Information Accountability Foundation

Assessments for an AI World

Legitimate Interest Assessment

Lynn Goldstein,

Peter Cullen

Steve Wood

November 2024

Executive Summary

The Information Accountability Foundation (IAF) for many years has advocated for a multi-dimensional stakeholder, interests, rights, freedoms, and risks balancing assessment. We believe this balancing approach is applicable to the balancing required by the legitimate interest lawful basis of processing under the EU General Data Protection Regulation (GDPR) and the UK GDPR (collectively the GDPR). With the expansion of artificial intelligence (AI), especially large language models (LLMs) such as ChatGPT, reliance on legitimate interests as a lawful basis of processing has expanded too. However, a gap remains with respect to processes that business might use to demonstrate this lawful basis, what regulators might expect, and in turn, regulators' perceptions of business capability to meet these requirements. This gap has led to an environment where trust and confidence in legitimate interests from both sides is low. As a result, the IAF established a project whose goal is to bridge this gap by developing a normative framework using multi-dimensional balancing to build capability in the business community that would create greater confidence by regulators in the use of legitimate interests as a fore-runner to what will be required in terms of demonstrable accountability in an AI world.

To help explain the required balancing, the IAF developed a directory of rights, interests, stakeholders, and consequences and used colors and symbols and a mathematic model to represent those factors. All these factors then are balanced – multi-dimensional balancing – and the results can be supported pictorially, mathematically, or both. While data protection authorities have begun issuing guidance on the proper application of legitimate interests as the lawful basis when AI is the processing activity and while this guidance is consistent with the IAF's multi-dimensional balancing, this guidance does not show organisations how to conduct this balancing. The draft model legitimate interest assessment (Draft Model LIA – see page 18) in this report, developed by the IAF with input from global businesses, regulators, civil society, academics, and nongovernmental organizations (NGOs), shows both business and regulators how to demonstrate legitimate interest multi-dimensional balancing when AI is the processing activity. In addition, while there are different purposes for different assessments, it is the view of the IAF the balancing or weighing requirements in the Draft Model LIA are the same or similar to what is required for the balancing part of a Data Protection Impact Assessment (DPIA) and what will be required in Fundamental Rights Impact Assessments (FRIA) called for by the EU AI Act for high-risk systems, many of the U.S. state privacy laws, and almost all emerging AI laws and regulations. It is the IAF's view that the balancing requirements of all these assessments which are similar call for multi-dimensional balancing.

It is the IAF's hope that the Draft Model LIA generated by this project will be used to socialize the goal more broadly and to support the development of AI regulatory guidance throughout Europe. An additional benefit of this project would be demonstrating the utility of legitimate interest as a lawful basis, particularly for AI scenarios, to those jurisdictions developing new data protection and privacy laws or considering revisions to existing laws.

Background

Data processing has become more complex with the growth of AI, especially with the prevalent use of LLMs. For organisations that collect and use data in a complex, powerful, and sometimes innovative manner, it increasingly is likely that legitimate interests will be the lawful basis for processing under the GDPR. To demonstrate that the risks associated with the processing, the balancing process required examines the full range of rights and interests of all stakeholders (so the interests or fundamental rights and freedoms of the data subject are not overridden), organisations often conduct an assessment (LIA).

The challenge of using lawful bases, such as contract and consent, for activities such as AI training when there may be no direct relationship between the organization and the data subjects, also highlights the importance of utilizing legitimate interest as the most sustainable approach to addressing safeguards, accountability, and responsible innovation. There is a perception amongst some regulators that legitimate interest is a weaker protection for data subjects than consent. Even though there is not a hierarchy of lawful bases in the GDPR, it has been implicit in the approach of some regulators. Recently, however, as the draft CNIL LI How-To Sheet observed, in recognizing that LI may be the most appropriate legal basis to develop an AI system, "it is often difficult to obtain the consent of individuals at a large scale or when personal data are collected indirectly." The European Data Protection Board in very recently issued guidance for consultation that covers legitimate interest in detail stated: "Article 6(1)(f) GDPR should neither be treated as a "last resort" for rare or unexpected situations where other legal bases are deemed not to apply, nor should it be automatically chosen or its use unduly extended on the basis of a perception that Article 6(1)(f) GDPR is less constraining than other legal bases."

However, despite the six years after passage of the GDPR, a gap remains with respect to processes that business might use to demonstrate this lawful basis, what regulators might expect, and in turn, regulators' perceptions of business capability to meet these requirements. This gap has led to an environment where trust and confidence in legitimate interests from both sides is low.

While the IAF has for many years advocated for a multi-dimensional stakeholder and interests balancing assessment, in practice, assessments such as LIAs have focused more on a narrow set of data protection rights with the primary focus being the individual data subject. As a result, a broader balancing approach, even one that addresses the multiple interests of individuals, historically has not been encouraged by regulators. Organisations have not shown that they have the competency to effectively perform a more comprehensive balancing assessment. In addition, a core issue is that there has been a **scarcity of guidance relating to the process of balancing.** All this has contributed to a self-reinforcing problem as depicted in the Exhibit below.

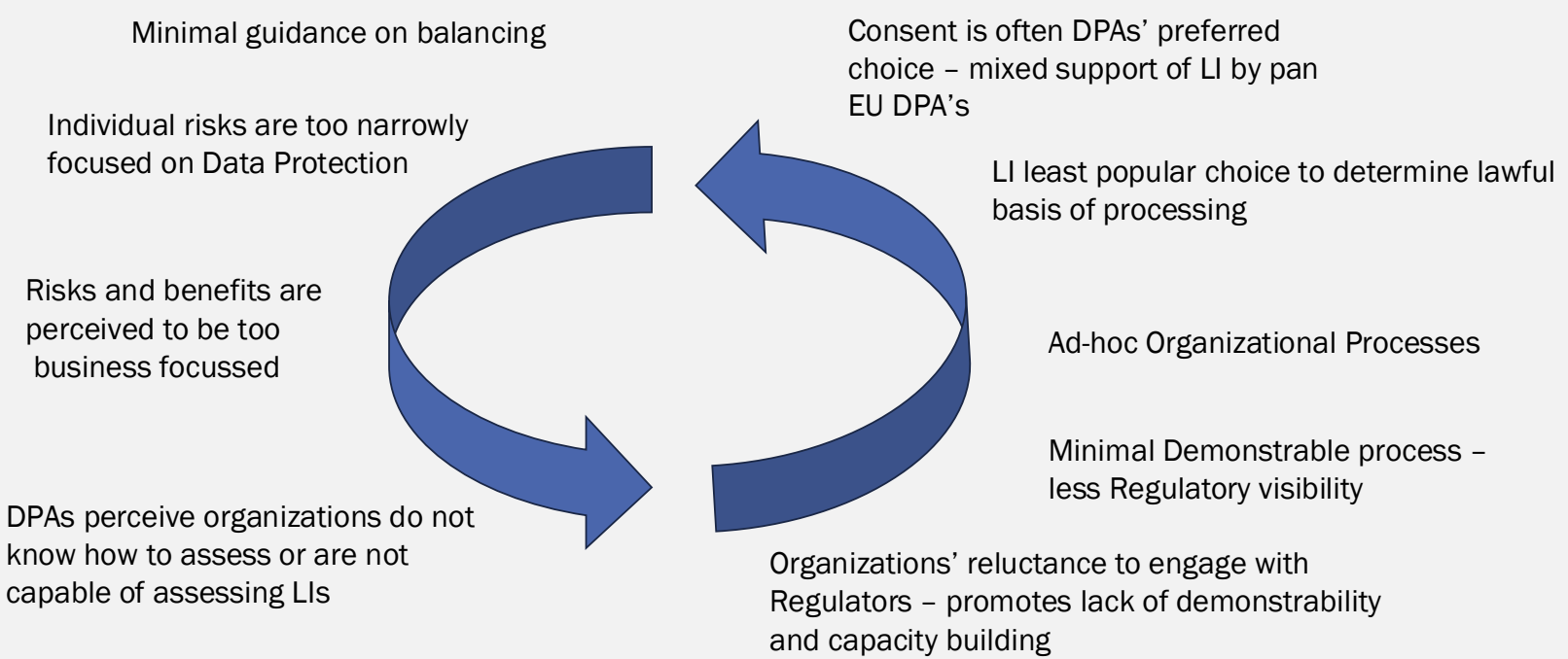


Exhibit 1

.

To help bridge this gap, the IAF created the AI Legitimate Interest Assessment Project (Project).

The Project

In 2017, the IAF developed the Legitimate Interests and Integrated Risk and Benefits Assessment. As part of that work, the IAF concluded that many of the same issues that come into play to determine if a processing is risky to individuals are the same issues that need to be considered when determining that legitimate interests is a lawful basis for processing.

The Ethical Data Impact Assessment was developed by the IAF in 2019 for use by organizations when they were conducting advanced analytics such as AI and machine learning. It looked at the full range of rights and interests of all parties in a data processing activity when advanced data analytics may impact people in a significant manner and/or when data enabled decisions are made without the intervention of people. It required careful consideration of all data activity benefits and risks to individuals and society and evaluated these benefits and risks by using a common risk assessment process often found in organisations' Enterprise Risk Management programs.

The IAF learned in the 2021 AI Impact Assessment and A Principled Approach to Rights and Interest Balancing that a methodology (Three-Dimensions Methodology) is needed to, in an orderly and repeatable fashion, identify and demonstrate three components:

- **The individual rights and benefits** (in a fundamental rights-based system) or established legitimate interests (in legal systems when fundamental rights are not established),
- **The full range of stakeholders whose rights or interests are involved**, and the impact to their interests, and
- **The adverse processing impacts that may be involved (to all stakeholders), their likelihood and level of consequence, recognizing that adverse processing impacts sometimes only may be reduced and not eliminated.**

To create that robust process, the IAF developed a directory of rights, interests, stakeholders, and consequences, and used colors and symbols and a mathematic model to represent those factors. In higher-risk scenarios, understanding the perspectives of representative groups of stakeholders via consultation and detailed analysis processes is an essential component of a LIA and understanding the risks more broadly. All these factors then are balanced – multi-dimensional balancing – and the results can be supported pictorially, mathematically, or both. For clarity, the detailed analysis of each stakeholders' rights, interests, and potential adverse impacts is key whether this analysis is supported by any pictorial or mathematical element.

The goal of the Project was to develop a normative framework using multi-dimensional balancing to build capability in the business community that would create greater confidence by regulators in the use of legitimate interests as a fore-runner to

what will be required in terms of demonstrable accountability in an AI world. The objective is that the Draft Model LIA generated by the Project would be used to socialize this goal more broadly and to support the development of AI regulatory guidance throughout Europe. An additional benefit of the Project would be demonstrating the utility of legitimate interest as a lawful basis, particularly for AI scenarios, to those jurisdictions developing new data protection and privacy laws or considering revisions to existing laws.

IAF's Three-Dimensions Methodology was combined with IAF's extensive track record of developing frameworks and assessments that include associated and requisite governance and controls. Through the multi-stakeholder convening process, the Project delivered the Draft Model LIA, a draft normative framework consisting of an assessment with governance controls as well as a suggested demonstrability framework that is intended to:

- expand business governance capability and sustainability
- increase regulator confidence
- Inform further guidance to business by regulators

The resulting Draft Model LIA benefits business and regulators who, with confidence from the multi-stakeholder process, can incorporate it into horizontal governance strategies, processes and procedures within their business organizations and into provided regulatory guidance.

The Project Process

Step 1: The IAF developed a draft model LIA addressing key problem areas.

- The IAF incorporated relevant global laws and regulations and guidance from Data Protection Authorities
- This step was supported by individual dialogue in 2024 with select business participants and regulatory authorities.

Step 2: The Draft Model LIA was reviewed through convening meetings with participating business organizations in Dublin and London.
Step 3: Based on feedback, an evolved Draft Model LIA was finalized, and the IAF convened a multi-stakeholder session that engaged the business community, the regulatory community (four Data Protection Authorities), and the NGO community (see Appendix IV). The Draft Model LIA was presented and discussed.
Step 4: The preparation and issuance of this Report and the final Draft Model LIA.

What was Learned from the Business and Regulator Meetings

Interviews with Business

Some organizations say legitimate interests is not a frequently used lawful basis for processing. The most common reasons are:

- The legitimate interest concept emanated from Europe's data protection law and the directive 95/46 before the GDPR. There are few parallels in Latin America or Asia. Therefore, legitimate interests does not work on a global basis.
- Some regulators have more nuanced views than other regulators. Therefore, it is easier consistently to use contract or consent as the lawful basis for processing.

However, other organisations said they rely on legitimate interests as the lawful basis for processing for the following reasons:

- Legitimate interests provides a more holistic approach. The business did research to understand how its customers go to market and the interests of its customers and to identify the harm from its data set.
- Legitimate interests allows a reasonable and sustainable course of action, especially since consent and contract do not work in all situations.
- Business needs a lawful basis that works consistently, and contrary to the view expressed by other organisations, these organisations are of the view legitimate interest is the only lawful basis that works consistently and can be used worldwide.

Many organizations are very interested in getting to an industry standard or best practice or a methodology and agreed steps. Organisations are happy to comply once they know what the bar is; they do not want to be flying blind. Regulators repeatedly have asked business to better explain the interests of the stakeholders and how the balancing was done. However, it is clear, more guidance from regulators is needed on these subjects.

From the perspective of organisations, there does not seem to be a widespread practical approach by regulators to legitimate interests as a lawful basis.

Interviews with Regulators

Regulators reported that business really does not fully consider or understand the interests from the data subject's point of view. It seems difficult for business to put itself in the position of the person whose data it is using.

There often is a lack of evidence to support the positions taken by business. Business needs to identify its legitimate interest and to provide a complete explanation of the minimum amount of data needed to support its position and to be able to demonstrate how it arrived at that conclusion.

Some are looking for ways to make the assessment by business of interests less subjective and more objective. Others want business to demonstrate their thinking but do not want to be prescriptive about how to do that. Frequently, to them, a LIA looks more like a DPIA. The level of attention is disappointing; basic content is missing. They often feel that the organisation has not done a full assessment so there is nothing to demonstrate.

Some felt that legitimate interests was being used as the lawful basis to process because it was the easiest lawful basis to use. Others felt there was a preference for consent as a lawful basis because it was easy to measure, i.e., three legal requirements, and it aligns with the individual having control; as a result, consent was being overused or misused in some contexts.

