



# TRANSPARENCY IN THE REGULATION OF ONLINE SAFETY

RESEARCH REPORT FOR OFCOM

May 2021



# Contents

<b>Executive Summary</b>	<b>6</b>
<b>Introduction</b>	<b>10</b>
<b>Methodology</b>	<b>10</b>
<b>Trust, Transparency and Accountability</b>	<b>10</b>
<b>The role of transparency</b>	<b>11</b>
<b>Meaningful Transparency</b>	<b>13</b>
<b>Framing Transparency</b>	<b>15</b>
<b>The current state of transparency reporting</b>	<b>18</b>
<b>Models and Mechanisms Handbook</b>	<b>20</b>
<b>Common Mechanisms of Transparency</b>	<b>20</b>
<b>Applying lessons from other sectors</b>	<b>23</b>
<b>Advertising</b>	<b>23</b>
<b>Asset &amp; Wealth Management</b>	<b>27</b>
<b>Aviation</b>	<b>30</b>
<b>Data Services</b>	<b>33</b>
<b>Environment</b>	<b>36</b>
<b>Fashion, Textiles &amp; Clothing</b>	<b>40</b>
<b>Food &amp; Drink</b>	<b>44</b>
<b>Media, Entertainment &amp; Film</b>	<b>48</b>
<b>Financial Services</b>	<b>51</b>
<b>Utilities</b>	<b>55</b>
<b>Applying lessons from legislation</b>	<b>58</b>
<b>Australia – Online Safety Act</b>	<b>58</b>
<b>France – Avia Law</b>	<b>61</b>
<b>European Union - Mandatory Non-Financial Disclosures</b>	<b>65</b>
<b>European Union – General Data Protection Regulation (GDPR)</b>	<b>68</b>
<b>European Union – Digital Services Act</b>	<b>71</b>
<b>UK Modern Slavery Act</b>	<b>75</b>
<b>Applying lessons from other approaches</b>	<b>79</b>
<b>Australia – ACMA Annual Compliance Priorities</b>	<b>79</b>
<b>The Santa Clara Principles</b>	<b>81</b>

<b>The Transparency Reporting Toolkit: Content Takedown Reporting</b>	<b>83</b>
<b>Summary of Principal Mechanisms Employed</b>	<b>86</b>
<b>Opportunities for Ofcom</b>	<b>87</b>
<hr/>	
<b>Looking Forward - Adaptive Regulation</b>	<b>87</b>
<b>Themes for exploration in transparency regulation</b>	<b>88</b>
<b>Annex</b>	<b>90</b>
<hr/>	
<b>Transparency Concept to Implementation Process Diagram</b>	<b>90</b>
<b>Defining Trust, Transparency and Accountability</b>	<b>90</b>
<b>References</b>	<b>91</b>
<b>Table of Figures</b>	
Figure 1: Summary of the principal transparency mechanisms employed in each regulated sector .....	8
Figure 2: The Trust Triangle Model .....	10
Figure 3: Mechanisms Icons .....	22
Figure 4: Summary of the principal transparency mechanisms employed in each regulated sector .....	86

## Foreword

In February 2021, Ofcom appointed PA Consulting to conduct a literature review of existing transparency schemes and reporting policies, and the extent to which they could be applied to improve online safety within a regulatory regime. Our priority for this research project was to answer the question: What models of transparency exist or have been suggested for use in transparency locally and internationally in other sectors, and online safety in particular?

This work was commissioned as a concise study to identify and summarise the primary mechanisms of transparency to inform approaches for transparency initiatives in online safety. To achieve this, we followed an agile, mixed-methodology approach using a combination of a desk-based literature review, supported by interviews with regulators, thinktanks and academics from international organisations including the UK, USA, Australia, and the Netherlands.

Interview participants were assured that their comments would not be attributable, which has enabled them to frankly share their considerable knowledge, insights, and observations. We also hope that these individuals will become an important peer review/reference group as Ofcom develop their thinking and proposals around the application of transparency mechanisms in online safety.

PA would like to thank all those individuals and organisations that participated in the research for their perspectives, support, and guidance throughout the project.

# Executive Summary

## Introduction

Transparency, particularly in the context of business and government, describes the concept of an entity, usually an organisation, sharing information (either on a voluntary or compulsory basis) to show its operations to a relevant audience. The act of 'being transparent' aims to discourage poor practice and build trust as organisations can be held responsible for their activities that are now open to scrutiny. The audience or recipient of information from transparency initiatives can vary, as can the method of communicating such information. For example, information may be published for all of the public to see live on a website or be written in an annual report.

When referring to 'models of transparency' throughout this report, we mean both a description of the main ideas underpinning transparency and the mechanisms by which transparency regimes are implemented and applied e.g., transparency reports, 'traffic light' nutritional labelling and complaints data disclosure.

## The nature and role of transparency

Despite the broad application of transparency as a regulatory tool in many sectors, the concept and efficacy of transparency is debated. Notable areas of consideration include:

- The logic of transparency assumes that observation of malpractice alone is enough to see subsequent well-intended action to resolve issues or risks identified
- Transparency can have a number of limitations and unintended consequences e.g., the impact of displacing the risk of a harm (for example pollution) from occurring within regulated spaces to unregulated spaces
- A causal link between a transparency initiative and its specific outcomes can be hard to attribute and measure
- There is a significant difference between genuine or 'meaningful' transparency initiatives that add value vs transparency for show or 'theatre'
- Despite debate and scepticism, transparency is fundamentally important for a range of reasons, including exposing malpractice, protecting human rights, and enabling consumer agency.

## Steering away from transparency in the abstract

As highlighted by Gaventa & McGee (2013) and also in several interviews, a challenge in transparency discourse stems from discussions of transparency as an abstract concept. Consequently, it is best to avoid talking about transparency 'in the abstract' and focus on the tailored application of individual transparency initiatives to their specific usage cases. For example, in the case of online safety regulation, this could involve tailoring transparency initiatives to specific areas, e.g. automated content moderation, rather than attempting to frame a singular method of reporting for online safety overall.

## Transparency in online safety

Whilst a mandated regulatory framework for transparency in online safety is a relatively new concept, many platforms have been voluntarily disclosing information under transparency initiatives for several years, starting with Google's first Transparency Report in 2010.

Given the relative immaturity of the social media sector, the variety in nature and quality of their disclosure is not unexpected. In this sense, it might be considered that despite scrutiny, current transparency reporting at least represents a start. As noted during interviews, 'any transparency is better than nothing'<sup>1</sup>. With that said, the current state of transparency reporting has been scrutinised for several reasons with the primary criticisms centred around three points:

- Current reporting focuses on disclosing aggregated quantitative data on the occurrences of harms (prevalence) and subsequent moderation interventions. This type of static reporting<sup>2</sup> lacks utility for both users in understanding platform safety and researchers attempting to assess the efficacy and impacts of content curation, amplification, and moderation.
- A lack of transparency in the design of content moderation, curation and amplification processes has led to calls for improvement in disclosure to understand the development of these processes end-to-end.
- The information disclosed under current transparency initiatives lacks auditing and therefore the credibility and accuracy of this transparency reporting (and subsequent judgements on the nature and efficacy of an organisation's processes) cannot be verified.

## Common transparency models and mechanisms

In conducting this research, during which we have examined theory and practice of transparency in a diversity of settings, we observed that despite variety in the environments that transparency as a regulatory tool is applied (e.g., aviation vs food safety), there exists commonality in the mechanisms of implementation.

Specifically, we've identified **14 common mechanisms** that were prevalent across the case studies explored: Engagement, Exposure, Education (consumers), Education (Industry), Empowerment, Performance and Behavioural Accountability, Investment, Confidentiality, Regulatory Transparency, Audit, Reporting (one-way, two-way and performance) and Researcher Access (outlined in section 3).

To demonstrate this commonality, in the 'handbook' section of this report, each of the 14 transparency mechanisms is represented with a designated icon, which appear throughout the models and mechanisms to highlight their usage in different scenarios. These are summarised in Figure 1 below.

---

<sup>1</sup> Interview 3, Interview 8, 2021

<sup>2</sup> Static reporting relates to reports created using historical data within a specific timeframe, unlike dynamic reporting which can use real-time data (for example via a live dashboard).

Case Studies	Principal Mechanisms Employed													
	Engagement	Exposure	Education (Industry)	Education (Consumers)	Empowerment	Performance and behavioural accountability	Investment	Confidentiality	Regulator Transparency	Audit	One-way occurrence reporting	Two-way reporting and feedback	Performance reporting	Researcher Access
Advertising	✓	✓		✓	✓	✓			✓			✓		
Asset & Wealth Management			✓	✓	✓	✓	✓					✓	✓	
Aviation	✓		✓		✓				✓			✓		
Data Services (The 5 Safes Framework)										✓				✓
European Pollutant Release and Transfer Register (E-PRTR)				✓	✓	✓						✓		
Fashion Textiles & Clothing				✓	✓	✓						✓	✓	
Food & Drink	✓			✓	✓							✓		
Media, Entertainment & Film	✓			✓	✓							✓		
Financial Services				✓	✓	✓						✓		
UK Modern Slavery Act				✓		✓						✓	✓	
Ofgem Complaints Reporting (Utilities)				✓		✓							✓	
EU Mandatory Non Financial Disclosures				✓		✓				✓		✓		
The Santa Clara Principles				✓		✓						✓		
The Transparency Reporting Toolkit: Content Takedown Reporting			✓	✓		✓						✓		
Australian - Online Safety Act			✓			✓						✓	✓	
Australian Communications and Media Authority (ACMA)						✓			✓			✓		
French AVIA Law						✓						✓		
German Network Enforcement Act (NetzDG)				✓	✓	✓						✓		
EU General Data Protection Regulation (GDPR)				✓	✓	✓								
EU Digital Services Act						✓						✓		

Figure 1: Summary of the principal transparency mechanisms employed in each regulated sector

## Adaptive regulation

A key reflection from this research is the potential value of an adaptive approach to online safety regulation. It can be observed that the very factors that pose significant challenges to transparency in this context (international scale, pace of innovation, skills & information asymmetry) also represent opportunities for a regulator to evolve to meet each challenge head-on. Through international collaboration, flexible frameworks and the iterative upskilling of technical and policy expertise, a regulator, in mirroring the fluidity of the organisations it seeks

to regulate, may find it easier to keep pace and evolve regulatory solutions. Additionally, to support operating in an environment whereby regulatory demands are constantly shifting, a regulator may consider frequently stress-testing and gaming their transparency regime to identify flaws, gaps, or risks in reporting requirements.

## **Themes for exploration in transparency regulation**

Over the course of this research, we also identified a number of themes that emerged (primarily from interview discussions) to explore as potential focuses in the development of a regulatory regime. These are detailed in section 4 of this report and include concepts such as ‘sandboxes’ for exploration of new technologies; tailoring reporting requirements to different platform sizes/capabilities; and, understanding the impact of skills and information asymmetry.

## **International Collaboration**

Given the range of activity in regulating online safety internationally, seeking international collaboration in approaches, particularly regarding the definitions and taxonomies of harm, could be valuable in strengthening the robustness of cross-border regulation and seeing the benefits of transparency (e.g. human rights protections) cascade beyond the borders of the UK.

# Introduction

## Methodology

The methodology of this research followed a mixed-methods approach using a combination of a desk-based literature review supported by 20+ interviews with regulators, thinktanks and academics from organisations across the world including the UK, USA, Australia, and the Netherlands.

