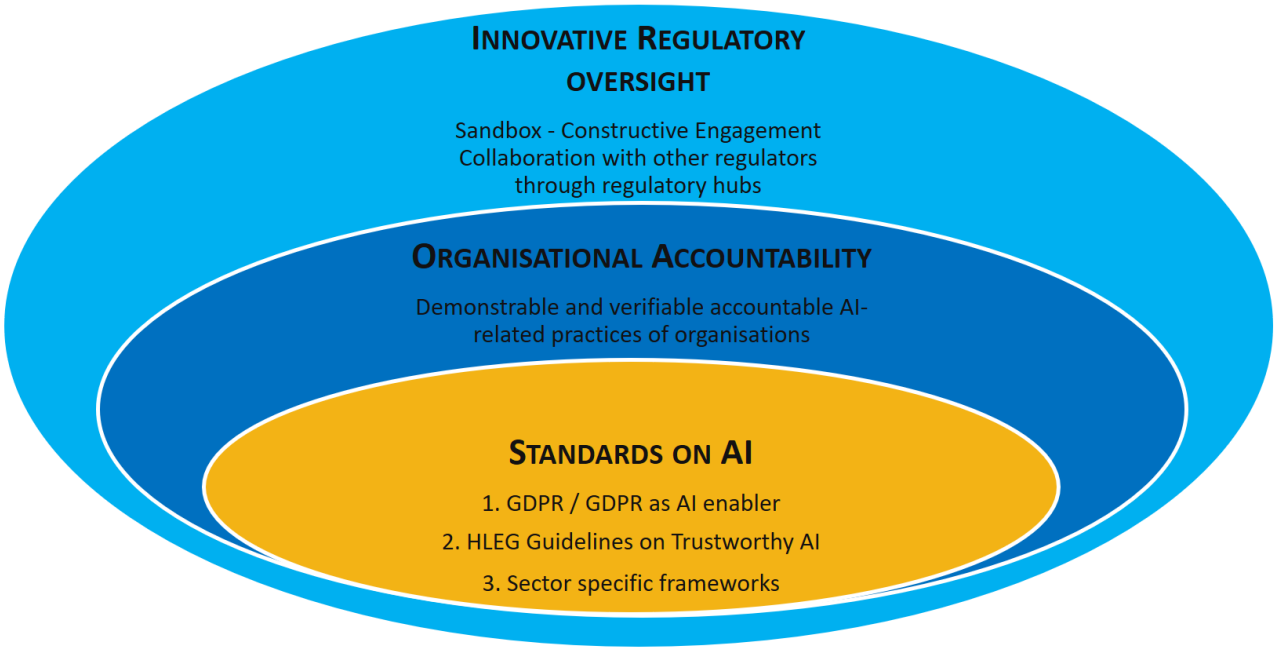
The EU AI regulatory framework must provide an **explicit statutory basis for regulators to set up regulatory sandboxes** in the context of AI with appropriate and relevant regulators including data protection, competition, audiovisual, consumer, health/pharma, telecom and financial regulators, at both the local and EU levels, leveraging the expertise of national notified bodies and centers of excellence. For EU-wide and cross-sector topics, this collaboration can be set up as part of the EU AI governance structure, where different AI experts could set up and participate in joint sandboxes.

Summary of CIPL Recommendations

- **Rely on existing authorities and avoid creating an additional layer of AI-specific regulators**;
- Rely on the current avenues for **redress and legal recourse at the national level;**
- Expand **expertise on AI within existing authorities;**
- Provide that the GDPR remains fully applicable **whenever AI processes personal data;**
- **Put in place regulatory hubs to avoid a** fragmented approach at the member state level and at the sector level;
- Build a **pragmatic and consistent** voluntary certification and labeling model **applicable at the EU level;**
- **Further incentivise data review boards to** promote a thoughtful dialogue and consideration of risks and benefits of AI systems by organisations; and
- Provide an **explicit statutory basis for setting up regulatory sandboxes** in the context of AI, including across countries and sectors.

Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

Appendix 1 – CIPL's Layered Approach to Regulating AI in the EU



Appendix 2 – CIPL Accountability Framework

Organizations must be able to demonstrate accountability – internally and externally

Accountability is not static, but dynamic, reiterative and a constant journey

Accountability requires comprehensive privacy programs that translate legal requirements into risk-based, verifiable and enforceable corporate practices and controls

Company values and business ethics shape accountability

Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

Appendix 3 - Mapping Best Practices in AI Governance to the CIPL Accountability Framework

This table outlines examples of accountable AI activities undertaken by selected organisations of different sectors, geographies and sizes based on the CIPL Accountability Framework and against each accountability element. The practices are not intended to be mandatory industry standards, but serve as specific examples that are calibrated based on risks, industry context, business model, size and level of maturity of organisations.

ACCOUNTABILITY ELEMENT	RELATED PRACTICES
Leadership and Oversight	• Public commitment and tone from the top to respect ethic, values, specific principles in AI development • Institutionalized AI processes and decision-making • Internal Code of Ethics rules • AI/ Ethics/Oversight Boards, Councils, Committees (internal and external) to review risky AI use cases and to continuously improve AI practices • Appointing Board member for AI oversight • Appointing Responsible AI Lead/Officer • Privacy/ AI engineers and champions • Set up an internal interdisciplinary board/senior working group (e.g. Lawyer, Technical teams, Research, Business units, internal audit, procurement, public affairs, thought leadership) • Appointment of privacy stewards to coordinate others • Ensuring inclusion and diversity in AI model development and AI product teams
Risk Assessment	• Understand AI purpose and use case in business/ processes—for decision making, or input into decision, or other • Understand impact (benefits and risks) on individuals and society • Algorithmic Impact Assessment / Algorithmic bias—tools to identify, monitor and continuous test, including sensitive data in data sets to avoid human bias • Fairness assessment tools to ensure biases are tested for, identified and any anomalies are mitigated to avoid concept drift in algorithms • Ethics Impact Assessment • Broader Human Rights impact assessment • DPIA for high risk processing • Assessment needs to include the benefits vs. risks of the AI autonomy, and challenge if such autonomy is necessary • Consider anonymization techniques • Document trade-offs (e.g. accuracy—data minimization, security—transparency, impact on few—benefit to society) for high-risk processing as part of the DPIA • Data quality assessment via KPIs • Framework for data preparation and model assessment – assessed and used by data scientists – including feature engineering, cross validation, back-testing, validated KPIs by business, etc. • Establishing controls and implementing safeguards to mitigate risks and trade-offs; • Working agile in close collaboration between business and data experts to assess regularly the needs and results accuracy – squad also includes data analysts, data engineers, IT and software engineers to ensure that the model can be properly used • Developing standardized risk assessment methodologies, which take into account the likelihood and severity of risk factors on individuals and/or society, level of human oversight involved in individually automated decisions with legal