The Draft Model LIA

The GDPR and Draft Legislative and Regulatory Guidance

Article 6(1)(f) of the GDPR provides that: "Processing shall be lawful only if and to the extent that at least one of the following applies: . . . processing is necessary for the purposes of the legitimate interests pursued by the controller or by a third party, except where such interests are overridden by the interests or fundamental rights and freedoms of the data subject which require protection of personal data, in particular where the data subject is a child." In a recent case, the Court of Justice of the European Union (CJEU) recapped its previous case law on the three cumulative conditions for relying on the legitimate interests lawful basis for processing:

1. The pursuit of a legitimate interest by the data controller or by a third party
2. The necessity to process personal data for the purposes of the legitimate interest pursued
3. The rights and interests of the data subject do not outweigh the legitimate interest of the controller or of a third party (Balancing Requirement)

As the above discussion of the interviews with business and regulators shows, the key problem area is the third condition – the Balancing Requirement. During the Project, the UK Government introduced a bill that sets out a limited list of legitimate interests, and several data protection authorities issued draft guidance on legitimate interests, including the Balancing Requirement.

The EDPB guidance for consultation and the Draft CNIL LI How-To-Sheet say that to satisfy the Balancing Requirement the controller must balance its anticipated legitimate interests against the data subjects' anticipated interests, rights, and freedoms. To do this, the controller must identify and describe, i.e., must measure:

- The controller's interests and the data subjects' interests, fundamental rights and freedoms. An "interest" is the broader stake or benefit that a controller or third party may have in engaging in a specific processing activity. "The fundamental rights and freedoms of the data subjects include the right to data protection and privacy, but also other fundamental rights and freedoms, such as the right to liberty and security, freedom of expression and information, freedom of thought, conscience and religion, freedom of assembly and association, prohibition of discrimination, the right of property, or the right to physical and mental integrity, which may be affected by the processing, either directly or indirectly (e.g. through a chilling effect . . . .) The interests of the data subjects to be considered as part of the balancing test include any interest that may be affected by the processing at stake, including, but not limited to, financial interests, social interests or personal interests." When addressing what is a legitimate interest, the EDPB guidance for consultation says: "There is no exhaustive list of interests that may be considered as legitimate. In the absence of a definition of that concept in the GDPR, a wide range of interests is, in principle, capable of being regarded as legitimate. Both the GDPR and the CJEU have expressly recognized several interests as being legitimate, such as having access to information online, ensuring the continued functioning of publicly accessible websites, obtaining the personal information of a person who damaged someone's property in order to sue that person for damages, protecting the property, health and life of the co-owners of a building, product improvement, and assessing the creditworthiness of individuals, among others." The UK Data (Use and Access) Bill proposes the following be recognized as legitimate interests: disclosures for purposes of process described in UK GDPR Article 6(1)(e), national security, public security and defence, emergencies, crime detection, investigation, prevention, and prosecution, and safeguarding vulnerable individuals.

- According to the EDPB guidance for consultation, the greater the anticipated benefits of the processing, the more likely the legitimate interest of the controller is to prevail over the rights and freedoms of individuals. The following factors make it possible to measure the positive impact of the interests pursued:
    a. The extent and nature of the expected benefits of the processing for the controller but also for third parties or the interest of the public or society. The UK ICO's guidance also recognizes wider public interest: "The legitimate interests of the public in general may also play a part when deciding whether the legitimate interests in the processing override the individual's interests and rights. If the processing has a wider public interest for society at large, then this may add weight to your interests when balancing these against those of the individual."
    b. The usefulness of the processing carried out in order to comply with other legislation.
    c. The specification of the interests pursued

- The impact of the processing on data subjects, including
    a. The nature of the data to be processed. Special categories of personal data (sensitive data) and personal data relating to criminal convictions enjoy additional protection under Articles 9 and 10 of the GDPR, respectively. The types of data that data subjects generally consider to be more private (e.g., financial and location data) or rather of a more public nature (e.g., data concerning one's professional role) can impact the processing. Generally, the more sensitive or private the nature of the data to be processed, the more likely it is that the processing of such data will have a negative impact on the data subjects, and the more weight should be attributed to it in the balancing test.
    b. The status of the data subjects and the status of the controller engaging in the processing
    c. The context of the processing and the specific data processing methods also may influence the impact that the processing may have on the rights and interests of the data subjects.
    d. Further consequences of the processing. Consequences of the envisaged processing may further the rights, freedoms and interests of the data subjects.
- The reasonable expectations of data subjects. They play an important role in the Balancing Requirement, to limit the risks that data subjects are unduly surprised by the processing or by its consequences or implications.
- The final balancing of opposing rights and interests, including the possibility of further mitigating measures. Once the controller has identified and assessed the legitimate interest(s) being pursued, the relevant interests, rights and freedoms of the data subjects, the impact of the processing, and the reasonable expectation of data subjects, the controller should be able to strike a balance between all the interests, rights and freedoms identified. If the outcome of this assessment is that the legitimate interest(s) being pursued are not overridden by the data subjects' interests, rights and freedoms, the envisaged processing may take place. However, if the data subjects' interests, rights and freedoms seem to override the legitimate interest(s) being pursued, the controller may consider mitigating measures which may be of a technical, organizational or legal nature to limit the impact of the processing on data subjects, in view of achieving a fair balance between the rights, freedoms and interests involved. If controller decides to implement mitigating measures, they should perform the balancing test anew, to assess whether the legitimate interest(s) being pursued are overridden by the data subjects' interests, rights and freedoms, after the adoption of the mitigating measures. The controller should aim to strike this balance as objectively as possible. If the data subjects' interests, rights and freedoms override the legitimate interest being pursued, and no sufficient mitigating measures can be taken, the processing cannot be based on Article 6(1)(f) of the GDPR.

The chart at the end of the Draft CNIL LI How-To Sheet is the beginning of the determination of the residual impact/risk for each impact/risk identified. The process set forth in the Enisa Interoperable EU Risk Management Framework explains how to determine residual risk. For each risk:

- The inherent risk is determined by identifying the risks and determining its likelihood and significance, and
- The residual risk for each inherent risk is determined by identifying and determining the effectiveness of the security measures for each risk.

The Organisation of the Draft Model LIA

The beginning for the Draft Model LIA is the three-part test enunciated by the UK Information Commissioner's Office (ICO): Purpose Test, Necessity Test, and Balancing Test (Sections 2, 6, and 9) and the three-part test discussed above and recapped in the recent CJEU case, Koninklijke Nederlandse Lawn Tennisbond v Autoriteit Persoonsgegevens, and set forth in the EDPB guidance for consultation and the CNIL LI How-To Sheet (Sections 1 and 3, 6, and 9 and 10). In developing the Draft Model LIA, business encouraged the IAF to find the delta between the various laws that require balancing for automated decision-making including profiling (ADM). Therefore, the content in Section 7 of the Draft Model LIA also includes requirements of the Colorado Rules, the California Generative AI Transparency Law, the California AI Transparency Act, and the Draft California Regulations and in the Balancing Section of the Draft Model LIA includes requirements in the FRIA required under the EU AI Act and in the UK Data (Use and Access) Bill. Also, additional AI governance controls are included in Section 8 of the Draft Model LIA for illustrative purposes, as they support some of the other operational elements outlined in the "assumptions" section of the Draft Model LIA and align with specific requirements in some of the U.S. state laws. The Draft Model LIA also refers to the 2021 AI Impact Assessment for a fulsome approach to assessing these types of issues. The other substantive parts of the Draft Model LIA are the Nature of the Data (Section 4) and the Operational Elements of the Processing Activity (Section 5).

The IAF's Multi-Dimensional Balancing

It has long been the view of the IAF that to do multi-dimensional balancing, consistent with Article 4 of the GDPR, it is necessary to take into account all interests and fundamental rights found in the Charter of Fundamental Rights of the EU (EU Charter) and the UN Universal Declaration of Human Rights (UN Charter). Examples of these rights and interests go beyond autonomy and other data protection rights and include, for example, safety, better education and healthcare, shared benefits from technology and more robust opportunities (the EU Charter specifically includes broader societal rights and interests such as freedom to conduct a business). This multi-dimensional balancing provides form and substance in a world that demands both protection and advanced probabilistic programs.

In an age of AI, it is too simplistic to say that because the word "risk" is not mentioned in the language of the GDPR that sets forth the legitimate interests lawful basis of processing, data subjects do not have an interest or a right in having the risks to their personal data considered as part of the Balancing Requirement. Indeed, one of the first freedoms protected in the EU Charter and the UN Charter is the right to security of person which includes negative impacts to data subjects' data security, such as unauthorized access, destruction, use, modification or disclosure of personal data or unauthorized activity resulting in

the loss of availability of personal data. Understanding whether processing will create substantial risks for a data subject requires consideration of the nature of risk and the broader interests of the data subject, groups of individuals and society. A determination of what is a risk is based on shared societal values. Data protection risks traditionally have been seen as harms-based, but the concept of free flow of data is predicated on the risks associated with lost opportunity as well. This breadth of coverage means consideration of what is legitimate or not cannot be a simple fulcrum looking to balance interests between a controller and a single data subject's desire for autonomy but rather must be a variable analysis. In addition, an assessment of risk should include the risk of a benefit or interest not being realized.

This thinking has not been reflected in previous regulatory guidance on legitimate interests. The advent of generative AI is causing data protection authorities to begin to think more broadly – to address all stakeholders in addition to data subjects such as groups of individuals and society (collectively Stakeholders), to think about not only interests but also benefits, to think not only about harms but also risks, to think about mitigations to risks, and to recognize that not all risks have to be eliminated. The IAF has been advocating this approach since the Ethical Data Impact Assessment in 2019. The EDPB guidance for consultation recognizes that the purpose of the Balancing Requirement "is not to avoid any impact on the interests and rights of the data subjects altogether. Rather, the purpose of the Balancing requirement is to avoid a disproportionate impact and to assess the weight of these aspects in relation to each other." The EDPB also acknowledges that the Balancing Requirement should use an objective assessment.

Put together the EDPB guidance for consultation and the chart at the end of the Draft CNIL How-To Sheet (CNIL Chart) starts to provide guidance on how to conduct this multi-step balancing of Stakeholders' benefits, interests, rights, freedoms, and of risks and mitigation. This approach is consistent with multi-dimensional balancing advocated by the IAF. The Draft Model LIA takes a step in both advancing this type of guidance and proposing a normative model balancing approach.

Business has told the IAF that the Balancing Requirement of the LIA often is part of their overall risk assessment process, and therefore it is important to keep in mind: (i) the risk-based nature of the GDPR; and (ii) "The chief objective of Information Security Management is to implement the appropriate measurements in order to eliminate or **minimize** the impact the various security related threats might have on an organization." Enisa Risk Management (emphasis added); *see generally* COSO Internal Controls.

Appendix III to the Draft Model LIA applies this methodology to the Balancing Requirement when scraping is used to collect user generated content and for model training. For clarity purposes, only the interests, benefits and risks identified in the Draft CNIL LI How-To Sheet and the Draft CNIL LI Web Scraping Focus Sheet are used. The analysis in Appendix III shows how Information Security Management processes can be used in the Balancing Section of the Draft Model LIA to weigh Stakeholders' interests, benefits, risks, and mitigations. This analysis is a tool for discussion purposes; it never can be the result. It is not done in a

vacuum, and consultation occurs as determinations are reached. The Information Security Management process is a means to help facilitate, organize, and visualise the multiple interests, benefits, risks, and mitigations. It never is a substitute for the human conversation and consultation that needs to take place in satisfying the Balancing Requirement.

Finally, the EDPB guidance for consultation also points out that there are different purposes for different assessments: for example, a LIA is to evaluate and determine lawful processing and a DPIA is to evaluate "high risk" processing. The LIA should be made at the outset of the processing with the involvement of the data protection officer (DPO), if designated, and the DPO likewise also should be involved in the DPIA process, ensuring consistency between the assessments, particularly for inputs around risks and mitigations. The IAF thinks this also is true of the balancing or weighing requirements of the FRIA called for by the EU AI Act for high-risk systems, many of the U.S. state data protection laws, and almost all emerging AI laws and regulations. It is the IAF's view that the Balancing Requirements of all these assessments which are similar call for multi-dimensional balancing.

What was Learned in the Multi-Stakeholder Session

Participants in the Multi-Stakeholder Session were from global businesses, academics, NGOs, and present and former regulators (see Appendix IV). The nature of the remarks made during the Multi-Stakeholder Session broke down into two categories: Overall Observations which are of general interest to participants in the Project and readers of this Report, and Specific Recommendations regarding the Draft Model LIA itself.

Participants' Overall Observations

- It is important to determine the lawful basis before processing begins. It is important to get the lawful basis correct at the start. It should be made clear why legitimate interest is the lawful basis chosen and why other lawful bases were not chosen. Switching a lawful basis, e.g., from consent to legitimate interests, is a problem.
- Business needs to understand what is going on before doing it, i.e., understanding why business wants to do a specific processing activity is essential. Business at the beginning needs to explain how it thought of this particular processing activity, how it wants to make it happen, and whether this approach is the best one. Business also needs to think about how data will be used in the future.
- There should be transparency about what personal data is seen by whom within the business. If it is a lot of data but there is uncertainty what the data will show, then a smaller amount of data should be started with, and the amount of

data can be expanded as more is learned. Also, it is helpful for business to say what it is not going to do with the data. The more transparency, the better.

