IAF

Information Accountability Foundation

29 February 2024
info@informationaccountability.org

The following comments are pursuant to the consultation of the Information Commissioner's Office (ICO) on the lawful basis for scraping data from the web or processing web-scraped data to train generative AI models (Consultation). The Information Accountability Foundation (IAF) welcomes the opportunity to provide its input.

1. Who We Are[1]

The IAF is the preeminent global information policy think tank, creating collaborative scholarship and education on the policies and processes necessary to use data responsibly in an observational age, while enabling a trusted digital ecosystem that serves people. It is not-for-profit and independent. The IAF is the incorporation of the Global Accountability Dialog, the multi-stakeholder project that developed the "Essential Elements of Accountability." The IAF believes:

- It is critical that organizations are able to think with data and engage in knowledge creation to enable and achieve the benefits of a global digital ecosystem.
- To be trusted, organizations must be accountable, responsible, and answerable, and be prepared to demonstrate their accountability.
- Frameworks based on risk assessment and effective data governance enable beneficial, data-driven innovation while protecting individuals and society from the. potential harms that may arise from data processing in the digital age.

Since 2014, the IAF has focused on public policy models and the governance of innovative uses of data pertaining to people in advanced analytics and artificial intelligence (AI). A trademark of IAF's work is its consultative approach, described generally as follows:

- Based on its experience, the IAF drafts a proposed document (e.g., a framework or an assessment) and gets the reactions of a group of representative private sector organizations.
- Based on the feedback of the organizations, the IAF revises the document several times.
- When a close to final version of the document has been achieved, the IAF conducts a multi-stakeholder session where the feedback of academics, civil society, and regulators to the near-final version of the document is obtained.
- After the multi-stakeholder session, the IAF revises the document and distributes the final version of the document through appropriate media channels.

[1] These comments were prepared by IAF staff and do not necessarily reflect the views of the IAF Board of Directors, funders, or members of the IAF extended community.

Thus, the IAF's work is developed on the basis of workshops and research with a wide range of private sector organizations and other stakeholders, drawing on practical experience and real-life challenges, testing its approaches, and finding forward looking and long-term solutions.

2. The Scope of Our Comments

The thoughtful approach taken in the Consultation is reflective of the UK's pro-innovation strategy regarding AI regulation and the ICO's Guidance on AI and Data Protection (AI Guidance). Since the ICO previously provided the AI Guidance, the IAF understands why the Consultation is limited to generative AI development and use. The IAF hopes that the final generative AI guidance will be incorporated into the AI Guidance and will not be standalone guidance. The IAF's Comments are provided on this assumption, and given the focus of the Consultation, the IAF's Comments will focus on the subjects raised in Sections 1, 3 and 4.

3. IAF's Views on the Proposed Regulatory Approach

A. The Model Development Lifecycle

The five-stage model development description[2] is a simplified view of the lifecycle which is very helpful to those who are unfamiliar with not only generative AI but AI model development in general. Important to the understanding of the impact of these actions is the distinction between "thinking" and "acting" with data." The IAF has written on this topic extensively, e.g., Big Data and Analytics 2013 and Advance Data Analytics Processing. Thinking with data is the robust use of data to create new insights, and the use of those insights to affect individuals is acting with data. It is important to understand this distinction because the risk associated with the two phases is different.

Although thinking with data (everything through creation of the first model in Figure 1) may use personal data, since that phase normally does not have any impact upon individuals, any additional risk beyond the organization's other uses of personal data is nominal. Until the hypothesis or the model is applied to individuals, there is little potential for additional harm that does not otherwise rise from the use of personal data and therefore little impact on individuals. Failure to understand the risk at each stage means a failure to understand the appropriate mitigations necessary. When the difference between thinking and acting with data is understood, it is less likely that mitigations will be overbroad and instead will be commensurate and relevant to the context of each stage. For example, the mitigations customarily used with personal data should be used when personal data are used for training purposes.

The important relevancy to this distinction is knowledge creation and knowledge application (aka thinking and acting with data) should not be treated the same way.

[2] This description is similar to the Discovery and Application flowchart on page 8 of Big Data and Analytics 2013.

Therefore, regulatory approaches, including the regulation, oversight, and enforcement, should treat these two processes differently. Both processes require controls, but since the risks are not the same in knowledge creation and knowledge application, the controls should not be the same. This construct was highlighted in the IAFs feedback to the Department for Digital, Culture, Media, and Sport ("DCMS") consultation entitled "Data: A new direction."