	effects as well as their explainability (according to the contextual factors of the AI decision) and auditability, etc.
Policies and Procedures	• High level principles for AI—how to design, use, sell • Adopting specific AI policies and procedures on how to design, use or sell AI; • Assessment questions and procedures • Accountability measures for 2 stages – training and decision taking • White, black and gray lists of AI use • Evaluate the data against the purpose—quality, provenance, personal or not, synthetic, in-house or external sources • Purpose and other contextual factors of AI determines how much human intervention is required • Level of verification of data input and output; • Check no bias or unfair discrimination in the operation or outcome throughout the entirety of AI lifecycles • Pilot testing AI models before release • Use of protected data (e.g. encrypted, pseudonymised or where useful synthetic data) in some AI/ML models • Use of high quality but smaller data sets • Where applicable federated AI learning models (data doesn't leave device), considering trade-off with data security and user responsibilities • Special considerations for organisations creating and selling AI models, software, applications • Due diligence checklists for business partners using AI tech and tools • Using external tools, guidelines, self-assessment checklists • Processes and procedures to receive and address feedback and complaints • Define escalation steps with regards to reporting, governance, risk analysis and handling, etc. • Reliability – process for the testing and verification of the reliability of the AI system documented and operationalized. • Exploring ways to anonymise, de-identify or tokenise data, or to use synthetic data to train AI models; • Baseline model (if possible explainable) to assess uplift of advanced ones (advanced models should be used only if needed, model decision/KPI should consider the model complexity to be avoided – under the Occam's Razor principle) • Ideation phase between all stakeholders (data scientists, business, final user, control functions etc.) where needs, outcomes, validations rules, maintenance, need for explainability, budget, etc. are discussed • Documenting the use of AI technologies, the categories of data used in connection with the technologies, the decision-making process, and the identified risks and mitigations • Application of privacy and security by-design in AI life cycle
Transparency	• Different needs for transparency to individuals, regulators, business /data partners and internally to engineers and, leadership at the different stages of AI lifecycle • Adequate disclosures communicated in simple, easy to understand manner • AI must be inclusive and thus also accessible and usable by those in special needs/disabilities • Explainability is part of transparency and fairness • Transparency trail: explainability of decision and broad workings of algorithm; more about the process than the technology; what factors and what testing to be fair; accountability for impact of decisions on a person's life; what extent of human oversight

	• Explain that it is an AI/ML decision, if possibility for confusion (Turing test) • Provide counterfactual information • Differentiated and flexible transparency—linked to context, audience/users, purpose of explainability and risk, severity of harm—prescriptive lists of transparency elements is not helpful • Understand customers' expectations and deploy based on their readiness to embrace AI—tiered transparency • From black box to glass box—looking at the data as well as algorithm /model; aspiration of explainability helps understand the black box and builds trust • Define criteria of deployment of AI technologies within the organization (e.g. usage scenarios) and communicate them to the user • Traceability trail to make the AI system auditable, particularly in critical situations • Model cards (short documents accompanying AI models to describe context in which model should be used, what is the evaluation procedure) • Data hub for transparency on data governance, data accessibility, data lineage, data modification, data quality, definition, etc. • Use of LIME, SHAP, etc. for interpretation
Training and Awareness	• Data scientist training, including how to avoid and address bias • Cross functional training – privacy professionals and engineers • Ad hoc and functional training • Fairness training to technology teams • Ethics training to technology teams • Uses cases where problematic AI deployment has been halted • Role of "Translators" in organizations, explaining impact and workings of AI
Monitoring and Verification	• Capability for human in the loop—in design, in oversight, in redress • Capability for human understanding of the business and processes using AI • Capability for human development of software and processes • Capability for human audit of input and output • Capability for human review of individual decisions with legal effects • Ongoing monitoring, validation and checks • Oversight committees even in design stage • Redress to a human, not to a bot • Monitoring the eco-system from data flow in, data process and data out • Reliance on different audit techniques • Counterfactual testing techniques • Version control and model drift, tracking of black box, algorithms by engineers • RACI models for human and AI interaction • Pre-definition of AI audit controls • Internal audit team specialized on AI and other emerging technologies • Processes must allow human control or intervention in the AI system where both technically possible and reasonably necessary • See the *Assertion-based Framework for the Audit of Algorithms*, Otto Koppius and Iuliana Sandu • Model monitoring (including back-testing and feedback loop) and maintenance process
Response and Enforcement	• Complaints-handling • Redress mechanisms and appropriate personnel for individuals to remedy AI decision • Feedback channel • Internal supervision of AI deployment

Centre for Information Policy Leadership
HUNTON ANDREWS KURTH

Appendix 4 - Overlap between the White Paper's suggested requirements for high risk AI and the existing requirements of the GDPR