- There should be separate legitimate interests and AI assessments. Content could be drawn from both, and consistency between the two is important, but thematically they should be separate documents.
- A DPIA and a LIA are different types of assessments. For a LIA, the Stakeholders are individuals, groups of individuals and society, but for a DPIA, the Stakeholder is the data subject. While the core purpose of these two assessments is different, the balancing approach of each assessment is similar if not the same.
- Scalability is a big problem. Since the volume of LIAs that will need to be done is great (e.g., there are more LLMs than there are people to assess them), business needs to be able to automate certain parts of a LIA, and standardized assessment profiles are needed. However, there must be a degree of human involvement at certain key stages.
- A formal LIA is a document for a regulator and likely prepared by legal staff or counsel. Important LIAs also may be considered by cross organisational privacy or data protection committees. While the content that might be part of this document will be sourced from other assessment and analysis processes, it should not be a goal for business to fully operationalize a LIA.
- A common format for a LIA that regulators can expect would be helpful. The organisation must be able to answer a regulator's questions.
- For demonstrability, business should use an IT system that can produce artifacts for regulators. Business should be able to provide proof of efficiency of the mitigation measures. Evidence of mitigation being effective is helpful, e.g., KPIs may be a way of demonstrating effectiveness. Reviews or consultation by an outside group (e.g., Data Review Boards) also may be useful, particularly in novel and innovative uses of technology. Business should address why synthetic data (and other privacy enhancing technologies that support data minimization) are not being used instead of personal data. Evidence that supports the purpose of the processing also is key.
- Preparation of a LIA is an iterative process involving many different parts of the organisation.
- Regarding necessity, a processing activity could be done in numerous different ways; describe why the business has decided to do it in this way. The less intrusive the better; "less intrusive" means "are there alternative means of processing available?".
- Proportionality means a focus on significance. Business should weed out risks as it goes along to get to the major risks. There should be mitigations for all risks; if not, business should explain why.
- For high-risk processing, senior management should sign off on the LIA. What is "senior management" may depend on the jurisdiction.

- Completion of a LIA using a recognised process does not guarantee that the regulator will agree with the conclusion of the business. There is a meaningful risk that regulators will disagree. Sections of the LIA inform the result, but they do not dictate the result.
- Overall, the Draft Model LIA is viewed as a very useful matrix tool, especially the visualisation of benefits and risks.
- Simplifying the scoring too much could give a false sense of security; on the other hand, cramming too much into the methodology was not embraced. One recommendation was to use the scoring as a storytelling tool and to use it to show how the decision moves from red to green. Another recommendation was that the scoring may be most useful for the project team itself, e.g., what mitigations work for what risks. These comments reinforced that the use of "scoring" is not a replacement or a substitute for a deep, detailed and descriptive analysis in the LIA.

Participants' Specific Recommendations

Several attendees had recommendations for specific revisions to the Draft Model LIA.

Recommendation: The document is contradictory, i.e., some parts are too broad, and some parts are missing

- IAF Response: There is greater consistency between what is in the main part of the Draft Model LIA and what is in the appendices to the Draft Model LIA.

Recommendation: There is redundancy in the Draft Model LIA, i.e., content is included in both the body and the appendices of the Draft Model LIA.

- IAF Response: This redundancy has been eliminated.

Recommendation: "What is the processing activity, i.e., what is the legitimate interest?" question needs to be added. At the beginning, business needs to be very specific about the processing activity.

- IAF Response: This question has been added.

Recommendation: The "What is the purpose of the processing?" question is quite high level. Need to be specific about the purpose. Ask "What is the stated end of the processing activity?"

- IAF Response: This question has been added.

Recommendation: In Section 5, the Operational Elements of Processing Activity, for each category of personal data collected,

list each personal data source, state whether each data subject reasonably would expect the data would be used for the purpose of the processing activity, add the details about the recency of the data, describe how data is being collected without consent, ask for detail regarding existing and new data retention periods, and add whether reused personal data is compatible with the purpose for which the personal data were originally collected.

- IAF Response: This detail has been added.

Recommendation: In the Balancing Section, ask for sources and more detail on Stakeholders. Workers Councils and Trade Unions have not been listed as Stakeholders for employment scenarios.

- IAF Response: Sources detail has been requested, and formatting has been changed so more detail on Stakeholders is provided. Workers Councils have been included and Unions have been changed to Trade Unions in the list of Stakeholders in Appendix II.

Recommendation: In the Balancing Section, the benefits and opportunities need to be expanded.

- IAF Response: In the Balancing Section, the formatting has been changed to make clear that the benefits/interests for all Stakeholders, including those to the public, need to be considered.

Recommendation: In the Balancing Section, organize the taxonomy of risks so that specific mitigations align against specific risks.

- IAF Response: Formatting has been changed so that mitigations for each risk are aligned.

Recommendation: Sets of prompts about what to consider when completing the LIA should be added.

- IAF Response: Detail regarding what to consider has been added.

Recommendation: Add who is part of the team completing the LIA and whether external parties were involved, what their roles were, what their input was, and whether it was followed, and if no external parties were involved, why not, and how that function was replaced.

- IAF Response: That language has been added.

Going Forward

There are two possible separate work streams going forward.

- Further socialisation of the Draft Model LIA

- Potential New Projects

The IAF will explore with business, data protection authorities, NGOs, and academics about using the Draft Model LIA to advance business practices, including institutional review boards, and future regulatory guidance.

IAF Draft Model Legitimate Interests Assessment

Introduction

The Draft Model Legitimate Interest Assessment (Draft Model LIA) that follows is an example of what could be provided to a regulator:

- when data processing has become more complex with the growth of artificial intelligence (AI), especially with the prevalent use of large language models (LLMs), and
- when these organisations increasingly use legitimate interests as the lawful basis for processing under the EU General Data Protection Regulation (GDPR) and the UK GDPR (collectively the GDPR).

Rather than a separate assessment, the Draft Model LIA should be thought of as a form of output of what may be contained in many existing organisational assessments that likely will be put together by the legal team in conjunction with the business team and the Data Protection Officer (DPO). It is designed with an AI processing activity in mind. Less complex data processing activities may not require all questions and may require different approaches.

The IAF's Multi-Dimensional Balancing

The GDPR requires the balancing of the legitimate interests of the controller or a third party and the interests or fundamental rights and freedoms of the data subject. To do multi-dimensional balancing, consistent with Recital 4 of the GDPR, it is necessary to take into account all interests and fundamental rights found in the Charter of Fundamental Rights of the EU (EU Charter) and the UN Universal Declaration of Human Rights (UN Charter). The IAF's impact analysis considers the interests of the data subject, groups of individuals, and society (Stakeholders) and weighs all the benefits/interests/freedoms, including the effectiveness of mitigating measures, and weighs the risks/harms and effectiveness of mitigations, using a 1 – 5 scale which aligns with many typical risk management approaches. IAF's multi-dimensional weighing is unique. It is capable of weighing each of the Stakeholders, benefits, and risks as are relevant to the processing being assessed. It is capable of weighing each Stakeholder vis-à-vis each of the other factors. It can demonstrate the results of the detailed and descriptive analysis mathematically or pictorially or both and supplement the required narrative response. Examples of how to conduct multi-dimensional balancing are found in Section 13 of the Draft Model LIA.

Organisation of the Draft Model LIA

In developing the Draft Model LIA, business encouraged the IAF to find the delta between the various laws that require balancing for automated decision-making including profiling (ADM). Therefore, the Draft Model LIA comes from numerous sources:

- The three-part test enunciated by the UK Information Commissioner's Office (ICO): Purpose Test, Necessity Test, and Balancing Test (Sections 2, 6, and 9),
- The three-part test discussed above and recapped in the recent CJEU case, and set forth in the EDPB guidance for consultation and the CNIL LI How-To Sheet (Sections 1 and 3, 6, and 9 and 10),
- The content in Section 7 includes requirements of the Colorado Rules, the California Generative AI Transparency Law, the California AI Transparency Act, and the Draft California Regulations,
- The Balancing Section includes requirements in the Fundamental Rights Impact Assessment (FRIA) required under the EU AI Act and in the UK Data (Use and Access) Bill,
- Additional AI governance controls are included in Section 8 for illustrative purposes, as they support some of the other operational elements outlined in the "assumptions" section and align with specific requirements in some of the U.S. state laws, and Section 8 also refers to the 2021 AI Impact Assessment for a fulsome approach to assessing these types of issues.
- The other substantive parts of the Draft Model LIA are the Nature of the Data (Section 4) and the Operational Elements of the Processing Activity (Section 5).

Some key assumptions are listed at the end of the Draft Model LIA. One of these assumptions is that the traditional linear or step process of assessing "legitimate interest," then "necessity," and finally "balancing of interests, rights and freedoms" associated with a traditional LIA is not well suited to today's data driven business models and that instead, an iterative, multistage LIA more aligned with an AI development lifecycle, such as the Draft Model LIA, may be appropriate.

	Foundational Questions to Identify the Legitimate Interest, Determine the Necessity of the Processing Activity, Assess the AI Development Lifecycle, and Balance the Competing Interests
1. **What is the Processing Activity** (legitimate interest)	What is the processing the organization wants to undertake? Why does it want to do it? What is the problem the business is trying to solve by engaging in this activity? Will this activity solve this problem? Describe with specificity why the controller or third party needs to conduct the processing and how the processing solves that problem. The EDPB guidance for consultation makes clear that the concept of "interest" is closely related to, but distinct from, the concept of "purpose." "Interest" is the broader stake or benefit that a controller or third party may have in engaging in a specific processing activity. The ICO suggests considering the following in identifying the legitimate interest: • Why does the controller or third party want to process the data – what is it trying to achieve? • Who benefits from the processing? In what way? • Are there any wider public benefits to the processing? • How important are those benefits? • What would be the impact if the processing couldn't go ahead? • Would the use of the data be unethical or unlawful in any way? See Appendix II for examples of legitimate interests.

2. What is the Purpose of the Processing? (stated end objective for this personal data use)	The objective of the processing activity, including the relationship between the controller or third party and the data subject whose personal data will be processed, those data subjects' reasonable expectations concerning the purpose of the processing of their personal data, how the controller will process the personal data, and the context of the processing. A key element to include is the "why" of the purpose. The purpose must not be identified or described in generic terms. Is there data or evidence that supports the need for the processing? As part of the purpose analysis, consider all the stakeholders and ask: • What benefit does the controller expect to get from the processing? • Do any third parties benefit from the processing? • Are there any wider public benefits to the processing? A short summary of the benefits, and any supporting objective evidence, resulting from the processing that may flow, directly and indirectly, to the controller (including any downstream risks as a result of the controller sharing the data), data subject, other expected stakeholders, and the public should be provided, including how important these benefits are and what the impact would be if the processing could not go ahead. These benefits must be identified and described with specificity in Section 9. The EDPB guidance for consultation makes clear that "purpose" is the specific reason why the data are processed: the aim or intention of the data processing. A FRIA requires a description of the deployer's processes in which the high-risk AI system will be used in line with its intended purpose. The "purpose" analysis should be reassessed at key stages of the development data processing lifecycle, and a different purpose may be appropriate at different stages of the development data processing lifecycle. A FRIA requires a description of the time and frequency in which each high-risk AI system is intended for use. By extension, the LIA should be updated as each activity develops

3. What is Legitimate Interests the Lawful Basis for Processing?	Why is legitimate interest the most appropriate lawful basis for processing? Why can't consent be obtained? Why can't contracts be entered into with the individual(s)? According to the Draft CNIL LI How-To Sheet and the EDPB guidance for consultation, the first condition to reliance on legitimate interests as a legal basis is that the interests pursued must be "legitimate," and "interests pursued may be presumed as legitimate if they cumulatively are: • Manifestly lawful under the law, i.e., not contrary to EU or Member State law; • Determined in a sufficiently clear and precise manner; • Real and present (i.e. non-hypothetical or proven) for the organization concerned.
4. What is the Nature of the Data Involved? (are any special categories of personal data included)	The categories of personal data (information relating to an identified or identifiable natural person) to be processed and whether they include any special categories of personal data, including personal data from a known child or other vulnerable groups. Special categories of personal data include personal data revealing racial or ethnic origin, political opinions, religious or philosophical beliefs, or trade union membership, genetic data, biometric data for the purpose of uniquely identifying a natural person, data concerning health or data concerning a natural person's sex life or sexual orientation. The details of the process implemented to make sure that special categories of personal data and inferences from special categories of personal data are not shared unless there is a separate lawful basis in place and deletion or another safeguard is in place (in many cases, it will be appropriate to delete special categories of personal data within 12 hours of the processing). For uses of ADM, the organisation must identify the actions it has taken or any actions it plans to take to maintain the quality of personal data processed by the ADM.

5. What are the Operational Elements of the Processing Activity?	In determining the level of detail and specificity to provide, the controller must consider the type, amount, and sensitivity of personal data processed, the impacts that operational elements will have on the level of risk presented by the processing activity, and any relevant unique relationships (including whether the data subject interacts with the organisation, how they do so (e.g., via websites, applications or offline), and the nature of the interaction (e.g., to obtain a good or service from the organisation)). Relevant operational details may include: a. Sources of personal data for each category of personal data, b. Method of collecting personal data (directly or indirectly), does data subject reasonably expect personal data will be used for the purpose of the processing, how data is being collected without consent; c. Method of using, disclosing (including when, how, and what), storing, retaining, and sharing personal data; d. The volume, recency and extent of the personal data – does it contain all customers or subset for example? Does it represent a major percentage of a country's population? e. Technology (including whether it is novel or well established, the logic (including any assumptions or limitations) of ADM, and the output of the ADM and how the organisation will use the output of the ADM)) or processors to be used; f. Names or categories of personal data recipients, including third parties, affiliates, and processors that will have access to the personal data, the processing purpose for which the personal data will be provided to those recipients, categorical processes that the controller uses to evaluate that type of recipient, and how the organisation has made or will make data subjects aware of the involvement of those recipients; g. Operational details about the processing, including planned processes for personal data collection, use, disclosure, storage (including data location, backups), retention, and sharing; h. Specific types of personal data to be processed; i. How the controller's processing of personal data complies with data minimization requirements (including why the controller needs to process the personal data and the relevance of the personal data to the processing); j. How long the controller will retain each category of personal data (existing v. new retention period), the criteria used to determine the retention period, why the controller needs to retain each category for that length of time, and the frequency of purging each category no longer necessary; k. The approximate number of data subjects whose personal data the controller plans to

	process and the context of the relationship with the data subject; I. If personal data is being reused, is the processing activity compatible with the purpose for which the personal data initially were collected.
6. Is the Processing Necessary?	Is the intended processing capable of achieving the interest pursued? Does this processing help to further the various stakeholders' interests? Is it a reasonable way to go about it? Is there no other less intrusive way of achieving the interest pursued than to carry out the intended processing? Is the processing targeted and proportionate? If the use of personal data is required, is the processing indeed necessary to achieve its objective? According to the EDPB guidance for consultation, assessing what is "necessary" involves ascertaining whether in practice the legitimate data processing interests pursued cannot reasonably be achieved just as effectively by other means less restrictive of the fundamental rights and freedoms of data subjects. In practice, it generally is easier for a controller to demonstrate the necessity of the processing to pursue its own legitimate interests than to pursue the interests of a third party, and that the latter kind of processing generally is less expected by the data subjects. The ICO suggests considering the following in applying the necessity test: • Does this processing help to further the identified interest? • Is it a reasonable way to go about it? • Is there another less intrusive way to achieve the same result? The "Necessity Test" analysis should be reassessed at key stages of the AI development data processing lifecycle as elements may change as the development lifecycle advances. By extension, the LIA should be updated. The initial Necessity Test content should be based on the preliminary benefits and risk analysis above. Most legal interpretations of the term "necessary" normally mean that the processing does not need to be essential but should go beyond merely useful or 'nice to do.' The Draft LI How-To Sheet recognizes the connection between "necessity" and "mitigation of risk" in its discussion of data minimization. The second condition, which is that the processing must be necessary, also is to be examined in connection with the principle of data minimization, i.e., "the controller must ensure that it is necessary to process personal

