Overview of Data Protection Harms and the ICO's Taxonomy

Information Commissioner's Office

Date: April 2022

Contents

1. Background 1

1.1. Approach 1

2. What are data protection harms? 2

2.1. Theory of Data Protection Harm 2

2.2. Likelihood and Severity 3

3. Why are data protection harms hard to identify and quantify? 4

4. Literature Review 6

5. A taxonomy of data protection harms 8

The ICO’s Data Protection Harms Taxonomy 9

1. Background

Discussion of data protection issues frequently involves the use of the term 'harm'. The ICO's Harm Project seeks to improve the understanding of harm in a data protection context.

In this document, we set out our framework for harms, our evidence base, and a taxonomy of data protection harms.

1.1. Approach

Our approach to understanding the concept of harm has been to draw legal, policy and economic insights from wide-ranging sources. In the development of our data protection harms taxonomy, we have:

- reviewed guidance and literature on risk management to develop our understanding of how harms occur and isolate them from other elements of risk;
- collated evidence on the approaches to understanding and addressing harm from other regulators and data protection authorities;
- developed an initial harms taxonomy and socialised it internally and with key external stakeholders to test its application across a range of the ICO's work areas;
- commissioned an independent literature review on data protection harms to broaden our understanding of the wider evidence base and check and challenge our initial work;
- refined the initial taxonomy based on the findings of the literature review and lessons from testing applications of the taxonomy, ready to be shared externally and built into future work on data protection harms.

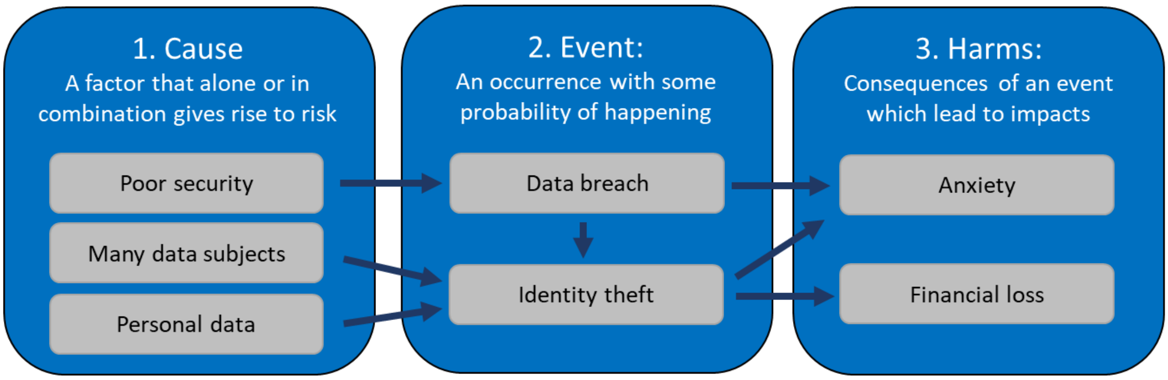
2. What are data protection harms?

The UK GDPR focuses, as set out in Recital 1, on the right to the protection of personal data. The focus on this right and the principles of UKGDPR is helpful in a legal context, however, to understand why these rights matter, it is necessary to take a step further. This is where the concept of data protection harms is useful. It allows us to focus on what the consequences of an infringement of data protection rights are for individuals and for society.

2.1. Theory of Data Protection Harm

General frameworks for considering risk and consequences can be found in organisational risk management guidance,[1] which make the distinction between causes, events and consequences. This is illustrated in Figure 1 below, which applies this thinking in a data protection context through a simplified example.

Figure 1: Theory of Data Protection Harm


1. Cause
A factor that alone or in combination gives rise to risk
Poor security
Many data subjects
Personal data
2. Event:
An occurrence with some probability of happening
Data breach
Identity theft
3. Harms:
Consequences of an event which lead to impacts
Anxiety
Financial loss

Source: ICO analysis.