In acknowledgement of sensitivities that may exist in the subject area of online safety, participants' individual responses have been anonymised when referenced throughout the report.

## Trust, Transparency and Accountability

The terms trust, transparency and accountability are often used in conjunction when applied to discussions of transparency initiatives, understood as closely linked concepts. Meijer (2014) notes how transparency is often connected to accountability<sup>3</sup> and Mabillard & Zumofen discuss how transparency and accountability have been extensively analysed together with academia and subsequently are often “treated together as pillars of good governance” (Mabillard & Zumofen, 2016)<sup>4</sup>. In addition to the association of transparency and accountability, trust forms the third element that may be considered a variable for the efficacy of transparency and accountability<sup>5</sup>, with differing levels of trust amongst citizens impacting the effectiveness of transparency and accountability initiatives. For example, if an organisation itself or the disclosure it offers is not trusted, the credibility of any transparency reporting can be undermined and therefore accountability may be less impactful.

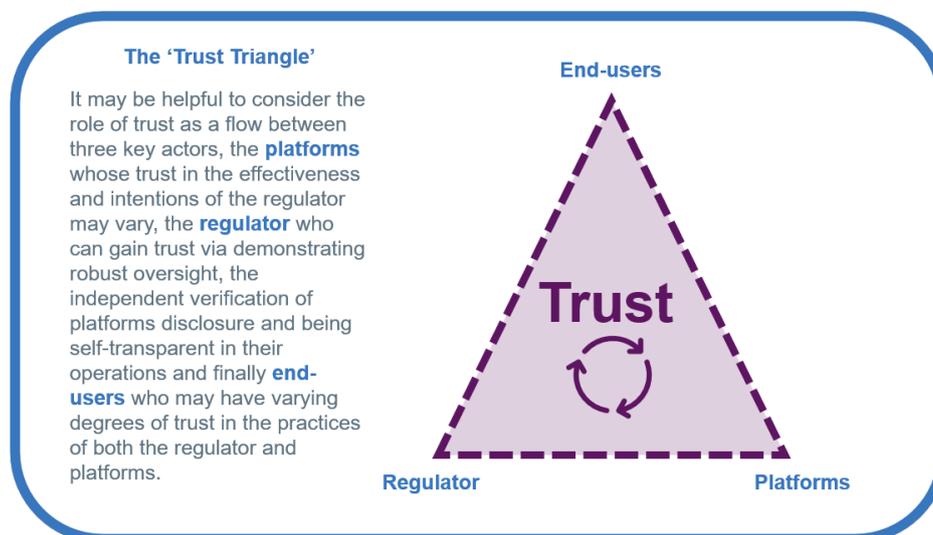


Figure 2: The Trust Triangle Model

<sup>3</sup> Meijer in Mabillard & Zumofen, 2016: 'The complex relationship between transparency and accountability: A synthesis and contribution to existing frameworks'  
<https://www.researchgate.net/publication/304070388> The complex relationship between transparency and accountability A synthesis and contribution to existing frameworks

<sup>4</sup> Ibid.

<sup>5</sup> Ibid.

## The role of transparency

- **Exposing Malpractice ‘Sunlight as Disinfectant’** – In the essay ‘What Publicity Can Do’ (published in 1914 as part of the series ‘Other People’s Money and How the Bankers Use it’), Louis Brandies discusses the role of publicity as a remedy in exposing and subsequently resolving malpractice in banking and other industries. In doing so, Brandies famously coined the phrase ‘Sunlight as Disinfectant’ in the reference “Sunlight is said to be the best of disinfectants; electric light the most efficient policeman”<sup>6</sup>. Subsequently, the notion of sunlight as disinfectant has become closely associated with transparency as a description of its value in exposing malpractice and driving accountability from organisations. With that said, this phrase despite popularity has not gone unscrutinised with scholars such as Etzioni noting the limitations of transparency and subsequently of the transparency as a disinfectant ideal<sup>7</sup>. Nonetheless, the analogy arguably maintains its usage as a brief explainer of the core purpose of transparency in exposing and subsequently remedying malpractice.
- **Enabling due diligence** – Transparency is recognized as a key component of information disclosure for mandatory due diligence regimes. For example, the European Parliament’s recent recommendations on corporate due diligence and corporate accountability “highlights comprehensive transparency requirements” as a “crucial element of legislation on mandatory due diligence”<sup>8</sup>.
- **Protecting Human Rights** – Transparency is understood as key for protecting human rights on several fronts including free speech, press autonomy, fair elections, and access to justice. Using freedom of information as a focus point, Birkinshaw (2006), in discussion of the transparency and the openness of information, notes how freedom of information is fundamental for the realization of other human rights such as freedom of speech and access to justice<sup>9</sup>. In this sense, we can see transparency as an enabling factor for the preservation of other human rights.

---

<sup>6</sup> Brandies, 1914: ‘Chapter V- What Publicity Can Do’ <http://louisville.edu/law/library/special-collections/the-louis-d.-brandeis-collection/other-peoples-money-chapter-v>

<sup>7</sup> Etzioni, 2010: ‘Is Transparency the Best Disinfectant?’ <https://www.uv.es/sasece/docum/mayo-sept2010/295Transparency.pdf>

<sup>8</sup> European Parliament, 2021: ‘Corporate due diligence and corporate accountability: European Parliament resolution of 10 March 2021 with recommendations to the Commission on corporate due diligence and corporate accountability’ [https://www.europarl.europa.eu/doceo/document/TA-9-2021-0073\\_EN.pdf](https://www.europarl.europa.eu/doceo/document/TA-9-2021-0073_EN.pdf)

<sup>9</sup> Birkinshaw, 2006: ‘Transparency as a Human Right’ [Abstract] <https://britishacademy.universitypressscholarship.com/view/10.5871/bacad/9780197263839.001.0001/upso-9780197263839-chapter-3>

- **Protecting Digital Rights** - Looking specifically to transparency for human rights in information and communications technology, transparency also features as a core component of the Global Network Initiatives “GNI Principles” on advancing and protecting the enjoyment of human rights in ICT globally<sup>10</sup>. Principle five on ‘Governance, Accountability and Transparency’ outlines “Participants will be held accountable through a system of (a) transparency with the public and (b) independent assessment and evaluation of the implementation of these Principles”<sup>11</sup>. In contrast, the importance of transparency for human rights online and particularly within the context of content moderation online can be demonstrated by examples whereby a lack of transparency has in-fact been detrimental to human rights. A notable example stems from calls for greater transparency in the moderation processes operated by the Global Internet Forum to Counter Terrorism (GIFCT) amidst concerns that moderation without due scrutiny could be detrimental to human rights, for example by automatically taking down and therefore unintentionally masking evidence of war crimes<sup>12</sup>.
- **Protecting Children’s Rights** - In February 2021, the UN Committee on the Rights of the Child adopted General Comment 25<sup>13</sup>, recognising the application of the 1989 UN Children’s Right Convention in relating to the ‘digital environment’<sup>14</sup>. This “game changer”<sup>15</sup> (Livingstone, 2021) introduces the need for considerations in catering for the rights of the child in the broader preservation of human rights online via transparency. The Children’s Research Network considers the role of a rights-based mechanism for children to have access to redress and transparency as part of “supporting children to lead positive online lives”<sup>16</sup>. Morgan, 2018 contemplates transparency for children in digital data protection as ‘the biggest transparency challenge of all?’<sup>17</sup> noting the challenge of raising awareness amongst children of the significance of data privacy online.
- **Promoting Consumer Agency** – Transparency in product or service offerings, including in manufacturing ethics, sustainability and pricing strategy can promote consumer agency in giving consumers the information necessary to compare and contrast between product or service providers. One of the clearest demonstrations of this is via price transparency which, As Boom, 2010 notes, forms a basic element of market competition<sup>18</sup>, “price *intransparency* can be detrimental to consumer decision making and may be associated with market failure” (Boom, 2010). However, it should also be considered that consumers may not change their behaviour based upon improved disclosure, as discussed throughout the FSA paper ‘Transparency’<sup>19</sup>. Baldwin discusses this as the ‘consumer sovereignty’ perspective, that the citizen-consumer knows best and therefore should be able to exercise informed choice<sup>20</sup>.

<sup>10</sup> Global Network Initiative: ‘The GNI Principles’ <https://globalnetworkinitiative.org/gni-principles/>

<sup>11</sup> Ibid.

<sup>12</sup> Human Rights Watch, 2020: ‘Joint Letter to New Executive Director, Global Internet Forum to Counter Terrorism’ <https://www.hrw.org/news/2020/07/30/joint-letter-new-executive-director-global-internet-forum-counter-terrorism>

<sup>13</sup> Souter, 2021: ‘Inside the Digital Society’: Children’s rights in the digital society’ <https://blogs.lse.ac.uk/medialse/2021/04/12/inside-the-digital-society-childrens-rights-in-the-digital-society/>

<sup>14</sup> Ibid.

<sup>15</sup> Livingstone, 2021: ‘Children’s rights apply in the digital world!’ <https://blogs.lse.ac.uk/medialse/2021/02/04/childrens-rights-apply-in-the-digital-world/>

<sup>16</sup> Children’s Research Network, 2020: ‘Do children have rights in the online world and can they be enforced? An analysis of the Council of Europe’s ‘Guidelines on Children’s Rights in the Digital Environment’ from an Irish perspective.’ <https://www.childrensresearchnetwork.org/knowledge/resources/article-child-rights-online-world>

<sup>17</sup> Morgan, 2018: ‘The Transparency Challenge: Making children aware of their data protection rights and the risks online’ <https://www.dataprotection.ie/sites/default/files/uploads/2019-02/TransparencyChallenge.pdf>

<sup>18</sup> Boom, 2010: ‘Price Intransparency, Consumer Decision Making and European Consumer Law’ <https://link.springer.com/content/pdf/10.1007/s10603-011-9163-8.pdf>

<sup>19</sup> FSA: ‘Discussion Paper DP13/1 Transparency’ <https://www.fca.org.uk/publication/discussion/fsa-dp13-01.pdf>

<sup>20</sup> Baldwin, 1999: ‘Understanding regulation: Theory, Strategy and Practice’

- **Promoting Market Competition** – As aforementioned in Boom, 2010 a generic depiction of the role of transparency is in the role of stimulating market competition through driving organisations to beat competitors. Møllgaard & Overgaard, 2001 describe this perhaps generalised viewpoint as the ‘consumer protection view’<sup>21</sup>, outlining “improved market transparency enables consumers to compare substitute goods or services and their prices. Consequently, it should make consumers more sensitive to perceived differences in the mix of price and characteristics and thus, intensify competition between rival suppliers. As a contrast, in a non-transparent market, rivals can free-ride on consumer ignorance, create artificial lock-in and stifle price competition.” (Møllgaard & Overgaard, 2001). However, later in their discussion Møllgaard & Overgaard conclude that increased market transparency is not always a good thing and assessments on efficacy should be based upon specific cases at hand<sup>22</sup>. Offering an alternate method of stimulating competition, Baldwin notes how regulators can promote ‘regulatory competition’ through variation in their regulatory regimes, for example changing standards to make regulation more cost effective or attractive to industry<sup>23</sup>.