	data or to store them in a form which permits the direct or indirect identification of individuals and that it is necessary to have recourse, where appropriate, to a technical solution which involves processing a large volume of personal data." While it may be appropriate to do a separate AI assessment, for the purpose of doing an LIA, it may be helpful to have a basic understanding of the AI involved in the processing. According to the Draft How-To Sheet, it is necessary to consider the nature of the AI system and the intended operational use. For the purposes of the LIA, ADM includes an AI System (machine-based system that infers from the input it receives how to generate outputs such as predictions, content, recommendations, or decisions that can affect physical or virtual environments) to the extent it includes personal data and it is without meaningful human involvement. ADM processing activity must include the following: 1. A plain language explanation of the specific types of personal data that were or will be processed by the ADM, including the personal data used to train the ADM and the sources of the personal data. (The Colorado AI Act requires a deployer, or a third-party contracted by the deployer (collectively deployer), to include a description of the categories of data the ADM processes as inputs and outputs the ADM produces.) 2. The decision to be made using ADM 3. The benefits of automated processing over manual processing for the stated purpose 4. A plain language explanation of why ADM is being used to achieve the purpose of the processing, the appropriate use(s) of the ADM, and any limitation on the appropriate use(s) of the ADM 5. A plain language explanation of the training data and logic of the ADM, including any assumptions of the logic and any statistics used in the analysis, either created by the controller or provided by a third party which created the applicable ADM 6. If the ADM is conducted by the third-party software purchased by the controller, the context in which software is being used and whether this affects the decision to be made by the ADM, and the name of the software and copies of any internal or external evaluations sufficient to show the accuracy and reliability of the software where relevant to the risks described below 7. A plain language explanation of the output(s) secured from the ADM 8. A plain language description of how the output(s) from the ADM are or will be used, including whether and how they are used to make a decision to provide or deny or substantially contribute to the provision or denial of financial or lending services, housing, insurance, education, enrollment or education opportunity, criminal justice,
7. What is the Nature of the AI Systems and ADM?	

	legal service, employment opportunity, health-care services, or access to essential goods or services, including essential government service 9. A plain language explanation of the steps taken or any steps to be taken to maintain the quality of personal data processed by the ADM, including personal data used to train the ADM; 10. If there is human involvement in the use of the ADM, a plain language explanation of the degree and details of any human involvement 11. A plain language explanation of how the ADM is evaluated for validity, reliability, fairness and disparate impact, and the results of any such evaluation 12. A plain language explanation of any safeguards that the controller plans to implement to address the negative impacts to data subjects' privacy that are specific to its use of ADM or for data sets produced by or derived from the ADM (The Colorado AI Act requires a deployer to include a description of any transparency measures taken concerning the ADM, including any measures taken to disclose to a data subject that the ADM is in use when the ADM is in use and of the post-deployment monitoring and user safeguards provided concerning the ADM, including the oversight, use, and learning process established by the deployer to address issues arising from the deployment of the ADM) 13. If personal data has been processed or is being processed to train ADM, and that ADM has been or is being made available to other data subjects for their own use, how a plain language explanation of the appropriate purposes for which the data subjects may use the ADM has been provided or will be provided to those data subjects and any safeguards that have been implemented or will be implemented to make sure that the ADM is used for appropriate purposes by those other data subjects for their own use (The Colorado AI Act requires a deployer to include an overview of the categories of data the deployer used to customize the ADM if the deployer used data to customize the ADM) 14. If personal data has been processed or is being processed to train ADM and is made or is being made available to other businesses (recipient businesses), how the necessary facts have been or will be provided to those recipient businesses to conduct the recipient businesses' risk assessments, including a plain language explanation of any requirements or limitations that the organisation identified as relevant to the permitted use of ADM 15. Following an intentional and substantial modification to an ADM, the Colorado AI Act requires the deployer to include a statement disclosing the extent to which the ADM was used in a manner that was consistent with, or varied from, the developer's intended

	uses of the ADM The developer of each generative AI system or service must post on its website a high-level summary of the datasets used in the development of the generative AI system or service: 1. Sources or owners of the datasets 2. Description of how the datasets further the intended purpose of the AI system or service 3. Number of data points included in the datasets, which may be in general ranges and with estimated figures for dynamic datasets 4. Description of the types of data points within the datasets 5. Whether the datasets include any data protected by copyright, trademark, or patent, or whether the datasets are entirely in the public domain 6. Whether the data sets were purchased or licensed by the developer 7. Whether the data sets include personal information 8. Whether the datasets include aggregate consumer information 9. Whether there was any cleaning, processing, or other modification to the datasets by the developer, including the intended purpose of those efforts in relation to the AI system or service 10. The time during which the data in the datasets were collected, including a notice if the data collection is ongoing 11. The dates the datasets were first used during the development of the AI system or service 12. Whether the generative AI system or service used or continuously uses synthetic data generation in its development Large generative AI providers (those with over one million monthly users) must give users the option to have a visible and easily perceived disclosure placed on any generative AI image, video, or audio content that the user has generated; must label the same type of generative AI content with a disclosure imperceptible to the human eye; must make an AI detection tool available at no cost that takes advantage of the either or both disclosures on their generated content, thereby enabling users to help determine whether the content was created or altered by the provider's generative AI system.
8. What are Other AI Governance Requirements?	Over and above the requirements in the LIA relating to ADM, since a full range of rights and interests refers to terms/issues that are not fully defined or included in the GDPR text, such as "test for fairness, negative inferences, adversarial use/attacks, unconscionable treatment, creation of anxiety, embarrassment, fear or other mental trauma, over collection, surveillance or use beyond what a reasonable data subject would expect," it may be helpful to include the following LIA governance topics in the LIA. These topics may lead to the need

	to create governance/control processes and questions that would appear in the functional part of a DPIA. Question themes found in an AI Impact Assessment, which includes ADM, are set forth below. Examples of questions in an AI Impact Assessment can be found here.

a. Lifecycle	The project plan should account for each stage of the model lifecycle and for regular assessment during each stage of the model lifecycle: plan and design model, collect and process data, build and use model, verify and validate model, deploy and use model, operate and monitor model, assess Impacts of model.
b. Fairness	The term "fairness" has been described, and steps are in place to measure and test for achieving fairness. This also will include an assessment of what the reasonable expectations of individuals are and how this has been informed by transparency actions.
c. Traceability	Traceability can be maintained across data, experiments, model versions and usage. Performance can be captured against success criteria.
d. Model Training	The quality of training data has been assessed. Were there enough total training samples? Were the samples representative of different social groups based on race, gender, color, age, income, etc.?
e. Model Testing	The performance of the model was tested. The model was well-trained and analyzed through different metrics – Precision, Recall, F1Score, Accuracy, Bias, Robustness and Sensitivity Training, Red Teaming, Postmortems, etc.
f. Equal treatment	All users are treated equally. If not – and the algorithms and predictive technologies prioritize certain information or sets prices or access differently for different users – describe how you would handle data subject/user demands or government regulations or contractual requirements that require all users be treated equally, or at least transparently unequally. The criteria for conducting Bias Audits under the Final Rule implementing New York City Local Law 144 may be helpful in determining whether a decision or other action is discriminatory.
g. Third Parties	If the organisation obtained models or datasets from a third party, describe how the risks of using third parties are assessed and managed and what the documentation requirements of using the third parties are.
h. End user	Describe how end-users or other subjects are made aware adequately that a decision, content, advice, or outcome is the result of an algorithmic decision.

i. Other Governance Requirements	In addition to the LIA for ADM, the GDPR and other laws also contain requirements regarding transparency (including model logic), opting out, and consent. Describe the additional governance processes put in place to address these additional requirements, such as substantive guidelines, policies, and procedures.
j. Governance Controls	Describe the governance controls in place that will enable a consistent, robust, repeatable development/implementation process, such as substantive guidelines, policies, and procedures. Describe how all project team members' roles have been established and communicated.
Balancing Process	
9. Is the Legitimate Interest Overridden by the Interests or Fundamental Rights and Freedoms of the Stakeholders? The LIA process must involve all relevant internal actors across the controller's organisational structure and, where appropriate, relevant external parties, to identify, assess and address the benefits and risks to individuals	According to the EDPB guidance for consultation, to "balance" the controller's or third party's legitimate interest in question against the interests or fundamental rights and freedoms of the data subjects (the individuals, groups of individuals, and society), the following must be identified and described: • The data subjects' interests, fundamental rights and freedoms • The impact of the processing on data subjects, including the nature of the data to be processed, the context of the processing, and any further consequences of the processing • The reasonable expectations of the data subject • The final balancing of opposing rights and interests, including the possibility of further mitigating measures After having identified the fundamental rights and interests that may be affected by the processing, the likely impact of the processing on the data subjects should be carefully assessed. This assessment should focus on the various way in which individuals may be affected – positively or negatively, actually or potentially – by the processing of their personal data The impact of the process on the data subjects may be influenced by: • The nature of the data to be processed – special categories of personal data and personal data relating to criminal convictions and offences enjoy additional protection, whether it is objectively possible to infer sensitive information from the data processed, the types of data generally considered to be of a more private (e.g., financial or location data) or of a more public (e.g., data concerning one's professional role) nature (the

	more sensitive or more private the nature of the data to be processed, the more likely it is that the processing of such data will have a negative impact on the data subject). • The context of the processing - the scale of the processing and the amount of personal data to be processed (in terms of overall volume of data, volume of data per data subject) and the number of data subjects affected), the status of the controller (including vis-á-vis the data subject), whether or not the personal data to be processed are combines with other data sets, the degree of accessibility and/or publicity of the data to be processed, and the status of the data subject. • Further consequences that the processing may have – potential future decisions or actions by third parties that may be based on the personal data to be processed the controller, the possible production of legal effects concerning the data subject, exclusion of discrimination against individuals, defamation or , more broadly, situations where there is a risk of damaging the reputation, negotiating power or autonomy of the data subject, financial losses which may be incurred by the data subject, exclusion from a service for which there is no real alternative, and risks to freedom, safety, physical and mental integrity or life of natural persons. These adverse outcomes may be ones that specifically can be foreseen or may be ones involving possible broader emotional impacts resulting from a data subject losing control over personal information or realizing that is has been misused or compromised. For each stakeholder, describe how the benefits of the processing outweigh the risks as mitigated by the safeguards. This description includes the sources and nature of risks to the rights of data subjects associated with the processing posed by the processing, the legitimate interests pursued by the controller, and why the controller's interests are not overridden by the interests or fundamental rights and freedoms of the data subject rights which require protection. The source and nature of the risks and benefits may differ based on the processing and type of personal data processed. The negative impacts to data subjects' privacy, or other data harms, associated with the processing, including the sources of these negative impacts, must be identified and the magnitude of the negative impacts and the likelihood of the negative impacts occurring must be described with specificity, including how the criteria used to determine the magnitude and likelihood of the negative impact. Consultations with stakeholders may inform people how to understand likely impacts, and how they may vary across different groups. Verified evidence from other organizations' use of similar technologies and systems may also provide evidence. This also may include

	evidence of unintended consequences. In doing the balancing test, the ICO suggests considering the impact of the intended processing and whether this impact overrides the interest identified: • What is the nature of the relationship of the controller with the individual? • Is any of the data particularly sensitive or private? • Would individuals expect their data to be used in this way? • Is the controller comfortable explaining its use of the data to individuals? • Are some individuals likely to object or find this use intrusive? • What is the possible impact on the individual? • How big an impact might this use have on them? • Is children's data being processed? • Are any of the individuals vulnerable in any other way? • Can any safeguards be adopted to minimize the impact? • Can an opt-out be offered? Use the specific Stakeholder Detailed Analysis, identify a full range of benefits and risks to key stakeholders. See **Appendix II** for examples of stakeholders and of benefits and risks. The EDPB guidance for consultation cautions: • When it is clear that many data subjects share the same interests, a combined assessment of such interests may suffice • The more intrusive a processing operation is, more specific circumstances should be factored into the assessment • The assessment of the interests at stake should not assume that all the affected data subjects share the same interests when there are – or should be -concrete indications of the existence of particular individual interests or when, from an objective perspective, it simply is not likely that all data subjects will have the same interest(s) assumed According to the Draft CNIL LI How-To Sheet, the following make it possible to measure the positive impact of the interests pursued: • The extent and nature of the expected benefits of the processing, for the controller but

	also for third parties, such as the end-users of the AI system or the interest of the public or society. A commercial interest may converge, to a certain extent, with a public interest (in general, the fact that a controller acts not only in its own interest but also in the interest of the community may give more "weight" to that interest). • The usefulness of the processing carried out to comply with other legislation. • The development of an open-source model, which, provided sufficient safeguards are in place, may have significant benefits for the scientific community, research, education and the adoption of these tools by the public. • The specification of the interests pursued. The more precisely an interest is defined, the more weight it will have in the balancing exercise, since it makes it possible to specifically apprehend the reality of the benefits to be considered. Conversely, an interest defined too broadly (e.g. "providing new services to its users") is less likely to prevail over the interests of individuals. The Draft CNIL LI How-To Sheet goes on to provide that those benefits must be balanced against the impact of the processing on data subjects. Specifically, the controller must identify and assess the consequences of all kinds, potential or actual, the development of an AI system and its subsequent use could have on the persons concerned as well as other concrete impacts of the processing on their situation. The actual impact of the processing on an individual are to be assessed according to the likelihood that the risks materialize and the severity of the consequences, which depend on the conditions of the processing as well as the AI system developed. To this end, it is necessary to take into account the nature of the data (e.g. sensitive data, highly personal data), the status of the data subjects (e.g. vulnerable persons, minors), the status of the organisation developing and/or deploying the AI system (the effects may be multiplied in the event of very wide use of AI), the way in which the data are processed (e.g. data crossing) or the nature of the AI system and the intended operational use. According to the Draft CNIL LI How-To Sheet and the EDPB guidance for consultation, the reasonable expectations of the individuals must be considered when assessing the impact of the processing on individuals. It is important to distinguish between the notion of reasonable expectations and what is considered common practice in certain sectors; the fact that certain types of personal data are commonly processed in each sector does not necessarily mean that the data subject can reasonably expect such processing. Reasonable

	expectations are a contextual aspect that the controller must consider when balancing the rights and interests at stake. Although not exhaustive, the following list illustrates contextual elements which can be considered in the assessment of the reasonable expectations of data subjects: • Characteristics of the relationship with the data subject or of the service – the very existence of a relationship with the data subjects and the date of termination of the relationships; the proximity of the relationship; the place and context of the data collection; the nature and characteristics of the service; applicable legal requirements in the relevant context • Characteristics of the "average" data subjects whose personal data is to be processed – the age of the data subject, the extent to which the data subject is a public figure, the (professional) position that the data subject holds and the level of understanding and knowledge of the envisaged processing that they are likely to have in a certain context Identify the safeguards that the controller plans to implement to address the negative impacts identified and explain how these safeguards address these negative impacts with specificity, including whether and how they eliminate or reduce the magnitude of the negative impacts or the likelihood of the negative impacts occurring; and any safeguards the controller will implement to maintain knowledge of emergent risks and countermeasures. A FRIA requires including a description of the implementation of human oversight measures, according to the instructions of use, and the measures to be taken in case of the materialization of these risks, including their arrangements for internal governance and complaint mechanisms. See **Appendix II** for examples of safeguards and for examples of additional mitigators Do the benefits of the processing outweigh the risks to the data subject, large groups of individuals, and society, as mitigated by the safeguards? A FRIA requires assessing the specific risks of harm likely to impact the categories of persons or group of persons identified above, considering the information given by the provider pursuant to Article 13 of the EU AI Act.