AI White Paper Requirements	GDPR Requirements
Training data Requirements aimed at providing reasonable assurances that the subsequent use of the products or services enabled by the AI system is safe, in that it meets applicable safety rules (e.g. that AI systems are trained on data sets that are sufficiently broad and cover all relevant scenarios needed to avoid dangerous situations) Requirements to take reasonable measures aimed at ensuring that such subsequent use of AI systems does not lead to outcomes entailing prohibited discrimination. Obligation to use data sets that are sufficiently representative Requirements aimed at ensuring that privacy and personal data are adequately protected during the use of AI enabled products and services	**Art. 5(1)(a)** – data should be processed fairly and transparently **Art. 5(1)(d)** – data should be accurate **Art. 5(1)(f)** – security of processing **Art. 25** – data protection by design and by default
Keeping of records and data Keep accurate records regarding the data set used to train and test the AI system, including a description of the main characteristics and how the data set was selected Keep data sets when justified Keep documentation on the programming and training methodologies, processes and techniques used to build, test and validate the AI	**Art. 5(2)** – accountability principle. The controller must be able to demonstrate its compliance with data protection principles **Art. 24** – Responsibility of controller to implement appropriate technical and organisational measures and demonstrate compliance **Art. 30** – records of processing
Information provision Ensure clear information is provided as to the AI system's capabilities and limitations, in particular the purpose for which the systems are intended, the conditions under which they can be expected to function as intended, and the expected level of accuracy in achieving the specified purpose Citizens should be clearly informed when they are interacting with an AI system and not a human being	**Art. 5(1)(a)** - data should be processed fairly and transparently **Arts. 13 & 14** – obligation to provide notice to data subjects, including meaningful information on logic involved in case of ADM with legal effects **Art. 22** - ADM with legal effects and right to obtain human intervention

Robustness and accuracy Requirements ensuring that the AI systems are robust and accurate, or at least correctly reflect their level of accuracy, during all life cycle phases Requirements ensuring that outcomes are reproducible Requirements ensuring that AI systems can adequately deal with errors or inconsistencies during all lifecycles phases Requirements ensuring that AI systems are resilient against both overt attacks and more subtle attempts to manipulate data or algorithms themselves, and that mitigating measures are taken in such cases	**Arts. 5(1)(d) &(f)** – data accuracy and security **Art. 5(2)** – accountability principle. The controller must be able to demonstrate its compliance with data protection principles **Art. 24** – responsibility of controller; technical and organisational measures have to be reviewed and updated where necessary **Art. 22** – ADM with legal effects and right to obtain human intervention **Art. 25** – data protection by design and by default **Art. 32** - Security **Arts. 35 and 36** – DPIA, risk mitigation and prior DPA consultation
Human oversight Requirement ensuring that the output of the AI system does not become effective unless it has been previously reviewed and validated by a human Requirement ensuring that the output of the AI system becomes immediately effective, but human intervention is ensured afterwards Requirement on monitoring of the AI system while in operation and the ability to intervene in real time and deactivate In the design phase, requirements to impose operational constraints on the AI system	**Art. 5(1)(a)** – data should be processed fairly and transparently **Art. 5(2)** – accountability principle. The controller must be able to demonstrate its compliance with data protection principles **Art. 22** – ADM with legal effects and right to obtain human intervention **Art. 35** – DPIA and risk mitigation

[1] CIPL is a global data privacy and cybersecurity think tank in the law firm of Hunton Andrews Kurth LLP and is financially supported by the law firm and 90 member companies that are leaders in key sectors of the global economy. CIPL's mission is to engage in thought leadership and develop best practices that ensure both effective privacy protections and the responsible use of personal information in the modern information age. CIPL's work facilitates constructive engagement between business leaders, privacy and security professionals, regulators and policymakers around the world. For more information, please see

CIPL's website at http://www.informationpolicycentre.com/. Nothing in this submission should be construed as representing the views of any individual CIPL member company or of the law firm of Hunton Andrews Kurth.

[2] For more information on CIPL's Accountable AI project: https://www.informationpolicycentre.com/ai-project.html

[3] https://ec.europa.eu/digital-single-market/en/news/white-paper-artificial-intelligence-european-approach-excellence-and-trust. CIPL's Paper provides recommendations for the future EU regulatory framework for AI with a focus on risks for fundamental rights, personal data, privacy protection, non-discrimination and safety. It does not specifically cover the liability and compensation regime.