## Meaningful Transparency

One of the hurdles of implementing transparency in any context is assessing the difference between reporting for ‘theatre’<sup>24</sup> vs genuine transparency, acknowledging the distinct and fundamental difference between the provision of information as a ‘tick-box’ compliance exercise vs information that can empower honest and genuine insights; ergo can be considered *meaningful transparency*. To understand if information disclosed will provide a mechanism for transparency that can be considered meaningful, two key checkpoints might be considered: Firstly, does the information disclosed *mean* anything to the intended audience? Secondly, will this type of information lead to effective outcomes for transparency and accountability?

**Meaningful Disclosure** - A broad-brush definition of transparency alone risks failing to capture the complexity of the fact that the demand of transparency reporting varies between audiences, i.e. *transparency means different things to different people*. Therefore, to measure the effectiveness of disclosure requirements, considering if a set of information will be meaningful to its specifically intended audience will help to assess if that information is truly useful or not.

**Meaningful Outcomes** – Having assessed if the information disclosed is relevant and meaningful to its intended audience, a second checkpoint should be considered to assess the broader efficacy of the information disclosed in driving the desired outcomes *e.g., does this information increase the accountability of an organisation and subsequently enable consumers the agency to make informed decisions and act freely in their consumption choices?*

<sup>21</sup> Møllgaard & Overgaard, 2001: ‘Market transparency competition policy’ <https://www.econstor.eu/bitstream/10419/208446/1/cbs-wp2001-06.pdf>

<sup>22</sup> Ibid.

<sup>23</sup> Baldwin, 1999: ‘Understanding regulation: Theory, Strategy and Practice’

<sup>24</sup> Douek, E, 2020: ‘The Rise of Content Cartels’ <https://knightcolumbia.org/content/the-rise-of-content-cartels>

## Accessible, Accurate, Manageable, Verifiable

Reflecting upon lessons from other sectors, the development of meaningful transparency appears dependant on **four core pillars (Accessible, Accurate, Manageable, Verifiable)** that rest upon the foundation of clear communication of information to all interested (and eligible) parties, recognising communication mechanisms vary to meet the demands of different audiences. In order to be considered 'meaningful', the information an organisation is disclosing needs to be presented in a format that is **accessible** (including physical access to the information and being able to understand it) and **accurate** to the extent that it can be utilised for the purpose intended, by the intended recipient audience without challenge and enable the production of accurate judgement or insights.

Applying these principles to online safety, Suzor et al. (2019) highlight the shortfalls of information not being accessible and understandable, drawing light on possibly the two most demanded functions for transparency in the broader online platforms regulation debate: transparency in the end-to-end moderation process and transparency in the deployment of algorithms for moderation, content curation and amplification<sup>25</sup>. In this discussion, Suzor et al. (2019) suggest that not only does there lack accessibility to understanding the logic and process of moderation processes and algorithmic functionality, but that existing platform transparency reports focused on disclosing aggregated statistical data fail to provide a level of understanding sufficient in clarity and detail to hold platforms accountable<sup>26</sup>.

Whilst it is important that information disclosed is **accessible** and **accurate**, another significant obstacle presents itself in not having the physical capability to manage either the type or quantity of data or information reported. This challenge requires a third pillar to consider - information should be reported in an appropriate format and quantity such that it is **manageable** to the recipient. Unmanageable information is not only laborious and frustrating, it poses a risk for exploitation as an opportunistic method for organisations to hide malpractice amidst swathes of data, safe in the knowledge that the recipient lacks the capability or manpower to critically assess all information disclosed; a concept that is often compared to the famous literary remark "*Where does a wise man hide a leaf? In the forest*"<sup>27</sup>.

The final pillar relates to the level that disclosure is **verifiable**, a matter of quality assurance to ensure that information disclosed is of satisfactory accuracy such that it is fit for purpose and will lead to accurate insights into an organisation's operations, processes, and compliance to reporting requirements. In the case of current transparency reports published by platforms, there lacks the opportunity to independently verify or audit information disclosure and subsequently the quality of the reporting, and the organisational processes behind the data, cannot be assured<sup>28</sup>.

---

<sup>25</sup> SUZOR, Nicolas P. et al. What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation. *International Journal of Communication*, [S.l.], v. 13, p. 18, mar. 2019..

<sup>26</sup> Ibid.

<sup>27</sup> G.K Chesterton, 1911: 'The Sign of a Broken Sword'

<sup>28</sup> Tony Blair Institute for Global Change, 2020: 'Online Harms: Bring in the Auditors' <https://institute.global/policy/online-harms-bring-auditors>

## Framing Transparency

Despite the broad application of transparency as a regulatory tool across many sectors, the concept and efficacy of transparency remains contested within academic literature. The following section outlines several key works on transparency, with the purpose of providing contextual understanding to the concept of transparency and the potential ‘lessons learned’ from these contextual literary works for application to online safety regulation.

- **Acknowledging assumptions** – Ananny & Crawford (2018) discuss that the logic of transparency rests upon the fundamental assumption that observation of a ‘phenomenon’ (e.g., a particular harm) alone is enough to see subsequent well-intended action to resolve the issue or malpractice initially observed. In part, this assumption is fuelled by attempts to resolve the fear of the unknown and fear of the uncontrollable, i.e., assuming that if there is awareness of harm via transparency, it will subsequently be corrected or resolved; “Transparency includes an affective dimension, tied up with a fear of secrets, the feeling that seeing something may lead to control over it” (Ananny & Crawford, 2018). Within the context of online safety, this understanding can be applied for the purpose of managing expectations and expected outcomes, acknowledging the difference between observing malpractice (transparency) and correcting malpractice (accountability and corrective action)<sup>29</sup>.
- **Acknowledging limitations** – Ananny & Crawford also highlight **10 key limitations of transparency** that should be acknowledged to demonstrate the constraints of transparency in achieving intended outcomes overall. These limitations are noted in summary below, although readers should see ‘Seeing Without Knowing’ (Ananny & Crawford 2016:7-10) for full detail<sup>30</sup>.
  - 1 **Transparency can be disconnected from power** – Without meaningful impact, transparency can lose power and purpose and if corruption continues following transparency, it can even contribute to growth in corruption.
  - 2 **Transparency can be harmful** – Full transparency can be harmful by threatening privacy and inhibiting honest conversations.
  - 3 **Transparency can intentionally occlude** – Transparency can be gamified to intentionally occlude, for example an organisation printing out hundreds of records in paper for manual review to reduce the likelihood of malpractice being identified and exposed, referred in text to as ‘resistant transparency’.
  - 4 **Transparency can create false binaries** – Where transparency initiatives lack in-depth understanding of the accountability and visibility intended to be produced, transparency can be undermined as a false choice between secrecy and openness.

---

<sup>29</sup> Ananny & Crawford, 2018: ‘Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability’

<sup>30</sup> Ibid.

- 5 **“The ideal of transparency places a tremendous burden on individuals”<sup>31</sup>** – The ideal of transparency can imply that individuals who have to themselves seek out, digest and assess the significance of information disclosed under transparency requirements assumes firstly that individuals are able to understand the systems and processes in which the information relates to, and secondly that individuals will subsequently make informed choices that will lead to positive social change<sup>32</sup>.
  - 6 **Transparency does not necessarily build trust** – Despite transparency being often thought to ‘engender trust of organisations and systems’ there is little evidence to confirm this, and trust differs between how stakeholders perceive a given set of information.
  - 7 **Transparency entails professional boundary work** – Effective transparency can be restricted by professionals protecting the ‘boundaries’ of their expertise, i.e., not wishing to openly share their expertise in fear of scrutiny.
  - 8 **Transparency can privilege seeing over understanding** - Transparency risks emphasising the value of access and observation to information over the physical ability of recipients to manage, understand and subsequently effectively use the information they have access too.
  - 9 **Transparency has technical limitations** – Transparency, particularly in the context of online safety and algorithmic transparency can be constrained by a lack of technical understanding.
  - 10 **Transparency has temporal limitations** – The period in time a transparency initiative is launched risks misrepresenting the overall picture of a given subject area as it alerts attention to only the present state of a process or operation, missing out key contextual understanding that could be gleaned from analysing previous iterations.
- **Steering away from abstract transparency** - A point highlighted by Gaventa & McGee (2013)<sup>33</sup>, and also in several interviews, suggests it may be more productive to avoid talking about transparency as an abstract concept and focus on the tailored application of individual transparency initiatives to their specific usage cases rather than theorising transparency as an over-arching and catch-all phrase. For example, in the case of online safety regulation, this could involve tailoring transparency initiatives to specific areas, e.g., automated content moderation, rather than seeking to apply a singular broad concept of transparency to resolve a multiplicity of issues under the wide-reaching category of online safety in its entirety.

**INTERVIEW 7** To talk about transparency in the abstract is useless. In theory, transparency could mean I have access to all of the information in your cell phone... this is obviously very different to the transparency that is needed to inform public policy.

<sup>31</sup> Ibid.

<sup>32</sup> ‘Seeing Without Knowing’ (Ananny & Crawford 2016:7-10)

<sup>33</sup> Gaventa & McGee, 2013: ‘The Impact of Transparency and Accountability Initiatives’

- **Understanding context** - When assessing the use of transparency initiatives and their subsequent implementation, caution might be aired when assessing the efficacy of a given transparency initiative based upon previous usage cases in different sectors or usage cases. As noted by Gaventa & McGee, judgements on the intended outcomes of a given transparency initiative should look at the specific environmental context of its potential application rather than assuming its efficacy based upon successes in other usage cases which cannot be guaranteed<sup>34</sup>.
- **Understanding causality** – Michener (2018) notes, despite there being some well-evidenced examples of the positive cause and effect of transparency initiatives, it is incredibly difficult to objectively measure the causality of outcomes as a result of a specific transparency initiative or policy, and research into the efficacy of transparency policies has in-fact overrepresented ‘direct, quantifiable evidence of causality’ whilst in reality most policies produce ‘diffuse and indirect impacts’. In application to online safety regulation, understanding the difficulty in assessing quantifiable causality vs the likelihood of indirect impacts may be useful to incorporate into planning the usage of different mechanisms and attempting to pre-empt likely ‘knock-on’ or unintended consequences of a given transparency initiative beyond that of the primary intended outcome<sup>35</sup>. One particular consideration for unintended consequences, as discussed by Baldwin<sup>36</sup>, might be given to the impact of displacing the risk of a harm (for example pollution) from occurring within regulated spaces to unregulated spaces as a result of ‘over-stringent’ regulatory demands.
- **Privacy or Safety?** – Wood (2021) highlights how discussions regarding freedom of expression online often begin by referencing the “trade off” between regulating digital media platforms whose products harm the public and preserving free speech. Wood notes how this argument is perceived as a zero-sum game – that increasing public safety means reducing freedom of speech. However, Wood notes that in fact this notion of a trade-off or ‘zero-sum game’ is false and undermines ‘proper’ government intervention attempts in the ‘information marketplace’. To avoid this, Wood argues for a ‘systems-based’ approach to regulation, focusing on moderating the systems of amplification rather than moderating user content– “rather than focusing on content, it is better to focus on (re)designing these curation systems so that harmful material stops being algorithmically and automatically spread at scale. (Wood, 2021)<sup>37</sup>. For transparency initiatives, this could apply to requesting disclosure on the processes behind a content moderation system rather than the number of interventions/takedowns or the nature of content itself.
- Alternatively, the use of privileged researcher access could alleviate privacy concerns, allowing access to sensitive data for necessary research or regulatory purposes by vetted researchers in a controlled environment whilst maintaining the privacy and confidentiality of user data and proprietary information<sup>38</sup>.