As shown in the illustration, this simple framework is useful in a data protection context for stepping out the process by which harms occur. We refer to this approach as a theory of data protection harm, similar to that used in competition law to understand how breaches of legislation lead to harms to competition.[2] There are a number of things to note here:

[1] See the 'Orange Book', the Government's risk management guidance (https://www.gov.uk/government/publications/orange-book), or ISO 31000:2018 risk management guidelines (https://www.iso.org/iso-31000-risk-management.html).

[2] 4 -DFF-Factsheet-Theories-of-harm-in-competition-law-cases.pdf (digitalfreedomfund.org)

- Causes and events in isolation are not harms, it is the resulting negative consequences or impacts of the events that are the harms. Events may not always lead to a harm, for example, if a data breach is never discovered, data subjects cannot feel anxious about the event and if the data from the data breach or identity theft is not used to steal money then financial loss may never occur.

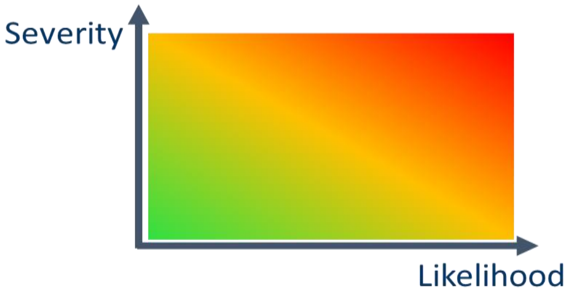
- Even in this simple example we can see that there can be complex chains of causes, events and consequences. One event might lead to other events, events can lead to multiple consequences, and multiple events can lead to the same consequences.

- It is also possible for some consequences to lead to others, for example the accumulation of loss of trust amongst individuals could result in a loss of trust at a societal level, with further harm of chilling effects on the use of services that would otherwise be beneficial to society. This is relevant to the distinction between harm and damage, discussed further below.

2.2. Likelihood and Severity

A further aspect of harm that is seen in risk management and is also inherent in GDPR is the idea of likelihood and severity as dimensions of harm. This is illustrated in Figure 2 below. An event and its consequences may have high likelihood and low severity, such as the receipt of spam emails can sometimes do, or low likelihood and high severity, such as a data breach and consequent financial harm.

Figure 2: Severity and likelihood as dimensions of harm



3. Why are data protection harms hard to identify and quantify?

In identifying data protection harms, even once a particular harm has been identified it can be challenging to quantify robustly for a number of reasons. These difficulties have parallels with the challenges of identifying the economic impact of data protection related regulatory interventions, as well as understanding the value of data (and data protection) more generally.

1. The **nature of the harms**: Our example in Figure 1 shows a number of consequences. Some of these are readily identifiable and directly quantifiable, such as financial harm of fraud leading to the loss of a certain amount of money. However, other aspects of harm resulting from the same event, such as distress or anxiety, are intangible and much more challenging to identify and quantify. Identification of harm can be challenging in itself when it is not apparent, for example when there is invisible processing.

2. Harms being **risk-based**: As noted above, data protection harms are probabilistic, which has important implications for the value that data subjects place on data protection. Individuals may incorrectly assess the likelihood of harm due to incomplete information, limitations on processing of information (so-called 'bounded cognitive ability'), and behavioural biases. This is compounded by the intertemporal nature of many privacy harms, whereby harm may be incurred long after the event that leads to it, and hence undervalued due to myopia, or 'present bias'.

3. Harms are **diffuse**: As noted above, the harm to any individual data subject may not be substantive, but when aggregated may lead to significant societal damage, or significant transfer of wealth from data subjects to controllers.[3] Or, conversely, many operators and events may inflict small insignificant harms on an individual that aggregate to a significant harm for which it is difficult to attribute blame or cause.

4. Harm **varies** by data subject: Individuals value their privacy to different extents, so the same event and consequences may have different impacts on different individuals, particularly when the harm is intangible. Harm can also vary dependent on other individual circumstances, such as whether or not someone is vulnerable, and the currency and relevance of their reputation to their livelihood or social standing.

5. Harm can be **difficult to avoid**: economic circumstances such as market power or barriers to switching can mean that harms are hard to avoid if

[3] This is a typical issue in impact assessment where benefits are spread thinly amongst many data subjects, whilst costs are concentrated amongst a small group of controllers. The latter are better resourced, motivated and capable of arguing their case.

even informed and unbiased consumers are unable to discipline providers by switching to alternatives.