[4] https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl-hunton_andrews_kurth_legal_note_-_how_gdpr_regulates_ai__12_march_2020_.pdf

[5] See https://ec.europa.eu/info/sites/info/files/communication-european-strategy-data-19feb2020_en.pdf page 12 "The Commission's approach to regulation is to create frameworks that shape the context, allowing lively, dynamic and vivid ecosystems to develop. Because it is difficult to fully comprehend all elements of this transformation towards a data-agile economy, the Commission deliberately abstains from overly detailed, heavy-handed ex ante regulation, and will prefer an agile approach to governance that favours experimentation (such as regulatory sandboxes), iteration, and differentiation."

[6] Recital 89 of the GDPR provides that: "Directive 95/46/EC provided for a general obligation to notify the processing of personal data to the supervisory authorities. While that obligation produces administrative and financial burdens, it did not in all cases contribute to improving the protection of personal data. Such indiscriminate general notification obligations should therefore be abolished, and replaced by effective procedures and mechanisms which focus instead on those types of processing operations which are likely to result in a high risk to the rights and freedoms of natural persons by virtue of their nature, scope, context and purposes […]"; See also https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_the_eu_commissions_consultation_on_a_european_strategy_for_data__29_may_2020_.pdf

[7] The concept of personal data is broad and receives an extensive interpretation from DPAs. Anonymisation techniques in parallel have to be increasingly robust to avoid possible re-identification. In most instances, an AI system will rely on "mixed data sets" that include both personal and non-personal data.

[8] The sector-based approach may however be relevant for use of AI by public authorities, such as, for instance, for the purpose of law enforcement or predictive policing.

[9] This also requires distinguishing between the developer and the deployer of an AI system. The role and responsibilities of the developer/deployer will depend on the level of control in the specific use scenario and its related risks.

[10] See Article 35(1) of the GDPR.

[11] See Recital 4 and Article 1(2) GDPR: "This Regulation protects fundamental rights and freedoms of natural persons and in particular their right to the protection of personal data." See Article 29 Working Party "Guidelines on Data Protection Impact Assessment (DPIA) and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679." WP 248 rev.01, 4 2017 at page 6: "… [t]he reference to "the rights and freedoms" of data subjects primarily concerns the rights to data protection and privacy but may also involve other fundamental rights such as freedom of speech, freedom of thought, freedom of movement, prohibition of discrimination, right to liberty, conscience and religion." http://ec.europa.eu/newsroom/document.cfm?doc_id=47711

[12] The EU Charter of Fundamental Rights lists: Human dignity, right to life, right to integrity, prohibition of torture and inhuman treatment, prohibition of slavery and forced labor, right to liberty and security, respect for private and family life, protection of personal data, right to marry and found a family, freedom of thought, conscience and religion, freedom of expression and information, freedom of assembly and of association, science and academic freedom, right to education, freedom to choose an occupation and right to engage in work, freedom to conduct a business, right to property, right to asylum, protection in the event of removal, expulsion or extradition, equality before the law, non-discrimination, cultural, religious and linguistic diversity, equality between men and women, rights of the child, rights of the elderly, integration of persons with disabilities, workers' right to information, right of collective bargaining and action, right of access to placement services, protection in the event of unjustified dismissal, fair and just working conditions, protection of child labor, family and professional life, social security and social assistance, healthcare, access to services of general economic interest, environmental protection, consumer protection, right to vote and stand as candidate at elections, right to good administration, right to access to documents, right to petition, freedom of movement and residence, diplomatic and consular protection, right to an effective remedy and to a fair trial, presumption of innocence and right of defense, legality and proportionality of criminal offenses and penalties, right not to be punished twice for the same criminal offense. https://www.europarl.europa.eu/charter/pdf/text_en.pdf

[13] See CIPL comments to WP29 Guidelines on ADM and profiling, 1 December 2017, pages 5 and 6 containing examples of legal effects/similarly significant effects/no legal or similar effects. https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_comments_to_wp29_guidelines_on_automated_individual_decision-making_and_profiling.pdf