	Legitimate interests of the controller includes the relevant and appropriate relationship between the data subject and the controller in situations such as where the data subject is a client or in the service of the controller. The existence of a legitimate interest needs careful assessment of whether the data subject can reasonably expect at the time and in the context of the collection of the personal data that processing for that purpose may take place. According to the Draft CNIL LI How-To Sheet, the greater the anticipated benefits of the processing, the more likely the legitimate interest of the controller is to prevail over the rights and freedoms of individuals. Consideration should also be given to the objective evidence that can inform this element of the assessment, for example evidence of benefits seen in previous pilots or tests, evidence from customer surveys of focus groups about expectations. See **Appendix II** for a list of processing with regards to the development of AI systems likely to exceed the reasonable expectations of individuals.

Interests or fundamental rights and freedoms of the data subject, large groups of individuals, society and others which require protection of personal data. **Detailed Analysis**		
Benefits/Interests	**Risks/Harms**	**Risk Mitigations**
Data Subject (add sources where appropriate)		

Large Groups of Individuals (add sources where appropriate)		
Society (add sources where appropriate)		
Other (add sources where appropriate)		
10. Summary of Weighted Balance (Include Heat Maps or other "Weighing" criteria)	A description of how the benefits of the processing outweigh the risks identified, as mitigated by the safeguards identified. This should include as many "evidence" points as to expand on the conclusion and analysis. According to the EDPB guidance for consultation, the purpose of the balancing exercise is not to avoid any impact on the interests and rights of the data subjects altogether; rather, its purpose is to avoid a disproportionate impact and to assess the weight of these aspects in relation to each other.	

	LIAs MUST IDENTIFY AND WEIGH THE BENEFITS THAT MAY FLOW, DIRECTLY AND INDIRECTLY, FROM THE PROCESSING TO THE CONTROLLER, THE DATA SUBJECT, LARGE GROUPS OF INDIVIDUALS AND SOCIETY AGAINST THE POTENTIAL RISKS TO THE RIGHTS OF THE DATA SUBJECT WHICH REQUIRE PROTECTION OF PERSONAL DATA, AS MITIGATED BY SAFEGUARDS THAT THE CONTROLLER CAN EMPLOY TO REDUCE THESE RISKS. THE CONTROLLER MUST FACTOR INTO THIS LIA THE USE OF DE-IDENTIFIED DATA AND THE RESONABLE EXPECTATIONS OF THE DATA SUBJECT, AS WELL AS THE CONTEXT OF THE PROCESSING AND THE RELATIONSHIP BETWEEN THE CONTROLLER AND THE DATA SUBJECT WHOSE PERSONAL DATA NEEDS PROTECTION. This narrative summary could be just a "narrative" or could also include some pictorial/numerical examples of the Multi-Dimensional/Stakeholder Balancing (see Section 13 for examples of this type of balancing).
11. Decision – Go/No Go (Approver)	How effective are the mechanisms that facilitate the processing auditability (e.g., traceability of the development process, the sourcing of training data and the logging of the processing system's processes, outcomes, positive and negative impacts)? Could the ADM be audited by independent third parties? What are the additional requirements surrounding the processing? Are there other legal, cross-border, policy, contractual, industry or other obligations linked to the collection, analysis, and use(s) of personal data? Have these all been addressed? If the processing involves ADM, does the assessment effectively address the reasonably foreseeable risk of: 1. Unfair or deceptive treatment of, or unlawful disparate impact on, the data subject; 2. Financial or physical injury to the data subject; 3. A physical or other intrusion upon the solitude or seclusion, or private affairs or concerns, of the data subject if the intrusion would be offensive to a reasonable person; or 4. Other substantial injury to the data subject. Is it foreseeable that the potential application of ADM might seem surprising, inappropriate, or discriminatory or might be considered offensive causing distress or humiliation?

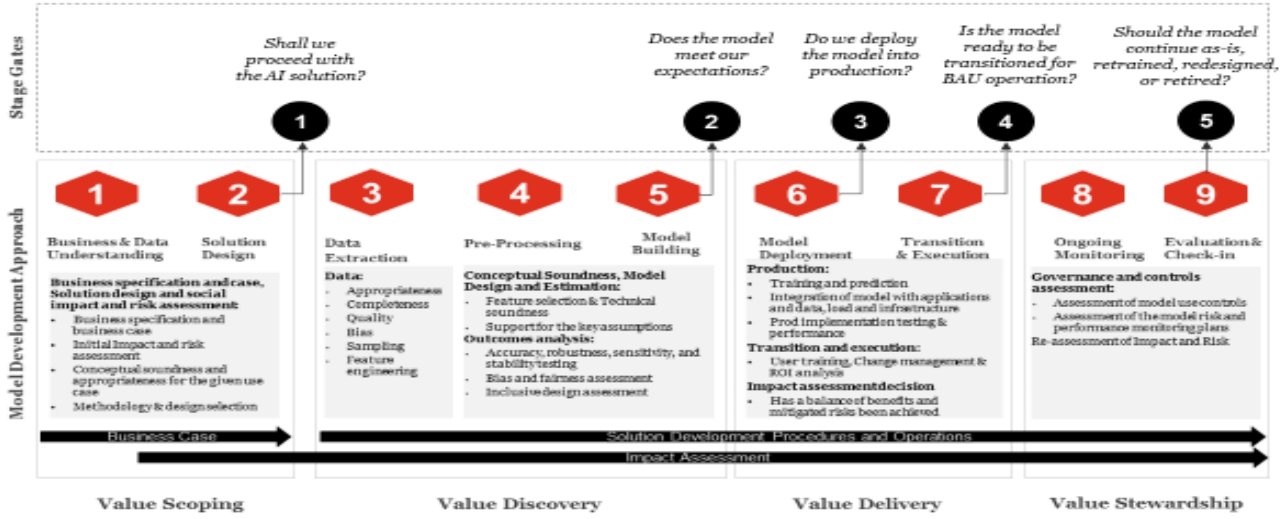
	How effective are the overall controls and safeguards in reducing risk? Describe with specificity how and why it was determined that the negative impacts do or do not outweigh the benefits, including how any specific safeguards identified in 10.C. affect this LIA. Identify who took part in the completion of the LIA and their roles. Identify external parties that were involved, what their roles were, what their input was, and what the response to that input was. If no external input was obtained, explain why and explain how that function was replaced. . **Decision** – Do benefits and mitigated risks support proceeding with the processing? Are there any other factors that should be considered? Have the interests, expectations and rights of stakeholders been effectively addressed? What evidence supports the weighing of impacts and benefits? Are potential risks to the data subject sufficiently mitigated? What level of confidence is there that mitigations will be effective? **If the iterative analysis done as part of the project lifecycle determines this is a new or evolved "purpose," the processing may be able to continue under legitimate interests if the new purpose is compatible with the original purpose. If it is not compatible, a new LIA should be performed and documented.** If the DPA were to review a copy of the LIA, is the organisation comfortable in sharing it? Keep a record/inventory of decisions (both Go/No Go decisions). This record/inventory builds up precedent on how similar situations have been decided in the past and is a resource tool to guide future decisions.
12.Signatories and Date	**Names, Titles, and Signatures of all Decision Makers (persons who make risk tolerance decisions within the organisation) and Contributors to the LIA:** **Relevant internal actors and external parties contributing to the LIA, including any internal or external committees** (The business must make sure that relevant individuals prepare, contribute to, or review the LIA, based upon their level of involvement in the processing activity

	that is subject to the LIA. Relevant individuals are those whose job duties pertain to the processing activity. These individuals must make good faith efforts to disclose all facts necessary to conduct the LIA and must not misrepresent in any manner any fact necessary to conduct the LIA. A LIA may involve external parties to identify, assess, and mitigate the risks.): **Any internal or external audit conducted in relation to the LIA, including the name of the auditor, the names and positions of individuals involved in the review process, and the details of the audit process.** **Dates the LIA was reviewed and approved, and names, positions, and signatures of the individuals responsible for the review and approval of the LIA** (Individuals responsible for the review and approval must include the individual who decides whether the organisation will initiate the processing that is the subject of the LIA. If the organisation presented or summarized LIA to the organisation's board of directors or governing body for review, or if no such board or equivalent body exists, to the organisation's highest-ranking executive who is responsible for oversight of LIA compliance for review, the organisation must include the date of that review).

13. Making a LIA and Benefits/Risk and Mitigations Operational in an AI World

Each organisation has its own way of making requirements operational, whether they be legally driven or internally policy driven. These operational methods are usually developed and designed consistent with the specific organisation's culture and rhythm of business. As AI grows as a part of business objectives and strategy, organisations will need to expand their own risk assessment and management processes. For example, an assessment process should cover the full development life cycle requirements from strategy and planning, model development, the specific issues related to training data, deployment, ongoing operation and monitoring issues, and governance. These requirements were outlined in the IAF's AI Assessment which included an example of a five-stage gate review to successful AI governance.1 Each stage gate review is designed to address specific questions and should involve a broad set of stakeholders and decision makers:

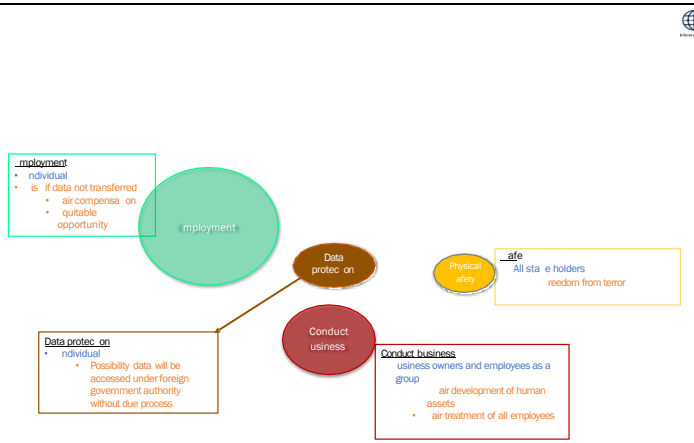
[1] Modified on original Six stage gates to a successful AI governance | by Anand Rao | Towards Data Science



By extension, a LIA should encompass a review at key parts of a project's development lifecycle. In effect, the assessment process required by a LIA should be iterative and matched to the cadence of an organisation's development process.

Formalizing an Assessment of Net Benefits and Risks

After a detailed analysis and narrative summary of all stakeholder's benefits, risks and risk mitigation tactics that involves multiple internal relevant actors as outlined in the above assessment model, many organisations can utilize a common risk assessment methodology to test or validate or support their conclusion, this could include use in stakeholder panels or data governance related committees within the organization. The IAF has developed **two examples of Multi-Dimensional/Stakeholder Balancing Output** – Both use a math determined balancing that aligns with a common risk management methodology (weighing), but one uses a mathematical depiction and the other one uses a pictorial depiction of the balancing (see [A-Principled-Approach-to-Rights-and-Interest-Balancing.pdf (secureserver.net)]) for more details).



The mathematical model uses a process common to many risk management programs and utilizes a defined scale. The answers to the five scale questions then are computed to reflect the "projected net benefit/risk" result. This form of mathematical assessment usually is best performed as an output of the internal stakeholder evaluation of the multi-stakeholder benefits, risks and mitigations assessment and can be used to support the narrative risk/benefit analysis.

