Identifying Algorithmic Harms when Creating DPIAs: A Quick Guide

This guide provides companies a landscape view of potential algorithmic harms[1] that should be considered when creating DPIAs, as an obligation under the GDPR.

GDPR

The European General Data Protection Regulation went into effect in May 2018, as the most comprehensive data protection law to date. Among the novelties brought by the GDPR to the existing data protection law framework in the European Union was the focus on accountability. This was reflected in several new legal obligations, such as the creation of a register of processing activities, implementing data protection by design and by default and conducting Data Protection Impact Assessments (DPIAs). The GDPR requires data controllers to carry out a **DPIA** when data processing is "likely to result in a **high risk** to the rights and freedoms of natural persons" (Article 35(1)).

What is a DPIA?

A DPIA is a structured assessment of a processing activity to help data controllers manage (identify and minimize) the risks that a project poses to the rights and freedoms of individuals. Guidance from the European Data Protection Board provides criteria and examples on what type of processing activities may result in a high risk to the rights and freedoms of persons[2] that would require controllers to conduct a DPIA.

What are the required elements of a DPIA?

Following Article 35, a DPIA must contain (1) a systematic description of the processing that it covers (2) an assessment of the necessity and proportionality of the processing, (3) an assessment of the risks to the rights and freedoms of data subjects, and (4) proposed measures to mitigate the identified risks. This Quick Guide focuses on point (3) and aims at supporting controllers to identify what kind of harms to the rights and freedoms of persons they should take into account when assessing the impact that a processing activity has and the type of risks it may create.

What is a "risk to rights and freedoms"?

There is no legal definition of what a risk to rights and freedoms of a person is in the context of the GDPR. The EDPB takes the view that a "risk" is "a scenario describing an event and its consequences, estimated in terms of severity and likelihood".[3] In order to assess the risks, they first need to be identified. The most effective way to identify risks to rights and freedoms seems to be focusing on the "consequences" part of the definition proposed by the EDPB. Therefore, in order to identify risks data controllers should prefigure and describe the consequences the processing activity they propose might have to the rights and freedoms of persons. In other words, they should think of any potential damage or harm to the rights and freedoms of individuals.

What type of damage/harm is relevant for including in a DPIA?

The GDPR explains in its Preamble that risks to the rights and freedoms of data subjects "may result from personal data processing which could lead to **physical, material** or **non-material** damage"[4]. Therefore, any type of damage or harm is relevant when conducting a DPIA. The GDPR also provides for some concrete examples of harm to the rights and freedoms of persons that may result from processing their personal data, but it mixes them with examples of risky processing activities, making it difficult for controllers to distinguish between potential damage/harm as a consequence of the processing on one hand, and the nature of the

[1] https://fpf.org/wp-content/uploads/2017/12/FPF-Automated-Decision-Making-Harms-and-Mitigation-Charts.pdf
[2] Article 29 Working Party (EDPB), Guidelines on Data Protection Impact Assessment and determining whether processing is "likely to result in a high risk" for the purposes of Regulation 2016/679, WP 248 rev.1.
[3] WP29 (EDPB) Guidelines on DPIAs, p. 6.
[4] GDPR, Recital 75.

FUTURE OF PRIVACY FORUM

processing activity on the other hand. In particular, Recital 75 of the GDPR specifies that risks to rights and freedoms can result from processing which could lead to damage, in particular:

- Where the processing may give rise to:
  - discrimination,
  - identity theft or fraud,
  - financial loss,
  - damage to the reputation,
  - loss of confidentiality of personal data protected by professional secrecy,
  - unauthorized reversal of pseudonymization,
  - Any other significant economic or social disadvantage.
- Where data subjects might be deprived of their rights and freedoms or prevented from exercising control over their personal data;
- Where personal data are processed which reveal racial or ethnic origin, political opinions, religion or philosophical beliefs, trade union membership ("special categories of data");
- The processing of genetic data, data concerning health or data concerning sex life or criminal convictions and offences or related security measures;
- Where personal aspects are evaluated, in particular analyzing or predicting aspects concerning performance at work, economic situation, health, personal preferences or interests, reliability or behavior, location or movements, in order to create or use personal profiles;
- Where personal data of vulnerable natural persons, in particular of children, are processed;
- Where processing involves a large amount of personal data and affects a large number of data.

What "rights and freedoms" are relevant for a DPIA?

The EDPB explained in the DPIA guidance that "the reference to the rights and freedoms of data subjects primarily concerns the rights to data protection and privacy, but may also involve other fundamental rights such as freedom of speech, freedom of thought, freedom of movement, prohibition of discrimination, right to liberty, conscience and religion"[5]. All rights and interests should thus be taken into account when conducting a DPIA, especially if they are protected as fundamental rights in the EU legal framework. For example, processing of personal data may not necessarily lead to the loss of privacy (as it would generally be the case of publicly available data being processed), but it may lead to loss of opportunity or to discrimination.