---

<sup>34</sup> Ibid.

<sup>35</sup> Michener, 2018: ‘Gauging the Impact of Transparency Policies’

<sup>36</sup> Baldwin, 1999: ‘Understanding Regulation: Theory, Strategy and Practice’

<sup>37</sup> Wood, 2021: ‘Online harms: why we need a systems-based approach towards internet regulation’

<https://blogs.lse.ac.uk/medialse/2021/02/19/online-harms-why-we-need-a-systems-based-approach-towards-internet-regulation/>

<sup>38</sup> Interview 7, 2021

## The current state of transparency reporting

Whilst transparency reporting in tackling online safety within a mandated regulatory framework remains a relatively new concept, many platforms have been voluntarily disclosing transparency information for a number of years, beginning with Google's first Transparency Report in 2010<sup>39</sup>. To this regard, whilst the nature of information currently disclosed via platforms transparency initiatives has been heavily scrutinised<sup>40</sup> for lacking utility<sup>41</sup>, it has also been noted that the recent uptake of transparency represents a strong start<sup>42</sup>.

Additionally, in the context of the age of some major tech platforms such as TikTok (5 years old), and Clubhouse (1 year old) vs the establishment of regulation in industries such as aviation (Civil Aviation Authority established 1972), it may be noted that platforms themselves are relatively new and therefore iteration in the quality of their disclosure is to be expected. In this sense, it might be considered that despite scrutiny, current transparency reporting represents a start, as noted during interviews, 'any transparency is better than nothing'<sup>43</sup>. With that said, the current state of transparency reporting has been heavily scrutinised for a number of reasons with the primary criticisms centred around three points outlined in summary below.

- **A lack of utility for aggregated data** – Information disclosed under current transparency initiatives tends to focus on aggregated occurrence and takedown data which lacks utility for both users in understanding platform safety and for researchers in attempts to assess the impact and efficacy of a technology company's activities.

**INTERVIEW 5** It's not clear who the audience for transparency reports actually is. As a result, the reports themselves do not really do a good job for anyone. They are too detailed for users but lack the detail for researchers.

- **A lack of transparency in end-to-end content moderation processes** - A lack of transparency in how content moderation, curation and amplification systems are designed and used has led to calls for improvement in disclosure throughout development processes, including: Clear understanding of the logic and justification for platforms developing specific moderation policies; the processes of implementing moderation, the design and usage of algorithms for automated moderation, curation and amplification and details outlining the role of human moderation (factors such as the number of moderators, training levels, skills and qualifications requirements, welfare processes and language capabilities); the processes of reviewing moderation decisions; user appeals processes for moderation decisions; information regarding the sharing of moderation frameworks and technology between platforms (as described by Douek, 2020 in 'The Rise of Content Cartels')<sup>44</sup>.

<sup>39</sup> Google: 'Transparency Report' [https://transparencyreport.google.com/about?hl=en\\_GB](https://transparencyreport.google.com/about?hl=en_GB)

<sup>40</sup> Delivering a Duty of Care, NSPCC (2021)

<sup>41</sup> Douek, E (2020), Suzor et al (2019), Ananny & Crawford (2018) in Suzor et al (2019), Ausloos et al (2020) etc.

<sup>42</sup> Tony Blair Institute for Global Change, 2020: 'Online Harms: Bring in the Auditors' <https://institute.global/policy/online-harms-bring-auditors>

<sup>43</sup> Interview 3, Interview 8, 2021

<sup>44</sup> Douek, 2020 'The Rise of Content Cartels' <https://knightcolumbia.org/content/the-rise-of-content-cartels>

- **A lack of independent verification in transparency disclosure** – As described in ‘Bring in the Auditors’ and mentioned in several interviews, a lack of auditing or independent verification of both the data disclosed and the processes behind it under current voluntary transparency initiatives means that the efficacy and accuracy of transparency reporting cannot be assessed<sup>45</sup>.

## **INTERVIEW 14 Transparency is meaningless unless the data has been checked and audited beforehand.**

In contrary to the position that current transparency represents a strong start, according to Access Now who curate the Transparency Reporting Index, growth of technology companies publishing transparency reports has actually been decreasing since 2013<sup>46</sup> which may be a consideration when assessing the overall state of transparency reporting to date.

---

<sup>45</sup> Tony Blair Institute for Global Change, 2020: ‘Online Harms: Bring in the Auditors’ <https://institute.global/policy/online-harms-bring-auditors>

<sup>46</sup> Access Now: ‘Transparency Reporting Index’ <https://www.accessnow.org/transparency-reporting-index/>

# Models and Mechanisms Handbook

The following section examines 20 case studies of the implementation of transparency initiatives in a range of settings covering two core areas: 'lessons from other sectors' (beyond online safety and communication technology) and 'lessons from international approaches' in regulating online safety.

The content from this section covers a broad range of insights gleaned from an extensive literature review in combination with 20+ interviews with regulators, thinktanks and academics from organisations across the world. It should be acknowledged, particularly with regard to discussions of the efficacy of transparency mechanisms, that where assessments and lessons learned across case studies are unreferenced, they stem from our own interpretations and analysis based upon the reviewed literature and content derived from interview discussions covered within the scope and timeframe of this research. To this end, within the context of analysing the fast-moving environment of online safety regulation, it should be noted that additional developments are likely to have emerged since the time of publication.

## Common Mechanisms of Transparency

Across the variety of transparency initiatives identified throughout this research, **14 common mechanisms** that were prevalent in the case studies were identified. Whilst the local application of these mechanisms varies across applications, with each emphasising different aspects, each of the 14 was prevalent. The development of these mechanisms stems from observations throughout our research (literature and interviews) with the purpose of acting as a guide to help demonstrate the common implementation of transparency mechanisms in a diversity of settings. There may be cases where an unflagged mechanism is occurring but has not been identified or highlighted in this report.

In the following sections, to highlight the application of these common mechanisms in different usage cases, each mechanism is represented with a specific icon (key below) and appears above each example alongside an additional **status** note to show whereby there are transparency initiatives that are yet to be implemented.

## Transparency Mechanisms Icon Key

### Engagement:

Engaging with a range of stakeholders in consultation of creating and updating regulatory initiatives.



### Exposure:

Exposing non-compliance, unethical practice, or mistreatment.  
Exposing scale of harm and measures in place to protect consumers.



### Education (Industry):

Educating product and service providers on general best practice to discourage non-compliance and malpractice.



### Education (Consumers):

Increasing consumer awareness of risks and issues of consumption choices and warning of industry non-compliance or malpractice.



### Empowerment:

Empowering consumers with improved agency and sufficient rights to be able to engage and hold accountable issues that impact them (e.g., data protection under GDPR).



### Performance and behavioural accountability (whole supply chain):

Social, economic, and environmental stewardship.

Diversity, equity, and inclusion.

Human rights: fair labour practices, staff welfare.

As gatekeepers of online discourse.

Application of 'Safety by Design' principles, training, and practices.



### Investment:

Direct investment towards safe and ethical products and services.



### Confidentiality:

Disclosure on a confidential basis that encourages safe pan-industry sharing of safety concerns or non-compliance information.



### Regulator transparency:

Regulator 'self-transparency' to be held account in the efficacy of their transparency initiatives and regulatory activity.



### Audit:

Independent inspection and verification of processes and data to ensure information disclosed under transparency initiatives is credible.

### Reporting:

**1.** One-way occurrence reporting (volume and trends). Disclosure of information that informs consumers and incentivises organisations to make beneficial changes to their conduct.

**2.** Two-way reporting and feedback between industry/regulators to address shared challenges and innovation (e.g., horizon scanning for emerging safety concerns, collaborative development of innovative service models, standards, and evaluation of effectiveness)

**3.** Performance information (including rankings and sanctions/penalties for non-compliance) to enable consumers to make informed decisions and better judge which products and services are most appropriate for their needs.

### Researcher Access:

Provisioning accredited or privileged researcher access to sensitive or proprietary information.

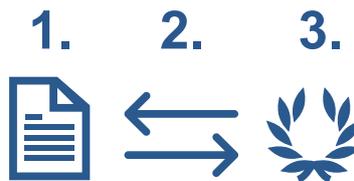


Figure 3: Mechanisms Icons

## Applying lessons from other sectors

The challenge of transparency is unique in its application to online safety. However, lessons can be learned from how transparency as a tool has been implemented in other sectors, as outlined in the subsequent section.

### Advertising

**Status:** Implemented

**Principal mechanisms employed:**



**Summary:** Transparency for the ASA can be understood from several perspectives: Self-transparency; transparency in the outcomes of their interventions; and, transparency for consumers in the advertising they are exposed to.

Mechanisms include annual reporting; stakeholder consultations; publication of rulings and education of consumers and advertisers. The ASA believe this works because they maintain an environment whereby understanding advertising standards and good practice is relatively easy and accessible, and malpractice is exposed contributing to a higher standard of advertising thus reducing the need for regulatory interventions across the board.

## Advertising Standards Authority

### General Overview

The Advertising Standards Authority have been operating to “keep UK advertising ‘legal, honest and truthful’ since their establishment in 1962<sup>47</sup>, focusing initially on ‘non-broadcast’ advertising to eventually expanding their remit in 2004 to have responsibility for TV and radio ads<sup>48</sup>.

The ASA have a wide range of regulatory measures at their disposal, including sanctions and naming non-compliance, but emphasise the fact that they seek to collaborate with advertisers to improve advertising standards in the first instance and see escalating intervention to measures such as publishing non-compliance as a last resort.

Transparency for the ASA can be understood from several perspectives:

- Self-transparency in their activities and efficacy as an organisation.
- Transparency in the outcomes of their regulatory interventions, including highlighting where advertising standards have been broken for specific interventions.

<sup>47</sup> ASA: ‘About ASA and CAP’ <https://www.asa.org.uk/about-asa-and-cap.html>

<sup>48</sup> ASA: ‘Our History’ <https://www.asa.org.uk/about-asa-and-cap.html>

- Transparency for consumers in the advertising they are exposed to - legal, honest, and truthful<sup>49</sup> advertisements.

**Transparency mechanisms:** The ASA operate a range of mechanisms including: self-transparency in their own regulatory performance, engagement with industry and consumer stakeholders, exposure of non-compliance (reputational pressure), the education and subsequent empowerment of consumers with pro-active informative advertisement campaigns and the use of confidential, two-way reporting and collaboration with advertisers to resolve non-compliance before having to escalate and publish breaches of advertising standards in a public sphere.

Looking specifically at methods of deploying these mechanisms there are several including:

- An annual performance report<sup>50</sup> in collaboration with CAP (The Committees of Advertising Practice) and a designated transparency section of their website<sup>51</sup> demonstrating regulatory transparency.
- Engagement via consultations (via CAP/BCAP) on proposed changes to advertising codes involving “everyone whose views should be considered, for example small and large advertisers, professional associations, marketers, consumer groups, media owners, public bodies and individual”<sup>52</sup>.
- The publication and exposure of rulings, and non-compliant online advertisers (including specific ad alerts)<sup>53</sup> on their website<sup>54</sup>.
- Education and empowerment of consumers with pro-active advertising campaigns, highlighting upcoming regulatory action to tackle non-compliance, in some instances with ASA ads appearing directly adjacent to the non-compliant ads in question.
- Education of advertisers on compliance measures, for example notifying social media influencers of ad rules<sup>55</sup> to improve the transparency of paid sponsorship and advertisement on social media platforms.