There are many approaches to this process. While individual risks could be evaluated against a common risk approach – using a 5-point scale (see below) likelihood times significance of risk (Inherent risk) minus effectiveness of controls = residual risk. Another approach is to evaluate each stakeholder and to include the "Benefits". This is determined using a similar 5-point scale and evaluating the likelihood and significance of the benefit and then subtracting the likelihood and significance of risk minus the effectiveness of mitigating tactics. The mathematical computation produces a the "projected net benefit/risk" score than can be reflected in a heat map or pictorial depiction

Scenario

Case Desciption	Internal FI Sharing Blockers
Stakeholder Type	Employee

Assessment Date:
2024-08-03 10:02

5 How significant is the benefit? (Pick one option from dropdown)
[1- Low Impact, 2 - Moderately Low impact, 3 - Moderate impact, 4 - Moderately high Impact, 5 - High impact]

1- Low Impact

6 Are the benefits likely to occur? How likely? (Pick one option from dropdown)
[1 - Slightly, 2 - Not likely, 3 - Likely, 4 - Highly Likely, 5 - Expected]

1 - Slightly

11 How significant is the risk? (Pick one option from dropdown)
[1- Low Impact, 2 - Moderately Low impact, 3 - Moderate impact, 4 - Moderately high Impact, 5 - High impact]

5 - High impact

13 How likely is the risk to be realized? (Pick one option from dropdown)
[1 - Slightly, 2 - Not likely, 3 - Likely, 4 - Highly Likely, 5 - Expected]

4 - Highly Likely

18 How effective are these controls and safeguards in reducing risk ? (Pick one option from dropdown)
[1 - Low Effectiveness, 2 - Moderately Low Effectiveness, 3 - Moderately Effective, 4 - Moderately High Effectiveness, 5 - High Effectiveness]

1 - Low Effectiveness

14. Assumptions/Caveats/Implications

- The organisation has in place a privacy impact assessment (PIA) and a data protection impact assessment (DPIA) (for higher risk processing) process that helps the organisation determine whether the requirements in the GDPR related to data subject rights (e.g., Transparency, Consent, Access, Portability) are sufficient.
- The core elements of an LIA (e.g., purpose, the risk impact and balancing content, information on data, controls, risk mitigation, necessity, and proportionality) are sourced from an established assessment process, OR all questions in

the LIA are output content, meaning production of an actual LIA happens if needed.

- The LIA is added to the organisation's already existing assessment processes, OR key aspects of the existing assessment are issued to supplement and/or create a separate LIA.
- The organisation has controls in place to monitor how the LIA is working in practice, particularly where LIAs are devolved to teams outside of the main privacy or DPO team.
- Organisations use a set of "triggering" questions to determine if a LIA is needed, based on how the legitimacy of data processing is determined. Practically speaking, ALL projects likely go through an initial risk assessment to determine if a LIA likely is required. This process reinforces the likelihood of designing an iterative, multistage assessment more aligned with an AI development lifecycle. The traditional linear or step process of assessing "purpose," then "necessity," and finally "balancing of rights" is not well suited to today's data driven business models.
- The explicit, implied aspects of an LIA coupled with issues that are not defined suggest the need for enhanced governance controls processes. Many of these are more associated with responsible AI governance.
- The **Data Protection Network (DPN**) have recently updated their established Legitimate Interests Guidance and Assessment, first published in 2017. The guidance helps organisations to assess whether they can rely on Legitimate Interests for a range of processing activities, both routine and bespoke. It includes core considerations, useful case studies and examples, plus a refreshed LIA template (in Excel) to use when conducting a Legitimate Interests Assessment.

15. Definitions and Explanations

A. "Plain language" means:
   1. Degree and details of any human involvement in the controller's use of ADM:
      a. Identifying who at the controller will be responsible for the use of the ADM and for what they are responsible
      b. Identifying and describing the human's qualifications, if any, to understand the controller's use of the ADM, including the personal data processed by, and the logic output(s) of the ADM
      c. Explaining whether and, if so, how the human evaluates the appropriateness of the personal data processed by, and the logic and output(s) of the ADM for the controller's proposed use(s)
      d. Explaining whether the human has the authority to influence whether or how the controller uses the output(s) of the ADM and, if so, how they exercise this authority
      e. If the human has the authority to influence whether or how the controller uses the output(s) of the ADM and, if so, how they exercise this authority

f. If a human is not involved in the controller's use of ADM, explaining why there is no human involvement, and which safeguards the controller has implemented to address the risks to data subjects' privacy that may arise from the lack of human involvement

2. How the controller evaluates its use of ADM for validity (confirmation that the ADM, including its input(s), performs as intended for the controller's proposed use(s), included the ADM's accuracy in performing as intended), reliability (ability of the ADM to perform as intended for the controller's proposed use(s), repeatedly and without failure, under time interval(s) and conditions consistent with the controller's proposed use(s), and fairness (equality, equity, and avoidance of discrimination harms)):
    a. The metrics the controller uses to measure performance, validity, reliability, fairness, known limitations, and disparate impact, and why the metrics selected are appropriate measures of validity, reliability, fairness and disparate impact
    b. If the controller uses data, hardware, software or other technical components provided by another person, including ADM, the controller must identify the name(s) of the person(s), the name(s) of the technological component(s) provided, and how the controller ensures the technological component(s) provided do not negatively impact the validity, reliability, fairness, or discriminatory impact of the controller's use of ADM
        i. This explanation shall include any copies of internal or external evaluations related to the technical component's validity, reliability, fairness, or discriminatory impact provided to or conducted by the controller
    c. Whether and, if so, how the controller evaluated other versions of the ADM or other ADM for validity, reliability, fairness, or discriminatory impact for the controller's proposed use(s)
    d. If the controller evaluated other versions of the ADM for validity, fairness, reliability, or discriminatory impact for the controller's proposed use(s), why the controller did not use the other versions of ADMs
3. The results of the controller's evaluations

B. "Quality of personal data" includes completeness, representativeness, timeliness, validity, accuracy, consistency, and reliability of the sources of the personal data for the organisation's proposed use of the ADM. Actions an organisation may take to ensure quality of personal data include: (1) identifying the source of the personal data and whether that source is reliable (or, if known, whether the original source of the personal data is reliable); (2) identifying how the personal data is relevant to the task being automated and how it is expected to be useful for the development, testing, and operation of the ADM; (3) identifying whether the personal data contains sufficient breadth to address the range of real world inputs the ADM may encounter; and (4) identifying how errors from data entry, machine processing, or other sources are measured and limited.
C. "Individual interests" are one of more interests of a single natural person. Smuha
D. "Group interests" are one or more interests of a collective or group of individuals. Just as a collective consists of the

sum of individuals, so do group interests consist of the sum of interests of the individual members of the collective. Smuha

E. "Societal interests" are one or more interests of society. In contrast with individual and collective interests, societal interests thus are not concerned with the interests of a particular individual or the interests shared by a collective of individuals. Instead, they concern an interest held by society at large, going over and above the sum of individual interests. Smuha

Appendix I – New Data Processing Activity

A new data processing activity is generated when existing processing is modified in a way that materially changes the level of risk presented. For example:

- The way that existing systems or processes handle personal data
- Processing purpose
- Personal data processed or sources of data subjects' personal data
- Method of collection of data subjects' personal data
- Personal data recipients
- Intentional and substantial modifications to an ADM
- Processor roles or processors
- Algorithm applied or algorithmic result
- Software or other systems used for processing
- The purpose of processing data subjects' personal data
- Data subjects' reasonable expectations concerning the purpose for processing their personal data or the purpose's compatibility with the context in which their personal data was collected
- The minimum personal data that is necessary to achieve the purpose of the processing
- The operational elements of the processing
- The benefits resulting from the processing to the controller, the data subject, other stakeholders, and the public
- The negative impacts to data subjects' privacy associated with the processing, including the sources of these negative impacts
- The safeguards that the controller has implemented or plans to implement to address the identified negative impacts
- The controller's assessment of whether the identified negative impacts identified, as mitigated by the safeguards, outweigh the identified benefits

- Why the controller is using or seeks to use ADM to achieve the purpose of the processing
- The output(s) secured from the ADM and how the controller will use the output(s)
- The steps the controller has taken or any steps it plans to take to maintain the quality of personal data processed by the ADM, including personal data used by the controller to train the ADM
- The logic of the ADM and the assumptions of the ADM's logic
- How the controller evaluates its use of the ADM for validity, reliability, and fairness
- The degree and details of human involvement in the controller's use of ADM
- The safeguards that the controller plans to implement to address the negative impacts to data subjects' privacy that are specific to its use of ADM or for data sets produced by or derived from the ADM

Appendix II – Stakeholders, Interests, Benefits, Risks, and Mitigations

A. **Stakeholders** – Examples of stakeholders are:

- Project manager, sponsor and team
- The customer (individual or organisation)
- Suppliers of material or other resources
- Creditors
- Employees
- Trade Unions
- Workers Councils
- City, community, or another geographical region
- Professional organisations
- Civil society organizations
- Any data subject or group impacted by processing
- Any data subject or group in a position to support or undermine the processing
- Internal or external, local or international

Stakeholders may be looked at organisationally, geographically, or by involvement with various processing phases or outcomes. Another way of determining stakeholders is to identify those who are directly impacted by the processing and those who may be indirectly affected.

Stakeholders should be grouped by geographic region, organisation, processing involvement, or whether they are directly or indirectly impacted. This analysis enables a close look at each stakeholder to gather more in-depth information to understand their impact, involvement, communication, requirements, and preferences.
In taking the public's or society's view into account, ask representatives of the public/society:

- Customers/clients/communities
- Employees
- Civil society/unions/consumer groups
- Community leaders/democratic representatives
- Marginalised/minoritised/vulnerable

Also consider adding lay members to boards and institutionalising public panels
Consider conducting Consequence Scanning (CS). In CS, three questions about the organisation's product are answered during roadmap, planning and feature creation:

- What are the intended and unintended consequences of this product or feature?

- What are the positive consequences the organisation wants to focus on?
- What are the consequences the organisation wants to mitigate

B. Benefits - What are the benefits to the data subject, large groups of individuals, and society? Could the use of this data be used in a way that could result in a specific group of individuals being treated differently in a positive way from other groups of individuals? Can the benefits obtained by various stakeholders be measured? Determine and describe the positive impacts on the various stakeholders that are expected to come from the application of this technology/data activity. Determine what the potential positive goal of the difference in treatment is (if any). Are areas of interest, such as integrity of the person, autonomy, respect for private life, liberty and security, or education access, affected in a positive way?

What is the benefit to the **organisation**? The EDPB guidance for consultation states that the "interest" is the broader stake or **benefit** that the controller or third party has in engaging in a specific processing activity**.** There is no exhaustive list of interests that may be considered as being legitimate. A wide range of interests is, in principle, capable of being regarded as legitimate a building, improving a product, and assessing the creditworthiness of individuals. Examples from cases of benefits to the organisation cited in the EDPB guidance for consultation are: having access to information online, ensuring the continued functioning of publicly accessible websites, obtaining the personal information of a person who damaged someone's property in order to sue that person for damages, protecting the property, health and life of the co-owners of a building, improving a product, and assessing the creditworthiness of individuals. Examples of benefits to third parties cited in the EDPB guidance for consultation are: establishment, exercise of defence of legal claims, disclosure of data for purposes of transparency and accountability, historical or other kinds of scientific research, public interest or third party's interest.

The Draft CNIL LI How-To Sheet lists the following interests as a priori legitimate for the development of AI Systems:

- Carry out scientific research (in particular for bodies which cannot rely on the public interest mission);
- Facilitate public access to certain information;
- Develop new systems and functionalities for users of a service;
- Offer the service of a conversational agent to assist users;
- Improve a product or service to increase its performance;
- Develop an AI system to detect fraudulent content or behaviour.

See the ICO's Consultation on Generative AI for a draft discussion of the use of LI as the lawful basis for training generative AI models on web-scraped data.

The UK Data (Use and Access) Bill lists the following as Recognised Legitimate Interests:

- Disclosure for purposes of processing described in UK GDPR Article 6(1)(e)
- National security, public security and defence
- Emergencies
- Crime detecting, investigating, preventing, apprehending, or prosecuting
- Safeguarding vulnerable individuals

The EU Charter recognizes the freedom to conduct a business which includes the right to produce products and to provide services which are of value to a data subject, groups of individuals, and society. Implicit in the fact that the EU AI Act requires the conduct of a FRIA means that operators deploying high-risk AI systems have such rights too. Examples of benefits to the organisation include; improved profitability, enhanced employee satisfaction, engagement and productivity, enhanced customer relationship, undamaged brand/reputation, enhanced brand/reputation, increased market share, prevention of cyber-crime and fraud, new/improved/innovative products/services, improved customer service. Examples of LIs given in recitals to the GDPR are fraud and direct marketing. Examples of benefits given in the Draft CNIL LI How-To Sheet are: improved healthcare, better accessibility of certain essential services, facilitation of the exercise of fundamental rights such as access to information, freedom of expression, access to education. According to the Draft CNIL LI Open Source Models Focus Sheet, open source in AI can have significant benefits for the controller: it allows them to leverage community contributions or enhance the attractiveness of their models by facilitating its adoption by other stakeholders, and it brings numerous advantages to research activities and scientific innovation (e.g. fosters knowledge sharing among developers in the open source community, improves accessibility for students, encourages the design and publication of related tools using open source models, and promotes the harmonization of practices and interoperability of models and systems). The benefits of open-source models, which may be considered in the assessment of the legitimate interest of the controller in the development phase, may:

- Strengthen the legitimate interest of data controllers with regard to the benefits for scientific research and to guarantee the quality of the model against risks of illegality in the use phase, in particular with regard to discrimination where the model contains biases;
- Provide an additional guarantee for processing operations in the development phase, where it has benefits in terms of transparency, accountability, or peer review.

Are there benefits for **society as a whole**? Consider factors such as increased revenue, lower costs, improved efficiency, enhanced employee satisfaction, engagement and productivity, enhanced citizen (or workforce) relationship, enhancement or maintenance of brand or reputation, assurance of compliance, fraud prevention, enhancement or maintenance of cyber or physical security, new or improved public services or citizen service, improved manner of marketing, improved ability to assess customer preferences, improvements to innovation or enabling greater, faster, more efficient innovation, improved research processes, improved ability to conduct research and find or enroll study subjects, or improved efficiency with studies, innovative ways to conduct research, better health care, improved education, positive impact on climate change, more accessible/usable technology, protection of reasonable expectation of privacy (including anonymity), protection of freedom of religion/thought/speech, protection of prohibition against discrimination on basis of race/national or ethnic origin/colour/religion/age/sex/sexual orientation/marital status/disability. Do the benefits of having the model and/or data use in production outweigh the costs of maintaining it? Are there any social interests served with the deployment of the processing? How does the processing contribute to or increase well-being? How will the processing contribute to human values?