The catalog of rights

The catalog of rights provided by the EU Charter of Fundamental Rights is particularly relevant when conducting such an analysis. Recital 4 of the GDPR specifies that the Regulation "respects all fundamental rights and observes the freedoms and principles recognized in the Charter, (…) in particular the respect for private and family life, home and communications, the protection of personal data, freedom of thought, conscience and religion, freedom of expression and information, freedom to conduct a business, the right to an effective remedy and to a fair trial, and cultural, religious and linguistic diversity."

Societal harms

Societal harms are not specifically mentioned in the provisions of the GDPR regulating DPIAs, but they are taken into account by some national data protection authorities providing advice on how to conduct DPIAs. For example, the UK Information Commissioner Office (ICO) considers that the focus should be "on the potential for harm to individuals or to society at large, whether it is physical, material or non-material."[6]

[5] WP29, DPIA Guidelines, p. 6.
[6] ICO DPIA Guidelines, available at https://ico.org.uk/for-organisations/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/data-protection-impact-assessments/

The "Chart of Potential Harms from Automated Decision-Making" as a DPIA tool

Analysis of personal data can be used to improve services, advance research, and combat discrimination. However, such analysis can also create valid concerns about differential treatment of individuals or harmful impacts on vulnerable communities. In 2017, FPF attempted to identify, articulate, and categorize the types of harm that may result from automated decision-making. To inform this effort, FPF reviewed leading books, articles, and advocacy pieces on the topic of algorithmic discrimination. We used this review to create a chart categorizing specific harms that can arise from automated decision-making.

We distilled the harms identified in the literature into four broad categories: (1) loss of opportunity, (2) economic loss, (3) societal detriment, and (4) loss of liberty—to depict the various spheres of life where automated decision-making can cause injury. It also notes whether each harm manifests for individuals or collectives, and as illegal or simply unfair.

"Loss of opportunity" harms occur within the domains of the workplace, housing, social support systems, healthcare, and education. "Economic loss" harms primarily cause financial injury or discrimination in the marketplace for goods and services, including credit discrimination, price discrimination, and narrowing of choice. "Societal detriment" harms impact one's sense of self, self worth, or community standing relative to others, and include filter and network bubbles, dignitary harms, stereotype reinforcement, and bias. Lastly, "Loss of liberty" harms constrain one's physical freedom and autonomy, including through the constraint of suspicion, surveillance, and incarceration.

The harms in this chart relate closely to the descriptions of "physical, material or non-material damage" detailed in Recital 75 of the GDPR, and may be helpful in identifying and assessing the harms caused by "high risk" processing.

The chart below provides some specific examples from FPF's chart that are relevant for a DPIA conducted following the requirements of the GDPR when analyzing automated decision-making.

Examples of Algorithmic Harms

Harm Category	Examples of Processing	Examples of Harm	Applicable "High Risk" Criteria from the WP29/EDPB guidance	Potentially Relevant EU Charter rights
Employment Discrimination ***Loss of Opportunity***	A hiring agency using an algorithm to select best candidates for an employer.	Filtering candidates by work proximity may lead to the inadvertent exclusion of minorities	• Evaluation & Scoring • Automated decision-making • Matching or combining data sets • Data processed on a large scale	• Article 15 – freedom to choose an occupation and right to engage in work; • Article 21 – Non-discrimination • Article 23 – Equality between women and men • Article 25 – The Rights of the elderly • Article 26 – Integration of persons with disabilities

Insurance & Social Benefit Discrimination ***Loss of Opportunity***	A biotechnology company offering genetic tests directly to consumers in order to assess and predict disease/health risks	Genetic information may be used to discriminate against individuals applying for life insurance	• Evaluation & Scoring • Sensitive data or data of a highly personal nature. • Automated decision-making	• Article 7: The right to respect for private life • Article 21: Non-discrimination • Article 34: Social security and social assistance • Article 35: Health care
Housing Discrimination ***Loss of Opportunity***	A search engine using artificial intelligence, machine learning and deep learning to aggregate publicly available data	Landlord decides to exclude minorities based on search results suggesting criminal history by race	• Evaluation & Scoring • Automated decision-making • Matching or combining data sets • Data processed on a large scale • New technologies	• Article 7: The right to respect for private life • Article 21: Non-discrimination
Education Discrimination ***Loss of Opportunity***	An advertiser selecting an ad audience	Presenting only ads on for-profit colleges to low-income individuals, potentially causing them to lose money and opportunity	• Evaluation & Scoring • Automated decision-making • Matching or combining data sets • Data processed on a large scale	• Article 14: Right to education • Article 21 – Non-discrimination
Credit Discrimination ***Economic Loss***	An institution creating a national level credit rating or fraud database	Denying credit to all residents in specified neighborhoods (“redlining”)	• Evaluation or scoring • Automated decision-making • Prevents data subject from exercising a right or using a service or a contract • Sensitive data or data of a highly personal nature	• Article 21: Non-discrimination • Article 38: Consumer protection
Filter Bubbles ***Societal Detriment***	The gathering of public social media data for generating profiles	Algorithms that promote only familiar news and information, magnifying social fissures	• Evaluation & Scoring • Data processed on a large scale • Matching or combining of datasets • Sensitive data or data of a highly personal nature	• Article 11: Freedom of expression and information