## What does transparency help with?

**What transparency helps with:** Ultimately, the goal of transparency in advertising standards is to provide consumers with legal, honest, and truthful<sup>56</sup> advertisements. The transparency

<sup>49</sup> ASA: ‘About ASA and CAP’ <https://www.asa.org.uk/about-asa-and-cap.html>

<sup>50</sup> ASA-CAP: ‘Annual Report 2019’ <https://www.asa.org.uk/uploads/assets/5d2b832d-8d12-4f28-9e9f5df54f7b93c1/809de510-6285-4c57-9c904b1c91f6d0ec/ASA-CAP-2019-Annual-Report.pdf>

<sup>51</sup> ASA: ‘Transparency’ <https://www.asa.org.uk/transparency.html>

<sup>52</sup> ASA-CAP-BCAP: ‘How we consult’ <https://www.asa.org.uk/about-asa-and-cap/the-work-we-do/how-we-consult.html>

<sup>53</sup> ASA: ‘Ad Alerts’ <https://www.asa.org.uk/advice-and-resources/resource-library/ad-alerts.html>

<sup>54</sup> ASA: ‘Non-compliant online advertisers’ <https://www.asa.org.uk/codes-and-rulings/non-compliant-online-advertisers.html>

<sup>55</sup> ASA: ‘Influencers’ guide to making clear that ads are ads’ <https://www.asa.org.uk/resource/influencers-guide.html>

<sup>56</sup> ASA: ‘About ASA and CAP’ <https://www.asa.org.uk/about-asa-and-cap.html>

of the ASA in how they enforce this, including the outcomes of intervention, works to improve understanding of advertising standards by industry and consumers alike, with the intended outcome of raising the bar for advertising standards overall.

**Why and how is transparency effective? What is the intended audience for the transparency information? Do the benefits spill out beyond that audience?**

**Why is transparency effective:** The ASA emphasise collaboration with industry and improving the standard of advertising prior to the need for formal intervention against non-compliance. Having transparency throughout this process, both in the operations of the ASA, their standards, and interventions themselves and the communication of these standards to consumers and industry alike creates an environment whereby understanding advertising standards and good practice is relatively easy and accessible.

**How transparency is effective:** Having a transparent regulatory environment between the ASA as a regulator, advertisers, and consumers means that malpractice is easier to identify which can contribute to a higher standard of advertising and reduce the need for regulatory interventions across the board.

**Why was transparency developed? Was it voluntary or mandated by Government?**

Advertising via commercial television broadcasting has been legislated since 1955 and for commercial radio broadcasting since 1973<sup>57</sup>. The Advertising Standards Authority have been operating to “keep UK advertising ‘legal, honest and truthful’ since their establishment in 1962<sup>58</sup>. Whilst ad standards, and subsequently the remit of the ASA, has had to evolve over time to meet changing demand (notably a 1995 extension to cover non-broadcast electronic media<sup>59</sup> and a 2010 extension to cover websites’ non-paid for ad space under their control<sup>60</sup>) the principle of legal, honest and truthful advertising remains the same.

In this sense, whilst the targets of advertising legislation and transparency have evolved from commercial television in 1955 to today’s social media influencers, the reason behind implementing transparency remains essentially unchanged.

**What is the key information / data that is reported?**

**Published rulings<sup>61</sup>:** A description of the ad, the specific issues that the ASA challenges, the response of the advertiser, a broken-down assessment stating whether specific challenges were upheld, a description of the final ASA action.

**Published non-compliant online advertisers<sup>62</sup>:** The background of the company in question (including name, website, and address) the specific advertising issues and indication of the ASA’s actions in response.

**ASA Advertisement Campaigns:** National ad campaigns informing consumers on the ASA’s purpose and operations<sup>63</sup> & ads specifically targeting non-compliant ads.

<sup>57</sup> ASA: ‘Our History’ <https://www.asa.org.uk/about-asa-and-cap/our-history.html#1961%20onwards>

<sup>58</sup> ASA: ‘About ASA and CAP’ <https://www.asa.org.uk/about-asa-and-cap.html>

<sup>59</sup> ASA: ‘Our History’ <https://www.asa.org.uk/about-asa-and-cap/our-history.html#1961%20onwards>

<sup>60</sup> Ibid.

<sup>61</sup> ASA: ‘Rulings’ <https://www.asa.org.uk/codes-and-rulings/rulings.html>

<sup>62</sup> ASA: ‘Non-compliant online advertisers’ <https://www.asa.org.uk/codes-and-rulings/non-compliant-online-advertisers.html>

<sup>63</sup> ASA: ‘National ad campaign’ <https://www.asa.org.uk/about-asa-and-cap/national-ad-campaign.html>

	<p><b>By social media influencers:</b> Ensuring paid ads and brand or product sponsorship are clearly labelled in social media posts through clear labelling e.g., '#ad'.<sup>64</sup></p> <p><b>Voluntary transparency on the ASA's own policies:</b> 'Who we are and what we do', 'What we spend and how we spend it', 'Our priorities and performance', 'How we made decisions', 'Our policies and procedures', 'Our lists and registers', 'The various services that we offer'.<sup>65</sup></p>
<p><b>What else is required to make the model work? What is the necessary context?</b></p>	<p>An additional mechanism that assists the ASA in increasing the reputational pressure upon non-compliant actors is the role of the media in publicising breaches of advertising standards. The role of the media increases the publicity of ASA interventions and therefore increases reputational pressure upon non-compliant advertisers to resolve non-compliance issues.</p>
<p><b>What were the processes and steps involved in implementing reporting?</b></p>	<p>None observed.</p>
<p><b>What are the risks? Are there any unintended consequences?</b></p>	<p>None observed.</p>
<p><b>What lessons can be learned from this model and be applied to Online Safety?</b></p>	<p><b>Regulator self-transparency –</b> Being transparent as a regulator in operations, policies, process of interventions and overall efficacy in online safety regulation.</p> <p><b>Collaborative approach -</b> Acknowledging the benefit of education and consultation with industry and consumers alike to encourage raising the bar for standards and increasing the likelihood of compliance before intervention has to occur.</p> <p><b>The role of reputational pressure –</b> Understanding reputational pressure via publicising non-compliance may be considered a last resort.</p>

<sup>64</sup> ASA: 'Recognising ads: Social media and influencer marketing' <https://www.asa.org.uk/advice-online/recognising-ads-social-media.html#3>

<sup>65</sup> ASA: 'Transparency' <https://www.asa.org.uk/transparency.html>

## Asset & Wealth Management

**Status:** Implemented

**Principal mechanisms employed:**



**Summary:** US Investment Management firm BlackRock signed up to the United Nations Global Compact-supported<sup>66</sup> Principles for responsible Investment (PRI) in 2008. In 2020, they committed to providing their stakeholders and clients more clarity and insight into their investment companies and stewardship practices. Transparency mechanisms include encouraging suppliers and companies in which they invest to provide transparency regarding diversity representation as appropriate to their business and the countries in which they operate.

Transparency helps with education, empowerment, and investment. Transparency reports are also being used to hold directors and businesses accountable for their performances and encouraging change.

Transparency has given insight to clients on the stewardship practices within BlackRock and how they are managing the companies in which they invest thus boosting accountability. Plus, it has driven companies seeking investment to publish crucial information across social, economic, and environmental aspects with the hope that this will lead to behavioural change.

Transparency provides greater information to clients (investors) but actually results in greater benefits for the general population as the transparency reports are made available to anyone.

Asset & Wealth Management - BlackRock	
<b>General Overview</b>	<p>In 2008, US Investment Management firm BlackRock signed up to the Principles for responsible Investment (PRI) with the primary objective of enabling signatory transparency on RI activities and facilitate dialogue between investors and their clients, beneficiaries, and other stakeholders <sup>67</sup>.</p> <p>However, in January 2020, BlackRock committed to providing their stakeholders and clients more clarity and insight into their investment companies and stewardship practices. Following this commitment, they have taken action to deliver increased transparency of their stewardship approach, enhanced reporting on their engagement and voting actions, as well as pushing the companies they invest in to provide regular transparency reports.</p>
<b>What does transparency help with?</b>	<p><b>Transparency mechanisms:</b> BlackRock encourages suppliers and companies in which they invest, to have in place programs and initiatives to advance diversity, equity, and inclusion within their business and to provide transparency regarding diversity representation as appropriate to their business and the countries in which they operate <sup>68</sup>. This is a mix of one and two-way</p>

<sup>66</sup> UNGC: 'Integrate the Principles for Responsible Investment' <https://www.unglobalcompact.org/take-action/action/responsible-investment>

<sup>67</sup> BlackRock: 'BlackRock Earns Straight A's for Transparency by Principles for Responsible Investment' <https://www.blackrock.com/corporate/newsroom/press-releases/article/corporate-one/press-releases/blackrock-earns-straight-a-for-transparency>

<sup>68</sup> BlackRock: 'Supplier Code of Conduct & Ethics' <https://www.blackrock.com/uk/individual/literature/publication/blackrock-supplier-code-of-conduct-and-ethics.pdf>

reporting with information being both just reported and also fed back to drive change.

**What transparency helps with:** Transparency helps with education, empowerment, and investment. As engagements and voting are two of BlackRock's core responsibilities as a fiduciary to their clients, transparency is being introduced to give better insight into their stewardship activities (thus educating and empowering their clients). Transparency reports are also being used to hold directors and businesses accountable for their performances and encouraging positive change<sup>69</sup>.

**Why and how is transparency effective? What is the intended audience for the transparency information? Do the benefits spill out beyond that audience?**

**Why is transparency effective:** In this instance, transparency has been effective on two fronts. Firstly, it has given insight to clients on the stewardship practices within BlackRock and how they are managing the companies in which they invest thus boosting accountability. Secondly, it has driven companies seeking investment to publish crucial information across social, economic, and environmental aspects with the hope that this will lead to behavioural change.

**How transparency is effective:** Pressure to produce transparency reports from BlackRock, with the potential to not receive investment, is hugely influential. Not only does this ensure companies are held to account through these reports, but the repercussions for business growth through the withholding of investment could be huge and provide strong incentives to comply.

**Who:** Both of these provide greater information to the clients (investors) but actually results in greater benefits for the general population as the transparency reports are made available to anyone.

**Why was transparency developed? Was it voluntary or mandated by Government?**

Transparency was developed to hold both BlackRock and the companies in which they invest accountable for their performances and behaviours. It was also developed in acknowledgement of the role that BlackRock can play as stewards for the companies they invest in and their responsibility for exposing and rectifying malpractice or below-standard behaviours.

Whilst the transparency initiative was developed by BlackRock (and therefore not mandated by a governing body), the power this has on companies seeking investment/ having gained investment is highly influential so is beyond simply voluntary.

**What is the key information / data that is reported?**

On top of transparency in their stewardship activities, BlackRock also disclose information about their sustainability and governance, social, and environmental practices through BlackRock's Annual Report<sup>70</sup>.