Examples of benefits to **individuals** are: more objective outcomes, safer interactions, better product selection, better access to new products and services, significant discounts, better product utilization, improved service, improved ease of use, engaged consumers/customers/employees, more convenience, appropriately linked to other choices, anticipating or meeting of a need, exercise of self-determination, public sector access, anonymous transportation. According to the Draft CNIL LI Open Source Models Focus Sheet, open source can have benefits for individuals whose data is used in the development phase, for model users, or for those whose data is used in the deployment phase:

- Increases the transparency of the model and its functioning, thereby facilitating individuals' exercise of their rights
- Enables verification of the model's capabilities and limitations
- Facilitates the verification and detection of biases to reduce or correct them
- Facilitates the detection of vulnerabilities in the model to improve its security

C. **Interests** - Interests or fundamental rights and freedoms of the data subject which require protection of personal data

Under the UN Charter, some of the rights to be considered are:

- Right to the security of person
- Right to own property

- The right to work, to free choice of employment, to just and favourble work conditions, and to protection against unemployment.
- The right to, without any discrimination, equal pay for equal work.
- The right to just and favourable remuneration.
- The right to reasonable limitation of working hours.
- The right to a standard of living adequate for the health and well-being of himself and of his family.
- The right to education

Under the EU Charter, some of the fundamental rights and freedoms to consider are:

- The right to be treated with dignity
- The right to security
- Respect for private life
- Protection of personal data
- Freedom to conduct a business
- Right to property
- Non-discrimination
- Cultural, religious, and linguistic diversity
- Equality between men and women
- Rights of the child
- Rights of the elderly
- Persons with disability
- Protection of pregnant workers and parents on maternity or parental leave

Examples of interests given in the Draft CNIL LI How-To Sheet are freedom of expression, freedom of information, freedom of conscience. The Draft CNIL LI How-To Sheet also provides that the following interests could be considered a priori legitimate for the development of AI systems:

- Carry out scientific research (in particular for bodies which cannot rely on the public interest mission);
- Facilitate public access to certain information;
- Develop new systems and functionalities for users of a service;
- Offer the service of a conversational agent to assist users;

- Improve a product o to increase its performance;
- Develop an AI system to detect fraudulent content or behaviour.

Conversely, certain interests cannot be regarded as legitimate, where the AI system envisaged has no connection with the mission and activity of the organisation or where it cannot be lawfully deployed.

D. **Risks** - Considering all the factors relating to the data, the metric or measure, the likely use, the associated processing, the identifiability and sensitivity of the data and its use, what are the risks (real and/or perceived) to the identified stakeholders/users? Could any metric of measure be used in a way that decides on a specific user and/or creates a profile on them (real or perceived)?

Specific risks to consider include:

- Constitutional harms, such as chilling or deterring data subjects' free speech or expression, political participation, religious activity, free association, freedom of belief, freedom to explore ideas, or reproductive freedom; and harms to data subjects' ability to engage in collective action or impede the right to unionize;
- Intellectual privacy harms, such as the creation of negative inferences about a data subject based on what an individual data subject reads, learns, or debates and using information inferred about consumers to manipulate them;
- Negative impacts to data subjects' data security, such as unauthorized access, destruction, use, modification, or disclosure of personal data or unauthorized activity resulting in the loss of availability of personal data;
- Discrimination harms, such as a violation of antidiscrimination laws;
- Unfair, unconscionable, or deceptive treatment;
- A negative outcome or decision with respect to an individual data subject's eligibility for a right, privilege, or benefit related to financial or lending services, housing, insurance, education enrollment or opportunity, criminal justice, employment opportunities, health-care services, or access to essential goods or services;
- Financial injury or economic harm, including limiting or depriving data subjects of economic opportunities; charging data subjects higher prices; compensating data subjects at lower rates; or imposing additional costs upon data subjects, including costs associated with the unauthorized access to data subjects' personal data;
- Physical injury (including processing that creates the opportunity for physical or sexual violence), harassment, or threat to an individual data subject or property;
- Privacy harms, such as physical or other intrusion upon the solitude or seclusion of the private affairs or concerns of data subjects, stigmatization or reputational injury, infer highly sensitive, latent information from seemingly innocuous surface data;

- Psychological harm, including anxiety, embarrassment, fear, emotional distress, stress, frustration, shame, feelings of violation, and other mental trauma;
- Other detrimental or negative consequences that affect an individual data subject's private life, private affairs, private family matters or similar concerns, including actions and communications within an individual data subject's home or similar physical, online, or digital location, where an individual data subject has a reasonable expectation that personal data or other data will not be collected, observed, or used;
- Inconvenience or expenditure of time;
- Disruption and intrusion from unwanted commercial communications or contacts
- Loss of autonomy through acts or practices that are not reasonably foreseeable by an individual data subject and that are intended to materially:
    - Alter that data subject's experiences
    - Limit that data subject's choices
    - Influence that data subject's responses, or
    - Predetermine results or outcomes for that data subject;
- Impairing data subjects' control over their personal data, such as by providing insufficient information for data subjects to make an informed decision regarding the processing of their personal data or by interfering with data subjects' ability to make choices consistent with their reasonable expectations or by giving controllers far more insight into their customers than their customers have into them, thereby producing a power imbalance between the controllers and their customers;
- Coercing or compelling data subjects into allowing the processing of their personal data, such as by conditioning data subjects' acquisition or use of an online service upon their disclosure of personal data that is unnecessary to the expected functionality of the service, or requiring data subjects to consent to processing when such consent cannot be freely given, or feeling pressure to conform to behaviours that data subjects think will please the algorithmic decisionmaker;
- Exploiting data subjects' vulnerabilities, such as age, employment or student status, immigration status, health status, or financial hardship;
- Producing erroneous models and predictions, as a result of inaccurate data or faulty algorithms, that can negatively impact data subjects;
- Creating opacity and procedural unfairness because most data subjects lack an understanding of, and opportunities to challenge, the controller's algorithmic determinations that can shape their life opportunities;
- Increasing automation which, in turn, displaces human labor; or
- Using analytics for intentional, harmful purposes.

Consider what factors about the processing have the highest impact on the likelihood any of these risks could be realized.

The Draft CNIL LI How-To Sheet states that the following impacts on people should be considered and the level of associated risks should be assessed in the development of AI systems:

1. Impacts on individuals related to the collection of data used to develop the system, where data have been scraped online
    a. Risks of infringement of privacy and rights guaranteed by the GDPR safeguards.
    b. Risks of illegal collection
    c. Risks of undermining freedom of expression

2. Impacts on individuals related to model training and data retention
    a. Risks of loss of confidentiality of the data contained in the dataset or in the model
    b. Risks related to the difficulty of ensuring the effectiveness of the data subject rights
    c. Risks associated with the difficulty of ensuring transparency towards data subjects

3. Impacts on persons at the training stage related to the use of the AI system
    a. Risks of memorization, regurgitation or generation of personal data when using certain AI system which may invade privacy
    b. Risks of reputational damage, spread of false information or identity theft where the AI system (particularly generative AI) produces content on an identified or identifiable natural person
    c. Risks of infringement of certain rights or secrets provided for by law in the event of memorization or regurgitation of protected data
    d. Serious ethical risks, which may impact certain general legal principles or the proper functioning of society as a whole related to the development of certain AI systems (e.g. discrimination, the safety of people in case of malicious use, incitement to hatred or violence, disinformation which may undermine the rights and freedoms of individuals or democracy and the rule of law, amplification of discriminatory biases in the training database, lack of transparency or explainability, lack of robustness or automation biases)

The Irish Data Protection Commissioner in its blog AI, Large Language Models and Data Protection (DPC Blog) list the following risks that people should be aware of:

- Use of large amounts of personal data during the training phase, sometimes unnecessarily and without the individual's knowledge, agreement or permission
- Issues for the individual or other arising from the accuracy or retention of personal data used (or generated)
- Issues for the individual if models based on the individual's data are shared with others for purposes the individual is not aware of or does not agree with or if the data are not properly secured
- Inaccurate or incomplete training data may cause biases in AI systems for the individual or for others that lead to decisions that affect the individual's rights in some way

- Where new personal data becomes part of training data for new fine-tuned versions of a model, the individual or others may be exposed to these kinds of risks where they were not before

With regards to the development of AI systems, according to the Draft CNIL LI How-To Sheet, certain processing is likely to exceed the reasonable expectations of individuals:

- Re-use of data published on the internet
- The variety of the potential uses of a dataset or a model

According to the Draft CNIL LI Open Source Models Focus Sheet, open source can present potentially significant risks:

- The unlawful or malicious reuse of the model
- Model security (e.g. risks of exploitation by attackers of the model's vulnerabilities, made apparent, risks in terms of traceability, risks of misuse of open-source models given the increased ease of removing or bypassing filters and security measures added to the systems)
- Where the model contains personal data or where the training data also is published:
    - Confidentiality of training data
    - Difficulties in ensuring the exercise of people's rights throughout the chain of use and redistribution of the model

According to the ICO's Consultation on Generative AI, which has not been finalized, individuals whose data is scraped for generative AI development can experience harm, either related to the collection of the training data or because of the use of the generative AI model. These harms can manifest in two ways:

- Upstream risks and harms: For example, people may lose control over their personal data, as they are not informed of its processing and therefore are prevented from exercising their information rights or evaluating the impact of that processing on them, including its fairness.
- Downstream risks and harms: For example, generative AI models can be used to generate inaccurate information about people resulting in distress or reputational harm, be used by hackers deploying social engineering tactics to generate phishing emails tailored to individuals or undertake other adversarial attacks.

Examples of risks to large groups and society:

- Overuse of technology (e.g., social media by teenagers)
- Utilization of disinformation and propaganda

- Economic and asset inequalities – Who will and will not have access to technology
- Utilization of technology by surveillance state
- Utilization of online tools by hateful and criminal actors- intentional, harmful use
- Algorithmic bias in technology against protected classes
- Who controls data subject data and who monetizes it – controllers have far more insight into their customers than their customers have into them producing a power imbalance between controllers and customers
- Loss of data subject trust
- Invasion of privacy – Inferring from innocuous surface data highly sensitive latent information about data subjects
- Manipulation – Using information inferred about data subjects to manipulate them
- Opacity and procedural unfairness – Most data subjects lack an understanding of, and opportunities to challenge, controllers' algorithmic determinations that can shape their life opportunities
- Displacement of labor – Increased automation can displace human labour
- Pressure to conform – Data subjects may feel pressure to conform to behaviours that they think will please algorithmic decisionmakers

The DPC Blog identifies risks to the organisation using AI systems:

- Unwanted, unneeded or unanticipated processing of personal data input to or used to train or fine tune an AI model may impact or involve several principles of the GDPR (e.g. lawfulness, fairness and transparency principle and purpose limitation principle)
- Inability to facilitate the exercise of data subject rights related to the engagement with AI products, especially if the organisation is inputting its own or others' personal data and needs to know where it is going or how it is being processed
- Where the organisation is using an AI product supplied by a third party, it may result in additional security or other data protection risks arising from the use of personal data that the organisation's employees or its data subjects input into the AI tool because the organisation does not fully understand both how personal data is protected when it processes the data and where it instructs another organisation to do so on its behalf
- The organisation unexpectedly or unnecessarily may be further processing data related to identifiable people when some AI models have inherent risks relating to the way in which they respond to inputs or "prompts," such as memorization, which can cause passages of (personal) training data to be unintentionally regurgitated by the product,
- The organisation needs to understand how to mitigate the effects when AI products rely on a process of "filtering" to prevent certain types of data (e.g. personal data, inappropriate data, and copyright data) to be provided to a user in response to a query or prompt and those filters are attacked or circumvented to cause that data to be made available to be processed in unintended, unauthorized, insecure or risky ways

- The organisation may be introducing "automated decision making" risks if the outputs of AI products, which can be prone to producing inaccurate or biased information, are relied upon without critical human analysis or intervention
- The organisation needs to protect personal data published on the organisation's website, either by its staff of its website users, from being collected and used for AI training or other processing where the staff or users have not already agreed to that purpose or if they do not have a reasonable expectation the data will be used for AI training

E. Safeguards/Additional Mitigators:

Compensatory or additional measures to limit the impact of processing on data subjects may be of a technical, organizational, or legal nature and must be able to limit the risk of harm to the interests, rights, and freedoms. At a minimum, the implementation of the following safeguards should be considered:

- Safeguards to protect personal data, such as encryption, segmentation of information systems, physical and logical access controls, change management, network monitoring and defences, and data and integrity monitoring;
- Use of privacy-enhancing techniques such as trusted execution environments, federated learning, homomorphic encryption, and differential privacy;
- Consulting external parties, e.g., experts in detecting and mitigating bias in ADM, to make sure that the organisation maintains current knowledge of emergent privacy risks and countermeasures and using that knowledge to identify, assess, and mitigate risks;
- Exploring the need for human involvement as part of the organisation's use of ADM and implementing policies, procedures, and training to address the degree and details of human involvement as necessary in that evaluation;
- Restrictions on the processing of personal data;
- Data subject/user testing or research

Some additional mitigators to consider are:

a. The use of de-identified data;
b. Measures taken pursuant to the controller duties, including an overview of data security practices the controller has implemented, any data security assessments that have been completed, and any measures taken to comply with the consent requirements of the GDPR;
c. Measures taken to ensure that data subjects have access to the rights provided in the GDPR;
d. Contractual agreements in place to make sure that personal data in the possession of a processor or other third party remains secure; or
e. Any other practices, policies, or trainings intended to mitigate processing risks.

If ADM is being assessed, the organisation must identify the following:

1) Safeguards used to reduce the risk of harms identified and safeguards for any data sets produced by or derived from the ADM (include a description of any other practices, policies, or trainings intended to mitigate processing risks)
2) Whether ADM was evaluated to make sure it works as intended for the organisation's proposed use and does not discriminate against protected classes
3) The policies, procedures, and training the organisation has implemented or plans to implement to make sure that the ADM works as intended for the organisation's proposed use and does not discriminate against protected classes
4) Where the organisation obtains the ADM from another person, the organisation must identify the following:
    a) Whether it reviewed that person's evaluation of the ADM and whether that person's evaluation included any requirements or limitations relevant to the organisations proposed use of the ADM
    b) Any accuracy and nondiscrimination safeguards that it implemented or plans to implement

In addition, are there any technical and/or procedural safeguards (mitigating controls) that could be implemented to prevent and mitigate risks should they occur (e.g., increased transparency, additional suggestions/guidance to the data subject, more choice, etc.)?