[14] See CIPL's paper "Risk, High Risk, Risk Assessments and Data Protection Impact Assessments Under the GDPR" https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_gdpr_project_risk_white_paper_21_december_2016.pdf
[15] See Article 6(1)(f) GDPR that requires balancing "the legitimate interests pursued by the controller or by a third-party" and "the interests or fundamental rights and freedoms of the data subject."
[16] For more detail see CIPL's paper at note 14, page 7.
[17] https://www.informationpolicycentre.com/organizational-accountability.html
[18] https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_final_draft_regulating_for_results_-_strategies_and_priorities_for_leadership_and_engagement.pdf
[19] https://ico.org.uk/media/about-the-ico/documents/2615039/project-explain-20190603.pdf
[20] https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_second_ai_report_-_hard_issues_and_practical_solutions_01.17.2020.pdf
[21] See note 4.
[22] See CIPL paper "Organisational Accountability – Past, Present and Future (October 2019) https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_on_organisational_accountability_%E2%80%93_past_present_and_future.pdf
[23] See https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_second_report_-_artificial_intelligence_and_data_protection_-_hard_issues_and_practical_solutions__27_february_2020_.pdf at page 4
[24] See note 18 at page 8.
[25] See https://ec.europa.eu/newsroom/dae/document.cfm?doc_id=60419 on page 19
[26] CIPL Accountability Discussion Paper 2 –Incentivising Accountability: How Data Protection Authorities and Law Makers Can Encourage Accountability, July 23, 2018 https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_accountability_paper_2_-_incentivising_accountability_-_how_data_protection_authorities_and_law_makers_can_encourage_accountability.pdf
[27] https://ec.europa.eu/digital-single-market/en/news/digital-innovation-hubs-helping-companies-across-economy-make-most-digital-opportunities
[28]https://www.canada.ca/en/government/system/digital-government/modern-emerging-technologies/responsible-use-ai/algorithmic-impact-assessment.html. The Algorithmic Impact Assessment is a questionnaire designed to help organisations assess and mitigate the impacts associated with deploying an automated decision making system. The 60 questions are focused on an organisation's business processes, data and system design decisions. The results provide the organisation with an impact level as well as a link to the requirements under applicable law.
[29] https://ec.europa.eu/futurium/en/ai-alliance-consultation/guidelines/2
[30] CIPL does not consider at this stage that a formal review of the GDPR is necessary as its current challenges, in particular as they relate to lack of harmonisation and consistency, can be resolved by the Commission, the EDPB and DPAs using the existing institutional and regulatory mechanisms and their wide interpretative powers. See https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_response_to_eu_commission_consultation_on_gdpr_evaluation__28_april_2020_.pdf
[31] https://www.iso.org/committee/6794475.html
[32] See Rachel Dockery, Fred Cate, & Stanley Crosley, "Why Data Review Boards Are a Promising Tool for Improving Institutional Decision-Making," IAPP (28 February 2020), available at https://iapp.org/news/a/why-data-reviewboards-are-a-promising-tool-for-improving-institutional-decision-making/#; see also CIPL Second Report, at p. 30, https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_second_report_-_artificial_intelligence_and_data_protection_-_hard_issues_and_practical_solutions__27_february_2020_.pdf;
[33] See note 25 at page 41.
[34] See CIPL Paper "Regulatory Sandboxes in Data Protection – Constructive Engagement and Innovative Regulation in Practice - March 8, 2019 https://www.informationpolicycentre.com/uploads/5/7/1/0/57104281/cipl_white_paper_on_regulatory_sandboxes_in_data_protection_-_constructive_engagement_and_innovative_regulation_in_practice__8_march_2019_.pdf
[35] https://ico.org.uk/about-the-ico/news-and-events/news-and-blogs/2020/03/combining-privacy-and-innovation-ico-sandbox-six-months-on/
[36] See note 32 at page 8.
[37] Belgium, Cyprus, Czech Republic, Denmark, Estonia, Finland, Malta, France, Germany, Italy, Latvia, Lithuania, Luxembourg, Portugal and Slovakia.