The model development lifecycle in the Consultation is too generalized. Once the model is created, then it is tested on actual or dummy data to make sure that it works as intended. Only once the model has been tested adequately is it deployed. It is this final development and deployment concept that is reflected in the EU AI Act.

Figure 1 also ignores the need to assess at each stage of the lifecycle. Conducting just one assessment at only one point in time is wholly inadequate. The IAF discusses how and when to conduct these assessments in Evolving AIA Impact Assessments. The lifecycle depicted in Figure 1 ends with deployment. As discussed later in the Consultation, further mitigations in deployment (e.g., output filters and organizational and contractual controls) play a substantive role in reducing risk. Moreover, mitigations should be proportionate and relevant to the context of each stage.

B. Training Data

The ICO limits the discussion of training data in the Consultation to publicly accessible data usually obtained through web scraping. This emphasis is understandable given the allegations in paragraph 85 of the complaint in The New York Times Company v. Microsoft Corporation that ChatGPT-2 "contains a staggering amount of scraped content from The Times." However, many organizations also use their own data, either as a developer or a deployer, as well as data from other sources to supplement and further train their models. Releasing guidance that focuses only on scraped data may not be helpful to organizations.

The statement "developers need to ensure the collection of the personal data they process to train models complies with data protection" assumes that all screen scraped data should be treated the same from a data protection perspective. This cannot possibly be the case. Significantly, it depends on what kind of data is collected and used. If special category data is processed, consent cannot have been the lawful basis for collecting and using the training data. In *GC, AF, BH, ED v CNIL, Case C-136/17,* the CJEU concluded that a balancing of the interests of the individuals concerned against the other interests involved should occur. The holding in this case suggests that additional balancing should occur when the organization is made aware by the data subject that special category data are contained within the web scraped data.

C. Legitimate Interest Three-Part Test

The ICO has developed the Sample LIA Template to help implement the Legitimate Interest "three-part" test: purpose, necessity and balancing. Consistent with the comments on the AI Guidance, the IAF hopes that the final generative AI guidance will be incorporated into the Sample LIA Template, which acts as a normative framework, and will not be standalone guidance.

Most of the analysis is generic in nature and no more specific than that already provided in the AI Guidance or than the questions asked in the Sample LIA Template. Examples of more specific questions/topics can be found in Evolving AIA Impact Assessments.

1. Purpose Test

   The Consultation talks about having a legitimate interest in collecting the data in one paragraph and in the next paragraph talks about the developer's interest in developing and deploying the model. These are two different forms of processing for which legal bases to process are necessary, and the analysis does not separate them out. There must be a legal basis for each processing operation (i.e., collection, recording, storage, use). Organizations are expected to be granular in their drafting, and it is respectfully suggested that the ICO should model that drafting in its guidance.

2 Necessity Test

   One of the questions to consider in determining necessity is whether the processing is proportionate to the purpose. Recently, the IAF has been considering what the proportionality principle means in the context of the private sector, e.g., A Principled Approach to Rights and Interests Balancing. In doing that work, the IAF distinguished the public sector proportionality analysis which considers two factors from the private sector which considers many factors. This multi-factor analysis is not limited solely to the factors of organizational processing and the individual to whom the data pertains. It requires organizations also to look at the rights of groups of people and society. This approach is consistent with ICO Guidance on Legitimate Interests which requires taking into account "the interests or fundamental rights and freedoms of the data subject . . ." and cautions organizations that if "the processing has a wider public interest for society at large, then this may add weight to your interests when balancing them against those of the individual." Thus, proportionality in the private sector needs to bring in all the factors that are involved when data are used for complex purposes, such as AI.