The following. Measures have been identified in the Draft CNIL LI How-To Sheet as relevant to limit the impact on data subject rights and freedoms (they must be adapted to the risks posed by the different processing of the development phase);

1. In response to the risks associated with the collection and compilation of the dataset
    a. Anonymise at short notice or, failing that, pseudonymize the data collected
    b. Where it does not adversely affect the performance of the model developed, synthetic data should be used
2. In response to risks related to model training and data retention
    a. Provide a discretionary right to object and, when collecting the data directly from data subjects, before the processing takes place
    b. Implement measures that ensure and facilitate the exercise of individuals' rights when the model is subject to the GDPR
    c. Where the model is shared or disseminated in open source, identify and implement measures to ensure the transmission of the exercise of rights through the chain of stakeholders
    d. Ensure greater transparency regarding the processing carried out for the development or

improvement of the AI system
   e. Implement measures and procedures to ensure a transparent development of the AI system, to enable auditability during the deployment phase
   f. Ensure effective peer review of the model's development
   g. Considering the severity and likelihood of the risks identified, establish an ethics committee, or, depending on the size and resources of the organization, and ethics referent
3. In response to the risks associated with the use of the AI system
   a. Implement measures to prevent the storage, regurgitation or generation of personal data, especially in the context of generative AI systems
   b. For general purpose AI systems, mitigate the risk of unlawful reuse by implementing technical measures
   c. Implement measures to address significant serious ethical risks

The Draft CNIL LI Web Scraping Focus Sheet recommends the implementation of several or even all of the following measures:

- Exclude, by default, the collection of data from certain websites
- Exclude collection from websites which clearly object web scraping and the reuse of their content for the purpose of building for AI training databases
- Create a "push-back list" managed by the controller allowing data subjects to object to the collection of their data on certain websites or online platforms by providing information that identifies them on those websites
- Provide data subjects with the option to object to the processing at their direction
- Limit the collection to freely accessible data (i.e. content accessible to any unregistered user without account creation) that were manifestly made public by the data subjects
- Apply anonymization or pseudonymization process immediately after data collection
- Disseminate as widely as possible information about data collection and data subjects' rights
- Prevent any combination of data based on individual identifiers
- Register contact details in a list that could be set up by the CNIL of organisations processing data collected through web scraping for AI development

The Draft CNIL LI Open Source Models Focus Sheet recommends:

- Ensuring that the published allow for a sufficient level of transparency, effective peer review and a real

contribution to the open-source community or scientific research, at minimum, by opening the following elements:

- The model parameter
- The code required to use the model
- A model description sheet, including information on its architecture, performance and limitations
- A descriptive sheet of the data used for training, fine-tuning or improving the model
- The publication of the training dataset
- Distributing the model under a license allowing contribution to the open-source community and authorising its download, modification, and reuse
- Implement legal measures (e.g. restrictive licenses) to limit model reuse and technical measures (e.g. digital watermarking) to trace and control certain reuses
- Implement technical data security measures (e.g. anonymization or pseudonymization of data or carrying out analyses to measure the risks of regurgitation or data leakage)
- Implement measures to ensure protection of data subjects' information and the exercise of their rights

Appendix III – Balancing Interests or Fundamental Rights and Freedoms When Web Scraping is Used to Collect User Generated Content and for Model Training

I. CONTROLLER'S INTERESTS

It is a priori legitimate for a Controller to develop new systems and functionalities for users of a service.

II. INDIVIDUALS' BENEFITS

Benefit Level = (Benefit Significance) x (Benefit Likelihood))

Benefit Significance (from the perspective of the individual): 1 = low impact; 2 = moderately low impact; 3 = moderate impact; 4 = moderately high impact; 5 = high impact

1. Improved healthcare
   - O 5
   - O 4
   - O 3
   - O 2
   - O 1

2. Better accessibility of certain essential services
   - O 5
   - O 4
   - O 3
   - O 2
   - O 1

3. Facilitation of the exercise of fundamental rights such as access to information, freedom of expression, access to education
   - O 5
   - O 4
   - O 3
   - O 2
   - O 1

4. Scores from all categories of benefits are totaled and the likelihood of the benefits is calculated as follows:

Low impact:	1 – 3
Moderately low impact:	4 - 6
Moderate impact:	7 – 9
Moderately high impact:	10 - 12
High impact	13 - 15

Benefit Likelihood (from the perspective of the individual): 1 = slightly; 2 = not likely; 3 = likely; 4 = highly likely; 5 = expected

5. Improved healthcare
   - O 5
   - O 4
   - O 3
   - O 2
   - O 1

6. Better accessibility of certain essential services
   - O 5
   - O 4
   - O 3
   - O 2
   - O 1

7. Facilitation of the exercise of fundamental rights such as access to information, freedom of expression, access to education
   - O 5
   - O 4
   - O 3
   - O 2
   - O 1

8. Scores from all categories of impacts are totaled and the benefit impact is calculated as follows:

   Slightly : 1 – 3

Not likely: 4 - 6
Likely: 7 - 9
Highly likely: 10 - 12
Expected: 13 - 15

9. The Benefit Level (Score in Question 4) x (Score in Question 8) is:

None: 0
Low: 1 - 4
Medium: 5 - 8
Moderate: 9 - 12
High: 13 - 15

III. INDIVIDUAL'S RISKS

Inherent Risk Level = (Risk Likelihood) x (Risk Significance):
Risk Likelihood: 1 = slightly; 2 = not likely; 3 = likely; 4 = highly likely; 5 = expected
Risk Significance: 1 = low impact; 2 = moderately low impact; 3 = moderate impact; 4 = moderately high impact; 5 = high impact
Residual Risk Level = (Inherent Risk Level) x (Effectiveness in Reducing Risk)
Effectiveness in Reducing Risk: 1 = low effectiveness; 2 = moderately low effectiveness;
3 = moderately effective; moderately highly effective; 5 high effectiveness

<u>Data Collection</u>

1. Risks
   A. Invasion of privacy
   - O 1
   - O 2
   - O 3
   - O 4
   - O 5

   B. Impact on freedom of expression
   - O 1
   - O 2
   - O 3
   - O 4

- O 5

2. What is the highest likelihood value of the risks associated with Data Collection?
   - O 1
   - O 2
   - O 3
   - O 4
   - O 5

3. What is the significance of the risks that Data Collection could have on individuals?
   - O 1
   - O 2
   - O 3
   - O 4
   - O 5

4. The Inherent Risk Level for Data Collection (the level in Question 2) x (the level in Question 3) is:

O	None:	1-3
O	Low:	4 - 6
O	Medium:	7 - 9
O	Moderate	10 - 12
O	High:	13 - 15

5. If the Inherent Risk Level for Data Collection is High, Moderate, or Medium, how effective are the following measures in reducing Data Collection inherent risk?

   A. Excluding unnecessary data categories
   - O 1
   - O 2
   - O 3
   - O 4
   - O 5

   B. Deleting irrelevant collected data

- O 1
- O 2
- O 3
- O 4
- O 5

6. Scores from all security measures in Question 5 are totaled, and the effectiveness of the security measures in reducing inherent risk is:

- O Low effectiveness: 1 - 2
- O Moderately low effectiveness: 3 - 4
- O Moderately effective 5 – 6
- O Moderate highly effective: 7 – 8
- O Highly effective 9 - 10

7. The Residual Risk Level for Data Collection (Score in Question 4) x (Score in Question 6) is:

- O None: 1 - 2
- O Low: 3 - 4
- O Medium: 5 – 6
- O Moderate: 7 - 8
- O High: 9 – 10

<u>Model Training and Data Retention</u>

8. Risks

A. Loss of confidentiality of training data

- O 1
- O 2
- O 3
- O 4
- O 5

B. Invasion of privacy and loss of confidentiality related to data memorization/regurgitation in the model

- O 1

- O 2
- O 3
- O 4
- O 5

C. Lack of transparency and opacity of processing

- O 1
- O 2
- O 3
- O 4
- O 5

D. Difficulty in guaranteeing the exercise of rights

- O 1
- O 2
- O 3
- O 4
- O 5

9. What is the highest likelihood value of the risks associated with Model Training and Data Retention?
   - O 1
   - O 2
   - O 3
   - O 4
   - O 5

10. What is the significance of the risks that Model Training and Data Retention could have on individuals?
    - O 1
    - O 2
    - O 3
    - O 4
    - O 5

11. The Inherent Risk Level for Model Training and Data Retention (the level in Question 9) x (the level in Question 10) is:
    - O None: 1-3

- O Low: 4 - 6
- O Medium: 7 - 9
- O Moderate 10 - 12
- O High: 13 - 15

12. If the Inherent Risk Level for Model Training and Data Retention is High, Moderate, or Medium, how effective are the following measures in reducing Data Collection inherent risk?
    A. Loss of Confidentiality of Training Data
        a. Anonymization/pseudonymization of data during collection
        - O 1
        - O 2
        - O 3
        - O 4
        - O 5

        b. Use of synthetic data
        - O 1
        - O 2
        - O 3
        - O 4
        - O 5

    B. Invasion of privacy and loss of confidentiality related to data memorization/regurgitation in the model
        a. Observing reasonable period between dissemination of training dataset and its use
        - O 1
        - O 2
        - O 3
        - O 4
        - O 5

        b. Providing periodic re-training of the model
        - O 1
        - O 2
        - O 3
        - O 4
        - O 5

C. Lack of transparency and opacity of processing
   a. Transparent development of the AI system and its auditability
      - O 1
      - O 2
      - O 3
      - O 4
      - O 5
   b. Peer review of model development
      - O 1
      - O 2
      - O 3
      - O 4
      - O 5

D. Difficulty in guaranteeing the exercise of rights
   a. Facilitation of exercise of rights if personal data is used to train model
      - O 1
      - O 2
      - O 3
      - O 4
      - O 5
   b. Discretionary right to object
      - O 1
      - O 2
      - O 3
      - O 4
      - O 5
   c. Reasonable time between creation of training dataset and its use
      - O 1
      - O 2
      - O 3
      - O 4
      - O 5
   d. Transmission of the exercise of rights

- O 1
- O 2
- O 3
- O 4
- O 5

13. Scores from all security measures in Question 12 are totaled, and the effectiveness of the security measures in reducing inherent risk is:

- O Low effectiveness: 1 – 9
- O Moderately low effectiveness: 10 - 19
- O Moderately effective: 19 - 29
- O Moderately highly effective: 30 - 39
- O Highly effective; 40 - 50

14. The Residual Risk Level for Model Training and Data Retention (Score in Question 11) x (Score in Question 13) is:

- O None: 1 - 150
- O Low: 150 - 300
- O Medium: 301 - 450
- O Moderate: 451 - 600
- O High: 601 - 750

<u>Use of the AI System</u>

15. Risks

A. Invasion of privacy and loss of confidentiality related to data memorization/regurgitation in the model

- O 1
- O 2
- O 3
- O 4
- O 5

B. Damage to reputation

- O 1

- O 2
- O 3
- O 4
- O 5

C. Regurgitation of protected data
- O 1
- O 2
- O 3
- O 4
- O 5

E. Discriminatory biases
- O 1
- O 2
- O 3
- O 4
- O 5

F. Unlawful reuse
- O 1
- O 2
- O 3
- O 4
- O 5

16. What is the highest likelihood value of the risks associated with Use of the AI System?
- O 1
- O 2
- O 3
- O 4
- O 5

17. What is the significance of the risks that Use of the AI System could have on individuals?
- O 1
- O 2

- O 3
- O 4
- O 5

18. The Inherent Risk Level for Use of AI System (the level in Question 16) x (the level in Question 17) is:

O	None:	1 - 5
O	Low:	6 - 10
O	Medium:	11 – 15
O	Moderate:	16 – 20
O	High:	21 - 25

19. If the Inherent Risk Level for Use of AI System is High, Moderate, or Medium, how effective are the following measures in reducing:

A. Invasion of privacy and loss of confidentiality related to data memorization/regurgitation in the model
   a. Observing reasonable period between dissemination of training dataset and its use
   - O 1
   - O 2
   - O 3
   - O 4
   - O 5

   b. Providing periodic re-training of the model
   - O 1
   - O 2
   - O 3
   - O 4
   - O 5

B. Damage to reputation
   a.
   - O 1
   - O 2
   - O 3
   - O 4

- O 5

b. Use of synthetic data

- O 1
- O 2
- O 3
- O 4
- O 5

C. Regurgitation of protected data

a. Observing reasonable period between dissemination of training dataset and its use

- O 1
- O 2
- O 3
- O 4
- O 5

b. Providing periodic re-training of the model

- O 1
- O 2
- O 3
- O 4
- O 5

D. Discriminatory biases

a. Ensuring the quality of the dataset

- O 1
- O 2
- O 3
- O 4
- O 5

b. Annotation as soon as the dataset is created

- O 1
- O 2
- O 3
- O 4
- O 5

c. Application of filters in the deployment phase

- O 1
- O 2
- O 3
- O 4
- O 5

E. Unlawful reuse

a. Reuse licenses

- O 1
- O 2
- O 3
- O 4
- O 5

Digital watermarking

- O 1
- O 2
- O 3
- O 4
- O 5

20. Scores from all security measures in Question 19 are totaled, and the effectiveness of the security measures in reducing inherent risk is:

- O Low effectiveness: 1 - 10
- O Moderately low effectiveness: 11 - 21
- O Moderately effective: 22 - 33
- O Moderately highly effective: 34 - 44
- O Highly effective: 45 - 55

21. The Residual Risk Level for Use of AI System (Score in Question 18) x (Score in Question 20) is:

- O None: 1 - 274
- O Low: 275 - 549

- Medium: 550 - 824
- Moderate: 825 - 1099
- High: 1100 - 1375

Overall Residual Risk Level

Scores from the effectiveness of all categories of security measures are totaled (Questions 7, 14, and 21), and the effectiveness of the security measures is calculated as follows:

None:	0
Low:	1 – 2,578,125
Medium:	2,578,126 – 5,156,250
Moderate:	5,156,251 – 7,734,375
High:	7,734,376 – 10,312,500

IV. SUMMARY

Controller's Interests are legitimate

Individuals' Benefit Level is _____

Individuals' Overall Residual Risk Level is _____

Appendix IV – Project Participants

Note - While their views and input played a valuable role in shaping the project, their involvement does not indicate endorsement of the contents of the report

Irish Data Protection Commission
Data Protection Network
Pembroke Privacy
Zoom Info
Mastercard
Merck
Belgium Data Protection Authority
Commission Nationale de l'Informatique et des Libertés (CNIL)
HP
Google
UK Information Commissioner's Office
Apple
Tik Tok
Bristol Myers Squib
Workday
Pembroke Privacy
Dun & Bradstreet
Crosley Law
Future of Privacy Forum
Interpublic Group
Johnson & Johnson
Cisco
Cognizant
Apple
Data and Marketing Association
Meta
Connected by Data