In the past year, BlackRock Investment Stewardship (BIS) has issued 45 bulletins to explain vote decisions and the engagement behind it. And, as of June 30, more than 80% of the approximate 5,600 active portfolios managed by BlackRock have integrated

<sup>69</sup>BlackRock: 'Delivering on our commitment to sustainability and stewardship' <https://www.blackrock.com/institutions/en-gb/solutions/sustainable-investing/our-commitment-to-stewardship>

<sup>70</sup> BlackRock: #BlackRock Sustainability Accounting Standards Board Disclosure' <https://www.blackrock.com/corporate/literature/continuous-disclosure-and-important-information/blackrock-2019-sasb-disclosure.pdf>

ESG criteria into standard investment practices such as portfolio risk reviews with BlackRock's Risk and Quantitative Analysis (RQA) group<sup>71</sup>.

For their suppliers and companies within their portfolio, BlackRock also encourages reporting programmes to be in place for all of the above information.

Just as BlackRock seek transparency from the companies in which they invest on behalf of their clients, BIS seeks to be transparent in its stewardship activities by publishing global governance and engagement guidelines, market specific voting guidelines, quarterly and annual activity reports, and a range of thought leadership pieces.

Outside of the work conducted by BlackRock, there is the Sustainability Accounting Standards Board (SASB) which acts an independent non-profit organization which sets standards to guide the disclosure of financially material sustainability information by companies to their investors. SASB Standards identify the subset of environmental, social, and governance (ESG) issues most relevant to financial performance in each of 77 industries. SASB also provides education and other resources that advance the use and understanding of its Standards <sup>72</sup>. The SASB's Standards are designed to:

**What else is required to make the model work? What is the necessary context?**

- Surface sustainability information which is financially material.
- Be industry specific because the issues that are most likely to impact financial performance vary by industry.
- Produce information that is decision useful for investors.
- Be cost effective for companies to use.
- Be developed using evidence based and market informed process that is modelled after the process used to develop financial accounting standards.

As transparency reporting between companies and their investors increases in the future, the demand for an independent body which helps regulate the content and language reported may increase to continue developing industry standards that are shared globally.

**What were the processes and steps involved in implementing reporting?**

Initially BlackRock signed up to an industry wide scheme (Principles for Responsible Investment) which initiated the transparency reporting. In order to ensure full coverage, BlackRock may engage with suppliers to track progress against their goals, promote best practices and advance diversity, equity, and inclusion within their businesses.

<sup>71</sup> BlackRock, BlackRock Earns Straight A's for Transparency by Principles for Responsible Investment, <https://www.blackrock.com/corporate/newsroom/press-releases/article/corporate-one/press-releases/blackrock-earns-straight-a-for-transparency>

<sup>72</sup> SASB (Sustainability Accounting Standards Board): 'About' <https://www.sasb.org/about/>

**What are the risks? Are there any unintended consequences?**

There is little evidence of unintended consequences. However, it has been noted that removing investment from non-complying companies does not solve the problem. Companies without the investment of one entity do not cease to exist and can still leverage investment from other sources. There therefore needs to be an industry-wide approach to ensure greatest impact<sup>73</sup>.

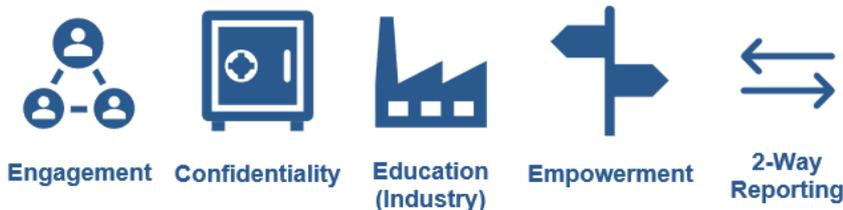
**What lessons can be learned from this model and be applied to Online Safety?**

A key lesson is the sentiment that to expect transparency, one has to be transparent themselves. BlackRock publishes self-reflecting reports across all of the issues that they expect their suppliers and companies invested in to publish. Whilst perhaps this is not directly applicable to Online Safety, it reflects the importance of a regulator being transparent in their processes and reasoning when asking for data from platforms and other businesses.

## Aviation

**Status:** Implemented

**Principal mechanisms employed:**



**Summary:** The aviation industry has established a ‘just culture’<sup>74</sup> whereby the reporting of safety concerns, incidents and near misses supersedes either the individual’s or an organisation’s concern for reputational damage or penalisation. There exists a confidential, two-way reporting structure between private entities and regulators. This mechanism often avoids public disclosure with the goal of educating industry players and achieving universal safety and compliance. This transparency ensures incidents and concerns do not go unreported and can be learnt from.

Essentially, the open and ‘just culture’ makes a safe environment for employees to report incidents. Stringent response protocols ensure reports are followed-up, promoting a sense of meaningful transparency. Therefore, transparency remains effective in preventing incidents through the engagement with, and empowerment of individuals and corporations to have an impact on safety protocols.

### Civil Aviation Authority

**General Overview**

The aviation industry is very heavily regulated throughout the entire supply chain from aircraft design & manufacture to flying operations. This regulation extends beyond the aircrew flying the aircraft, including regulations for groundcrew, air traffic control and for the manufacturers and suppliers of aviation equipment. Despite stringent regulations, a confidential, trust and safety centred approach has created a culture whereby the reporting of safety concerns, incidents and near misses supersedes either the

<sup>73</sup> Ruche, D. 2020, Green investing ‘is definitely not going to work’, says ex-BlackRock executive, <https://www.theguardian.com/business/2021/mar/30/tariq-fancy-environmentally-friendly-green-investing>

<sup>74</sup> CAA: ‘Just Culture for the Team’ <https://caainternational.com/course/just-culture-team/>

individual's or an organisation's concern for reputational damage or penalisation.

### What does transparency help with?

**Transparency mechanisms:** There exists a confidential, two-way reporting structure between private entities and regulators. This mechanism often avoids public disclosure with the goal of educating industry players and achieving universal safety and compliance.

**What transparency helps with:** Ultimately, transparency ensures incidents and concerns do not go unreported and can be learnt from. The safety of aviation operations and preservation of an open Just Culture in incident reporting across the sector, including the collaboration between competitors on safety issues and concerns promotes a pro-active and positive attitude towards reporting safety concerns.

### Why and how is transparency effective? What is the intended audience for the transparency information? Do the benefits spill out beyond that audience?

**Why transparency is effective:** The open and 'just culture' between the regulator and sector employees of reporting incidents or concerns mitigates the motivations for masking them in fear they may bring an individual (e.g., an employee) or an organisation (e.g., an airline) into disrepute with the consequence of public reputational damage or via legal or employment disciplinary action. Furthermore, stringent response protocols on investigations ensures reports are followed-up, promoting a sense of meaningful transparency whereby people value the reporting process, safe in the knowledge that reports will be sufficiently addressed and not ignored. This ensures that transparency remains effective in preventing incidents through the engagement with, and empowerment of individuals and corporations to have an impact on safety protocols.

**How transparency is effective:** Whilst 'reputational pressure' remains in the toolkit of the CAA, their regulatory approach prioritises safeguarding the confidentiality of issues reported; "Occurrence information can only be used to maintain or improve aviation safety. This means that we can't release occurrence information to the general public or to the media, including in response to Freedom of Information Act (FOIA) requests"<sup>75</sup>

This confidentiality is central to the principal of the Just Culture in aviation, based upon the theory that organisations and individuals who fear punishment may attempt to unsafely hide incidents or safety concerns - "Just Culture is a culture in which front line operators or other persons are not punished for actions, omissions, or decisions taken by them, that are commensurate with their experience and training, but in which gross negligence, wilful violations and destructive acts are not tolerated."<sup>76</sup>

**Who:** Whilst the initial **intended audience** for this Just Culture approach was for operators in the UK aviation sector, the global nature of the aviation industry means the benefits clearly spill out to other stakeholders, notably customers (passengers) internationally with increased aviation safety and for international aviation providers whom the CAA offers training and consultancy.

<sup>75</sup> CAA: 'Occurrence Reporting' <https://www.caa.co.uk/Our-work/Make-a-report-or-complaint/MOR/Occurrence-reporting/>

<sup>76</sup> CAA: 'Just Culture for the Team' <https://caainternational.com/course/just-culture-team/>

**Why was transparency developed? Was it voluntary or mandated by Government?**

The CAA were established in 1972 where the need for an open and transparent aviation safety reporting culture become readily apparent in light of the significant scale of aviation disasters, notably the 1972 Staines air disaster.

Transparency reporting follows a mixture of mandated and voluntary requirements dependant on the severity of the concern at hand, however the culture of reporting focuses on positive and pro-active voluntary reporting to identify issues prior to them being identified amidst mandatory reported data.

**What is the key information / data that is reported?**

The CAA have a range of transparency mechanisms that each report different information.

- **Occurrence reporting** – A combination of voluntary and mandated requirements, this reporting covers “any safety-related event which endangers or which, if not corrected or addressed, could endanger an aircraft, its occupants or any other person.”<sup>77</sup>
- **Innovation Hub**<sup>78</sup>- The CAA’s innovation hub is designed to foster innovation via allowing two-way transparency and insight between the CAA and emerging aviation technology companies. This is intended to promote the safe testing of aviation innovation and the opportunity to understand potential regulatory compliance challenges, and where either the CAA or the technology developers need to adapt to meet compliance demands. The hub includes an **innovation sandbox** and a **regulatory toolkit**.

**What else is required to make the model work? What is the necessary context?**

In order to manage the mixed taxonomy of data and information from native aviation provider systems, the CAA requires information disclosed to be sent in a common format, using a common taxonomy, to ensure information is compatible and can be understood and managed effectively by the CAA and ECCAIRS (European Co-ordination centre for Accident and Incident Reporting Systems) systems.

**What were the processes and steps involved in implementing reporting?**

The CAA have spent nearly 50 years putting into process the transparency reporting that currently exists. Key to the ongoing implementation is the ability to be adaptive to new challenges. It is also crucial to have multiple channels for reporting to ensure coverage of all measures.

**What are the risks? Are there any unintended consequences?**

With the CAA approach no unintended risks arose in the research. However, in interview 15 it was noted that where the introduction of limited reporting events (e.g., ‘3 strikes you’re out’) occur, this could damage the model through fears of retribution.

**What lessons can be learned from this model and be applied to Online Safety ?**

This example of operating transparency via a Just Culture approach, preserved via confidentiality, demonstrates an example whereby the safety processes of a provider can be improved via an open and collaborative approach towards safety improvement rather than disclosing safety issues or shortfalls to the public and

<sup>77</sup> CAA: ‘Occurrence Reporting’ <https://www.caa.co.uk/Our-work/Make-a-report-or-complaint/MOR/Occurrence-reporting/>

<sup>78</sup> The: ‘CAA Innovation Hub’ <https://www.caa.co.uk/Our-work/The-CAA-Innovation-Hub/>

risking subsequent reputational damage and a loss of public trust in the sector overall.

## Data Services

**Status:** Implemented

**Principal mechanisms employed:**



**Summary:** The '5 Safes framework' outlines a set of principles that provides researchers sufficient (transparent) access to sensitive data whilst preserving the security and anonymity of data accessed. The framework promotes principles in which accredited researchers can be granted access to sensitive data in a controlled manner, allowing research which can educate (and subsequently expose) issues & trends in society which would otherwise be hidden.