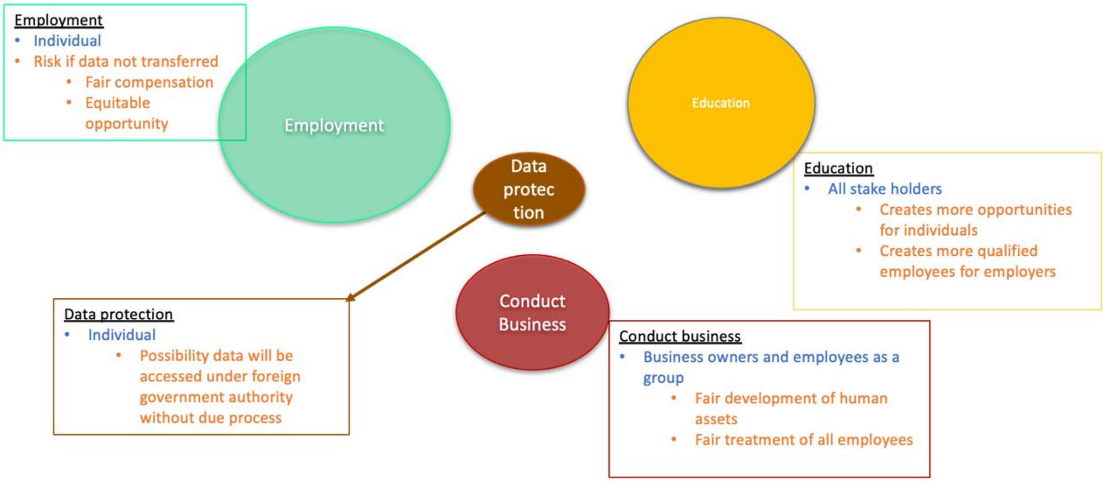
An example of this private sector proportionality based on the transfer of human resource data is set forth in A Principled Approach to Rights and Interests Balancing. Many U.S. organizations have employees based in the UK, and before the UK Extension to the EU-US Data Privacy Framework, they took a risk-based approach in transferring human research data. The stakeholders in this proportionality analysis are the employer, the individual employee, other employees, regulators, and national security agencies. Some of the employee interests in the transfer of. human resource data are:

- Recruitment and interview processes
- Hiring
- Onboarding
- Access management
- Employee directory
- Development and performance management
- Compensation
- Talent and succession planning
- Training
- Health and other benefits
- Emergency contacts
- Volunteer opportunities.

Using the UN Universal Declaration of Human Rights, the rights that relate to the transfer of human resource data are:

- 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.

The below diagram shows these rights weighted,


Employment
Individual
Risk if data not transferred
Fair compensation
Equitable opportunity
Employment
Education
Data protec tion
Education
All stake holders
Creates more opportunities for individuals
Creates more qualified employees for employers
Conduct Business
Data protection
Individual
Possibility data will be accessed under foreign government authority without due process
Conduct business
Business owners and employees as a group
Fair development of human assets
Fair treatment of all employees

Showing the proportionality of one of the factors visually helps the organization understand the impact of the processing.

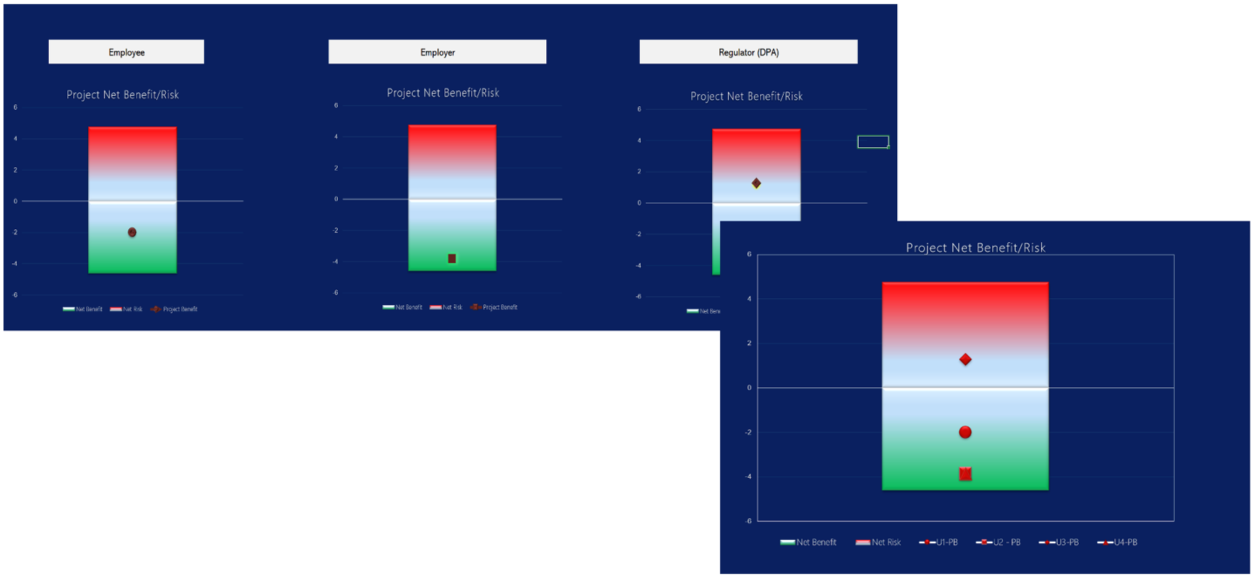
3 Balancing Test

In A Principled Approach to Rights and Interests Balancing, the IAF developed a methodology to, in an orderly and repeatable fashion, identify and demonstrate:

1. The individual rights (in a fundamental rights-based system) or established individual interests (in legal systems where fundamental rights are not established).
2. The stakeholders whose rights or interests are involved.
3. The adverse processing impacts that may be involved and their likelihood and level of consequence.

This methodology demonstrates the various stakeholders involved, the rights and interests of each of these various stakeholders, and how the likelihood and severity of risks impact the rights and interests of each of these various stakeholders. The IAF's impact analysis takes into account all of the stakeholders and weighs the benefits/interests against the risks/harms using a 1, 3, 5 scale and shows the results in two different ways: math determined and visual. IAF's weighing factors in as many stakeholders, benefits and risks as are relevant to the processing being assessed. It is capable of weighing each of the stakeholders vis-à-vis each of the other two factors. It can demonstrate the results mathematically or visually or both.

The below diagram shows the benefits and risks to the rights and interests of the employee, the employer, and the regulator – weighted - in the transfer of human resource data.



Visually and mathematically, the diagram shows that the net benefits/risks of the employee and the employer outweigh those of the regulator in transferring the human resource data.

The net benefit/risk reflected in this diagram have considered the mitigating actions. The Consultation suggests numerous times that contractual controls may mitigate risks. As Cambridge Analytica showed, contractual restrictions can be limited. The sole recourse is breach of contract. No matter how rigorous an organization's due diligence program is, if a vendor is intent on breaching a contract and has determined the upside of the breach outweighs the downside, then contractual controls are of limited use. Furthermore, as the Consultation makes clear, contractual controls only work when the developer has a contractual relationship with the deployer or through an API. While developers could require third parties to have the same controls in their contracts, those third-party contracts are not part of the developer's due diligence program and therefore it is impossible for developers to monitor them. Even when the developer has a direct relationship with the deployer, if the developer has thousands of deployers (or more), then it is unrealistic to expect individualized audits of all deployments. A more realistic approach is to expect each organization to act responsibly with respect to its use of the model, thereby acting in accordance with the accountability principle.

Recent experience with Legitimate Interest Assessments has been unsuccessful because the assessments have failed to address the full range of rights and interests/benefits versus harms/risks. The examples in A Principled Approach to Rights and Interests Balancing and the Colorado Data Protection Assessment show how to do the balancing called for by the Consultation.

D. Demonstrability

Completion of the Legitimate Interest Assessment is one form of demonstrability, but organizations need to be able to demonstrate that the structure is in place within the organization, the resources have been allocated and are in place within the organization, and the policies and procedures are in place within the organization so that the organization can demonstrate that the Legitimate Interest Assessment was conducted with integrity. Participation in an ICO AI Audit might be an additional way of demonstrating integrity. It is only when the regulatory focus shifts from an assessment of the outcome to an assessment of an organization's process and capabilities will the trust deficit regarding organizational collection and use of personal data be reduced. The reduction in this trust deficit is necessary for AI, including generative AI, to be a trusted form of processing.

Furthermore, as the Consultation points out, invisible processing and AI related processing are both seen as high-risk activities that require a DPIA under ICO guidance. if an organization is engaging in processing that should be the subject of a DPIA, the conduct of a legitimate interest assessment should be the first step in determining whether a DPIA is necessary. Generative AI is the type of new technology that is likely to require the carrying out of a DPIA prior to the processing. The connection between the legitimate interest assessment and the DPIA should be called out explicitly in the legitimate interest assessment.

4. Final Comments

It is the IAF's understanding that the Consultation is the first in a series of chapters which will outline the ICO's emerging thinking on how the ICO interprets specific requirements of the UK GDPR and Part 2 of the DPA (e.g., purpose limitation principle, accuracy principle, data subject rights) The IAF commends the ICO's undertaking this thoughtful effort and urges the ICO to act expeditiously and deliberatively in the issuance of each chapter so that the final versions will be timely.

Thank you for the opportunity to submit comments on the Consultation.