4. Literature Review

The ICO Commissioned an independent review of literature related to data protection harms. The purpose of review was to build on the ICO's initial work by developing a formal evidence base. This includes both theoretical and empirical evidence.

In total, 111 references were eligible for the review. The full literature review can be found here: Review of literature relevant to data protection harms (ico.org.uk) with the key findings summarised below:

- The evidence comes from a range of disciplines – including sociology, healthcare, marketing, information technology, economics, and law. This indicates that a range of sectors are considering this issue.

- The evidence found primarily focuses on the US, UK and EU jurisdictions. In the US, articles from law journals formed a key part of the evidence base. In the UK and EU, grey literature is an important source of evidence, but there are key contributions from the fields of cybersecurity and sociology.

- A range of studies assess individuals' awareness and concerns about data protection harms. In general, individuals express concern over the use of their personal data and say privacy is important to them. In particular, individuals express high levels of concern about loss of financial data or online fraud. However, outside of financial harms, their awareness of specific harms that may result from a breach of their personal data appears to be limited.

- The risks of certain types of data protection harms occurring does not appear to have been explored in a systematic manner. This is likely to reflect the challenge in linking a specific data event (such as a data breach, or the act of sharing data with a digital service) to a subsequent harm, which may occur months or years later.

- Whilst some harms are relatively well explored in empirical studies, such as loss of personal data, unwarranted intrusion and chilling effects, other harms have limited empirical evidence. For example, no empirical studies assess damage to law and justice. There are evidence gaps for harms that are otherwise well explored in the theoretical literature such as discrimination and loss of confidentiality.

- Surveys generally ask about individuals' views, concerns, or perceptions of harms, rather than experiences. However, some surveys ask individuals whether they have experienced data protection harms. In general, the proportion who report awareness of experiencing harms is substantially lower than the proportion expressing concerns (though it should be noted that some individuals may not be aware they have experienced data protection harms).

- The evidence generally supports the idea of a mismatch between individuals’ reported concerns about use of personal data and their actions and behaviours (sometimes termed the ‘privacy paradox’). This is variously attributed to information asymmetries, the transaction costs associated with evaluating the costs and benefits of disclosing personal data, or a sense of ‘digital resignation’ and ‘digital fatalism’.

- A variety of studies discuss or explore the potential impacts and implications of a particular type of harm occurring (e.g. fraud or identity theft). No study in the review has attempted to explicitly quantify the impact of a particular data protection harm at an aggregate level. However, a number of studies have attempted to quantify the value individuals place on different types of personal data (e.g. the payment users are willing to accept to permit use of their data).

This points to a wide and varied evidence base but not particularly advanced, leaving plenty of room for the ICO to drive forward the thinking in this area. The review also found that the ICO’s initial work in this area is relatively advanced compared to other Data Protection Authorities internationally.

5. A taxonomy of data protection harms

As explained in the previous sections, data protection harms are complex, wide-ranging and frequently overlapping. A simple definition would not suffice in helping us to develop our understanding or communicate the breadth of the issue internally or externally. This mirrors the positions of other DPAs, and the vast majority of other UK regulators.[4] As such, we propose a taxonomy of data protection harms which allows the flexibility to cover the varied and evolving scope of data protection harms, whilst still enabling a shared understanding of what data protection harms are. This is supplemented by a framework for considering harms which includes:

- Figure 1 above, and the theory of data protection harm;
- Figure 2 and consideration of severity and likelihood

Whilst reviewing the literature and commentary around data protection harms it quickly became apparent that elaborating all of the complex and case-specific chains of causes, events and consequences was unmanageable in a framework intended for use across the full range of data protection issues. We therefore focus on a taxonomy of consequences, which also has the advantage of avoiding the lack of clarity between events and consequences sometimes seen in the literature.

The taxonomy presents consequences at a relatively high level, identifying harms by type and providing examples, but not attempting to make this exhaustive. Equally, no hierarchy of harms is implied by the order of the table, with the likelihood and severity of harms dependent on the circumstances of the case.