The initial rationale for creating this framework was in recognition of the value of sensitive data access as a major research resource alongside the simultaneous need to strictly preserve confidentiality of the data accessed and monitoring how it is subsequently used.

### Data Access for Sensitive Research - The 5 Safes framework

#### General Overview

The 5 Safes framework is a set of principles that outline the steps that owners of sensitive data should take to provide complete assurance in the protection of this data whilst permitting controlled researcher access.<sup>79</sup>

In action, the framework provides researchers sufficient access to sensitive data whilst preserving the security and anonymity of data accessed. For data considered too confidential for download, the framework has been applied to 'Secure Lab' environments, whereby researchers can be permitted access to data under supervision and where research results can be screened by third-parties to ensure that results do not infringe upon the confidentiality of the original data.<sup>80</sup> In the case of the UK Data Archive at the University of Essex, researchers can only be granted access on the basis they are 'trusted' and are conducting research that services 'public good' which mitigates risk of misuse of data disclosure.<sup>81</sup>

#### What does transparency help with?

**Transparency mechanisms:** Transparency involves the access to confidential data for researchers in controlled environments.

**What transparency helps with:** The 5 Safes framework helps accredited researchers who require sensitive data access to conduct their research. It simultaneously preserves the security

<sup>79</sup> UK Data Service: 'Five Safes' <https://www.ukdataservice.ac.uk/manage-data/legal-ethical/access-control/five-safes.aspx>

<sup>80</sup> UK Data Service: 'Access to sensitive data for research: The 5 Safes' <http://blog.ukdataservice.ac.uk/access-to-sensitive-data-for-research-the-5-safes/>

<sup>81</sup> Ibid.

and anonymity of data as not to breach or undermine expectations and protections for individual's data confidentiality.

This is helpful for understanding trends in areas such as health or economics, whereby personal data in aggregate is useful to contribute to building 'big picture' understanding that can benefit the public overall, whilst ensuring individual sensitive data is protected, remains confidential and is not identifiable to any individual in research results.

**Why and how is transparency effective?  
What is the intended audience for the transparency information?  
Do the benefits spill out beyond that audience?**

**Why is transparency effective:** The 5 Safes framework promotes principles in which accredited researchers can be granted access to sensitive data in a controlled manner, allowing research which can educate (and subsequently expose) issues & trends in society which would otherwise be hidden.

For example, in deployment of 5 Safes by the Office for National Statistics, the accessibility to large sensitive datasets (such as surveys) by accredited researchers can serve the public good in improving understanding of trends and issues in society, subsequently empowering more informed public participation and decision making.<sup>82</sup>

**How is transparency effective:** The nature of the framework enables research projects that genuinely benefit the public overall without compromising the protections necessary to preserve sensitive data. In the context of transparency, this provides a mechanism for 'light' in the 'sunlight disinfects'<sup>83</sup> analogy as a channel of accessibility to study and raise awareness of a given subject without undermining anonymity and data protections. Furthermore, the auditing of research to determine that no infringements occur reinforces the security needed to ensure its effectiveness.

**Who:** The framework is aimed at researchers and research institutions that either hold or wish to access sensitive data. Arguably, the benefits of this mechanism for researcher access spill out beyond this intended audience, as outcomes of research conducted as a result can benefit broader understanding of trends and issues in society.

**Why was transparency developed? Was it voluntary or mandated by Government?**

**Why was transparency developed?**

The 5 Safes framework is an iteration of the Virtual Microdata Laboratory model initially implemented in 2002/2003 by the Office for National Statistics to "provide secure access to confidential business survey data for research purposes"<sup>84</sup>. The rationale for creating this was in recognition of the value of accessing sensitive data as a major UK research resource alongside the simultaneous need to strictly preserve confidentiality of data accessed.<sup>85</sup>

**Was it voluntary or mandated by Government?** Voluntary.

<sup>82</sup> Office for National Statistics: 'The Five Safes- Data Privacy at ONS' <https://blog.ons.gov.uk/2017/01/27/the-five-safes-data-privacy-at-ons/>

<sup>83</sup> Brandeis, 1914: 'Chapter V- What Publicity Can Do' <http://louisville.edu/law/library/special-collections/the-louis-d.-brandeis-collection/other-peoples-money-chapter-v>

<sup>84</sup> Ritchie, 2008: Secure access to confidential microdata: four years of the Virtual Microdata Laboratory <https://link.springer.com/content/pdf/10.1057/elmr.2008.73.pdf>

<sup>85</sup> Ibid.

<p><b>What is the key information / data that is reported?</b></p>	<p>N/A</p>
<p><b>What else is required to make the model work? What is the necessary context?</b></p>	<p>From a logistical perspective, implementation and operation of a 5 Safes system requires the initial expertise and understanding of operating the framework effectively and the physical tools to manage the process, which may require technical skills for the development of secure lab environments etc.</p> <p>In the context of understanding the efficacy of accredited or privileged research access for increasing the transparency of information to the public, you may need to assess the scale of public engagement with specific research outcomes.</p>
<p><b>What were the processes and steps involved in implementing reporting?</b></p>	<p>The <b>five stages</b> of the 5 Safes framework are as follows<sup>86</sup>:</p> <ul style="list-style-type: none"> <li>• <b>Safe data</b> – Data is assessed and treated to preserve confidentiality.</li> <li>• <b>Safe projects</b> – Projects are assessed for ethicality and use for ‘public good’.</li> <li>• <b>Safe people</b> – Researchers are trained and vetted for safe and ethical data usage.</li> <li>• <b>Safe settings</b> – Secure Lab environments control data access and prevent misuse.</li> <li>• <b>Safe outputs</b> – Research results are screened and subsequently approved to ensure confidentiality and anonymity are maintained.</li> </ul>
<p><b>What are the risks? Are there any unintended consequences?</b></p>	<p>The use of accredited or privileged researcher access opens up the potential for bias in the process of choosing which researchers, institutions or research projects are granted access to sensitive data. This has the potential to undermine the fairness of the process.</p>
<p><b>What lessons can be learned from this model and be applied to Online Safety?</b></p>	<p>In response to platforms displaying a hesitancy to share data with researchers due to privacy and confidentiality concerns, the 5 Safes framework provides an example of a workable solution. Implementing an approach whereby research requiring sensitive data can be carried out in a manner that mitigates the risks unethical data misuse or privacy/confidentiality concerns could be used in the regulation of platforms. This may be particularly useful in application of understand trends in the impact of online safety for specific user groups, including marginalised or vulnerable users.</p> <p>Furthermore, secure researcher access also has the potential to operate as a measure of quality assurance for the quality of data collection and disclosure under transparency initiatives. In the case of operations within the ONS Virtual Microdata Laboratory, several studies that utilised secure data access identified inconsistencies and inaccuracies in data collection across several</p>

<sup>86</sup> UK Data Service: Secure Lab ‘Five Safes framework’ <https://www.ukdataservice.ac.uk/use-data/secure-lab/security-philosophy>

subject areas, e.g., Ormerod (2007) highlighting inconsistency in self-employment data collection.<sup>87</sup>

## Environment

**Status:** Implemented

**Principal mechanisms employed:**



**Summary:** E-PRTR is a Europe-wide database that provides easily accessible data and information on environmental pollutants.

Transparency is in the form of a publicly available, one-way reported database which contains corporate information about industrial facilities pollutant behaviours. The theory is that by increasing access to information about industrial pollutants, it will increase public capacity for participation in environmental decision making. Based upon theory set out in the Kiev Protocol, this public availability of pollutant data not only drives accountability and educates consumers, but it subsequently incentivises companies to reduce their pollution as not to be identified as the greatest polluter in their sector.

### European Pollutant Release and Transfer Register (E-PRTR)

#### General Overview

The European Pollutant Release and Transfer Register (E-PRTR) is a Europe-wide database that provides easily accessible data and information on environmental pollutants covering 30,000 industrial facilities over 65 economic activities across Europe<sup>88</sup>.

#### What does transparency help with?

**Transparency mechanisms:** A publicly available, one-way reported database which contains corporate information about their pollutant behaviours.

**What transparency helps with:** The theory behind the register is that by increasing access to information about industrial pollutants, it will increase public capacity for participation in environmental decision making. Based upon theory set out in the Kiev Protocol, this public availability of pollutant data not only drives accountability and educates consumers, but it subsequently incentivises companies to reduce their pollution as not to be identified as the greatest polluter in their sector.<sup>89</sup>

<sup>87</sup> Ormerod in Ritchie, 2008: Secure access to confidential microdata: four years of the Virtual Microdata Laboratory <https://link.springer.com/content/pdf/10.1057/elmr.2008.73.pdf>

<sup>88</sup> The European Pollutant Release and Transfer Register (E-PRTR): 'What is E-PRTR?' <https://ec.europa.eu/environment/industry/stationary/eper/legislation.htm>

<sup>89</sup> Ibid.

**Why and how is transparency effective?  
What is the intended audience for the transparency information?  
Do the benefits spill out beyond that audience?**

**Why is transparency effective:** The European Pollutant Release and Transfer Register (E-PRTR) is effective through the empowerment the public with information on pollution in their environment, allowing the public to be educated on the scale of environmental pollutants and to subsequently participate in environmental decision-making processes from an informed standpoint.

**How is transparency effective:** The utility of the E-PRTR website provides instantaneous, free of charge ease of access to pollutant disclosure data for all who seek to understand more about the scale of pollutants. The website, in addition to the format of data presented, is designed in such a manner that is easy to find and understand. As a result, the friction between a user wanting to find out pollutant information and accessing it is minimal.

**Who:**

Public access

**Why was transparency developed? Was it voluntary or mandated by Government?**

**Why was transparency developed?**

Transparency via the E-PRTR was developed to empower public participation in environmental decision-making but was adopted as a method of implementing the Kiev Protocol at an EU level.<sup>90</sup> The Kiev protocol is the only international legally binding initiative on pollutant and transfer registers and seeks “to enhance public access to information through the establishment of coherent, nationwide pollutant release and transfer registers (PRTRs)”<sup>91</sup>.

Despite the E-PRTR not being operational until 2007, the principles it was founded upon under the Kiev Protocol date back to 1998, with the United Nations Aarhus Convention. The Aarhus Convention sets out three core pillars relating to the right for public participation in environmental decision-making; Firstly, “access to environmental information; Secondly “public participation in environmental decision-making”; And thirdly “access to justice” – relating to the judicial accountability of environmental decisions<sup>92</sup>.

**Was it voluntary or mandated by Government?**

Reporting pollutants under E-PRTR regulation is compulsory for operations above specified capacity thresholds. Article 5 of E-PRTR Regulation stipulates organisations who undertake one or more of the designated activities above the specified capacity thresholds are obliged to report<sup>93</sup>.

“The operator of each facility that undertakes one or more of the activities specified in Annex I above the applicable capacity thresholds specified therein shall communicate to its competent authority the information identifying the facility in accordance with

<sup>90</sup> Ibid.

<sup>91</sup> UNECE: ‘Introduction to the Kyiv Protocol on Pollutant Release and Transfer Registers

<https://unece.org/env/pp/former%20bodies-working-group-prtrs-introduction>

<sup>92</sup> European Commission: ‘The Aarhus Convention’ <https://ec.europa.eu/environment/aarhus/>

<sup>93</sup> European Commission: ‘Guidance Document for the implementation of the European PRTR’ [https://ec.europa.eu/environment/industry/stationary/e-prtr/pdf/en\\_prtr.pdf](https://ec.europa.eu/environment/industry/stationary/e-prtr/pdf/en_prtr.pdf)

Annex III unless that information is already available to the competent authority.”<sup>94</sup>

There is also the option to report information voluntarily, including data on production volume, number of installations, number of operating hours/year, number of employees and website or an information box<sup>95</sup>.