We have drawn on these in the development of the taxonomy of data protection harms below, which distinguishes in turn harm to individual, and harms to society, providing examples in each case.

[4] UK regulators we benchmarked against are: EHRC, Ofcom, FCA, CMA, Gambling Commission, ASA, GMC, CQC, SRA, Charity Commission, Environment Agency. International DPAs are the USA, Canada and Australia.

The ICO's Data Protection Harms Taxonomy

It is important to note that the taxonomy:

- is intended to provide a common language and starting point for further research
- is non-hierarchical and non-exhaustive, although we provide examples
- contains some closely related potentially overlapping categories of harm, and that some harms can lead to others.

Type	Category	Description	Examples
Individual	Financial harm	Negligently, knowingly, or purposefully paving the way for financial losses to occur	• Breach leading to fraud • Impact on credit rating • Extortion through use of personal data • Targeting those with gambling addiction problems with gambling adverts • Loss of income/employment due to reputational damage
	Bodily harm	Negligently, knowingly, or purposefully paving the way for physical injury to occur	• Suicide or other self-harm • personal data used to track someone's location, leads to assault • Medical malpractice caused by negligence or inaccuracies
	Costs of avoiding/mitigating harm	The cost in terms of time or money incurred in the avoidance or mitigation of	• Time spent avoiding harm/risk of harm • Security costs associated with protecting personal data

		harms or vulnerabilities related to data privacy	
	Discrimination	Harms arising from discrimination or bias (either conscious or unconscious)	• Entrenched bias in automated decisions • Price discrimination
	Unwarranted intrusion	Unwanted communications or intrusions that disturb tranquillity, interrupt activities, sap time or increase the risk of other harms occurring	• Unwanted targeted advertising • Nuisance calls or spam • Unwarranted surveillance
	Loss of control of personal data	Harms from thwarted expectations, through misuse, repurposing, unwanted retention or continued use and sharing of personal data, including a lack of commitment to the accuracy of data or lack of transparency	• Injury to peace of mind and ability to manage risk • Restrictions on ability to access or review use of personal data • Incompatible repurposing leading to emotional distress
	Lack of autonomy; manipulation and influence	Restriction, coercion, or manipulation of people's choices or their ability to make an informed choice	• Unwarranted nudging leading to poor decisions • Restriction of choice due to power and information asymmetry
	Psychological Harms	Negligently, knowingly, or purposefully paving the way for emotional distress or disturbance (embarrassment, anxiety, fear) to occur	• Detriment to mental health • Loss of sense or control of identity • Distressed relationships • Loss of confidence • Reputational loss/loss of standing • Harassment or bullying
	Chilling effects	Reduced use of services or activities due to an actual or perceived risk of potential	• Reduced activities requiring good credit rating

		harm. And hence, a reduction in the benefits that may have been achieved.	• Reduced use of beneficial products or services that require sharing of data due to perceived risk
	Adverse effects on rights and freedoms	Negative impacts on rights and freedoms in and of themselves	• Restrictions to data privacy rights • Restrictions to freedom of assembly • Chilling effects on freedom of expression
Societal	Damage to law and justice	Restrictions on or subversion of legislative intent, or legal or judicial process	• Creating a route for widescale subversion of a law • Chilling effects on victims or witnesses
	Damage to media, democracy, information and public discourse	Negative impacts on media, democracy information and public discourse at a societal level	• Mistrust in handling of electoral role influencing elections or voter turnout • Widespread mistrust leading to chilling effects on freedom of expression
	Damage to public health	Harms resulting in adverse health outcomes for society	• Mistrust in handling of health data leading to chilling effects on health service use
	Damage to the economy	Negative impacts on the economy that are significant at the local, regional, or national level, or for a specific sector	• Loss of trust from widespread privacy abuses leading to chilling effects on major services • Misuse of personal data leading to unfair competitive advantage
	Damage to the environment	Negative impacts on the environment either directly or indirectly resulting from misuse of data or mitigation of associated risk.	• High energy use associated with data mining, storage and sharing • Loss of ecological diversity and/or green space due to land use for server farms