### What is the key information / data that is reported?

The database covers a range of a data disclosed by more than 30,000 industrial facilities covering 65 economic activities within energy, production and processing of metals, mineral industry, chemical industry, waste and wastewater management, paper and wood production and processing, intensive livestock production and aquaculture, animal and vegetable products from the food and beverage sector and ‘other activities’<sup>96</sup>. Of these activities, data for 91 pollutants covering greenhouse gases, other gases, heavy metals, pesticides, chlorinated organic substances, other organics substances and inorganic substances is provided<sup>97</sup>.

### What else is required to make the model work? What is the necessary context?

The E-PRTR model is based upon the principle of empowering the public to participate in environmental decision-making via access to information. For this to work, the public need to be willing to seek the information and the information disclosed needs to be understandable. In acknowledgement of this fact, the European Commission have undertaken research to assess the efficacy of the E-PRTR against its intended purpose and found that it is in fact fit for purpose<sup>98</sup>. That said, within the aforementioned research, the understandability of data reported to the public was noted as an area for improvement; “The E-PRTR was judged to be a fairly effective tool for a range of informed stakeholders, and, in combination with other instruments, contributes to the achievement of wider environmental objectives. However, data interpretation could be further supported. For the general public, additional background information would be needed to better understand and use E-PRTR data. Additional context would be beneficial if E-PRTR data is to be more used for benchmarking the environmental performance of industrial activities (i.e., specific activity data, production capacity)”<sup>99</sup>.

Information disclosed is held on the E-PRTR website<sup>100</sup> which acts as the key facilitator and enabling factor for public accessibility to pollutant data. Therefore, from a logistical standpoint, the maintenance of this database and the subsequent website is fundamental to the model’s success.

<sup>94</sup> The European Parliament and Council of the European Union: ‘Regulation (EC) No 166/2006 of the European Parliament and of the Council’ <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:033:0001:0017:EN:PDF#page=8>

<sup>95</sup> European Environment Agency: ‘European Pollutant Release and Transfer Register- a harmonised PRTR across Europe’ [https://unece.org/fileadmin/DAM/env/pp/PRTR%20Bureau/GRT2013-Item4-1-EU-EEA\\_E-PRTR.pdf](https://unece.org/fileadmin/DAM/env/pp/PRTR%20Bureau/GRT2013-Item4-1-EU-EEA_E-PRTR.pdf)

<sup>96</sup> E-PRTR: ‘What does the E-PRTR cover?’ <https://prtr.eea.europa.eu/#/static?cont=about>

<sup>97</sup> Ibid.

<sup>98</sup> European Commission: ‘REFIT evaluation of Regulation (EC) No 166/2006 concerning the establishment of a European Pollutants Release and Transfer Register (E-PRTR)’ <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52017SC0710&from=EN>

<sup>99</sup> Ibid.

<sup>100</sup> E-PRTR: <https://prtr.eea.europa.eu/#/home>

**What were the processes and steps involved in implementing reporting?**

As the E-PRTR is a cross-border initiative spanning Europe, all member states were initially required to adopt national legislation and establish procedures to meet the requirements of the E-PRTR Regulation.<sup>101</sup> At an EU level, physical steps for implementation included:

- The design and development of the E-PRTR website,
- Consultation with member states
- The design and dissemination of reporting requirements (including data quality) and quality assurance mechanisms across member states.<sup>102</sup>

The European Commission also produced a guidance document one year ahead of reporting to educate and inform stakeholders on the purpose of the E-PRTR and information on reporting requirements and how to physically go about reporting - 'Guidance Document for the implementation of the European PRTR'<sup>103</sup>.

The process of the E-PRTR follows pollutant data being reported to designated "competent authorities" within member states. These authorities report to the European Environment Agency who manage the E-PRTR website. Subsequently, within member states, a further implementation task consisted of identifying relevant and competent authorities raised issues with some member states lacking the resources to implement effective reporting.<sup>104</sup>

**What are the risks? Are there any unintended consequences?**

None observed.

**What lessons can be learned from this model and be applied to Online Safety ?**

The E-PRTR demonstrates an example whereby a large range of transparency information spanning multiple countries is successfully fed into a singular database that is easily accessible by the public. The use of a central website or database that is publicly accessible could provide a way for consumers to quickly compare and contrast the safety processes of different platforms. Additionally, the principle enforced via the Kiev protocol (and the Aarhus Convention) that establishes a public right to information that facilitates considered, informed decisions about the environment could potentially be replicated for the digital environment.

<sup>101</sup> European Commission 'EFIT evaluation of Regulation (EC) No 166/2006 concerning the establishment of a European Pollutant Release and Transfer Register (E-PRTR)' <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52017SC0710&from=en>

<sup>102</sup> European Commission 'Guidance Document for the implementation of the European PRTR' [https://ec.europa.eu/environment/industry/stationary/e-prtr/pdf/en\\_prtr.pdf](https://ec.europa.eu/environment/industry/stationary/e-prtr/pdf/en_prtr.pdf)

<sup>103</sup> European Commission 'Guidance Document for the implementation of the European PRTR' [https://ec.europa.eu/environment/industry/stationary/e-prtr/pdf/en\\_prtr.pdf](https://ec.europa.eu/environment/industry/stationary/e-prtr/pdf/en_prtr.pdf)

<sup>104</sup> European Environment Agency: 'European Pollutant Release and Transfer Register- a harmonised PRTR across Europe' [https://unece.org/fileadmin/DAM/env/pp/PRTR%20Bureau/GRT2013-Item4-1-EU-EEA\\_E-PRTR.pdf](https://unece.org/fileadmin/DAM/env/pp/PRTR%20Bureau/GRT2013-Item4-1-EU-EEA_E-PRTR.pdf)

## Fashion, Textiles & Clothing

**Status:** Implemented

**Principal mechanisms employed:**



**Summary:** Concerns over ethics in apparel production, notably regarding fair labour practices and environmental sustainability, have seen the demand for transparency of fashion supply chains develop significant momentum.

Manufacturers provide transparency voluntarily via one-way occurrence reporting and also performance information where they might be falling short of expected standards. Transparency enables others to scrutinise what companies say they are doing to address human rights and protect the environment.

Transparency is effective in the context of fashion supply chains as it educates and empowers consumers to take action against businesses who do not participate good, ethical practice. There is also a strong role for accountability as the publishing of performance information allows both consumers and legislators to hold businesses to the standards that they claim to keep.

Ethical Supply Chains in fashion	
<b>General Overview</b>	<p>Concerns over ethics in apparel production, notably regarding fair labour practices and environmental sustainability, have seen the demand for transparency of fashion supply chains develop significant momentum. Coming to light largely in response to whistleblowing and private investigations amplified by the media, the theory of transparency in fashion supply chains relies on the idea of consumer empowerment and encouraging consumers to change their behaviours and adjust their consumption habits to make ethical purchases in line with the information provided. It is hoped that this pressure, and the additional accountability placed on manufacturers will drive them to improve poor practice to prevent reputational damage.</p> <p>Whilst it cannot be guaranteed that this behavioural change (from both consumers and manufactures/retailers) is effective in all applications, there is research to suggest that transparency positively effect consumers consumption choices<sup>105</sup>.</p>
<b>What does transparency help with?</b>	<p><b>Transparency mechanisms:</b> Transparency is occurring from both the one-way occurrence reporting by manufacturers and also performance information where they might be falling short of expected standards. It's worth noting that performance information sometimes comes from third parties, such as the media, through investigative journalism.</p> <p><b>What transparency helps with:</b> "Transparency enables others to scrutinise what companies say they are doing to address</p>

<sup>105</sup> Kim et al. (2020): 'Is Honesty the Best Policy? Examining the Role of Price and Production Transparency in Fashion Marketing' [https://www.mdpi.com/2071-1050/12/17/6800?type=check\\_update&version=1](https://www.mdpi.com/2071-1050/12/17/6800?type=check_update&version=1)

human rights and protect the environment. It means that there is information available for which others (consumers, investors, lawmakers, journalists, NGOs, trade unions, workers themselves) can hold brands and retailers to account for their policies and practices, especially when things go wrong like it did that day at Rana Plaza.”<sup>106</sup>

**Why is transparency effective:** Transparency is effective in the context of fashion supply chains as it educates and empowers consumers to take action against businesses who do not participate good, ethical practice. There is also a strong role for accountability as the publishing of performance information allows both consumers and legislators to hold businesses to the standards that they claim to keep.

Increased supply chain transparency from leading fashion brands has the potential to ‘raise the bar’ of industry standards by encouraging competitors to also increase their transparency standards to match. This was the case in the response to Nike’s increased efforts towards transparent operations with subsequent transparency activity stemming from key competitors as a result.<sup>107</sup>

**Why and how is transparency effective?  
What is the intended audience for the transparency information?  
Do the benefits spill out beyond that audience?**

**How is transparency effective:** In the instance of fashion, transparency is effective due to the scale of consequence that can occur if consumers take action against a business. The intended outcome of reputational pressure can drive the manufacturer/retailer to resolve issues of poor practice to reassure both consumers and investors.

### **Case Study: Boohoo Group PLC**

In July 2020, the Sunday Times published an article<sup>108</sup> outlining the outcome of an undercover investigation into fashion group Boohoo’s unethical manufacturing practices in Leicester, including accusations of slavery for workers being paid far below minimum wage. Subsequently, in the three days following the article, the value of the company fell by almost £2bn<sup>109</sup>. In response, the group commissioned an independent review into their Leicester supply chain, publishing a publicly available report of the findings<sup>110</sup>. Having accepted the findings of this review, Boohoo have since cut ties with a number of its manufacturers who are unable to demonstrate sufficient transparency information.<sup>111</sup> As a result of this decision to increase transparency within their supply chain, the groups share price significantly rebounded by more than 27%<sup>112</sup>.

<sup>106</sup> Fashion Revolution: ‘The Fashion Transparency Index 2020: What do we mean by transparency?’

<https://www.fashionrevolution.org/about/transparency/>

<sup>107</sup> Network for Business Sustainability: How Nike Turned a Supply Chain Crisis into Opportunity

<https://www.nbs.net/articles/just-do-it-how-nike-turned-a-supply-chain-crisis-into-opportunity>

<sup>108</sup> The Times: ‘Boohoo: fashion giant faces ‘slavery’ investigation’ <https://www.thetimes.co.uk/article/boohoo-fashion-giant-faces-slavery-investigation-57s3hxcth>

<sup>109</sup> The Guardian: ‘Boohoo shares bounce back after pledge to improve factory conditions’

<https://www.theguardian.com/fashion/2020/jul/09/boohoo-shares-bounce-back-after-pledge-to-improve-factory-conditions>

<sup>110</sup> Independent Review into the boohoo Groups PLC’S Leicester supply chain <https://www.boohooplc.com/sites/boohoo-corp/files/final-report-open-version-24.9.2020.pdf>

<sup>111</sup> Reuters: ‘Boohoo cuts number of UK suppliers’ <https://www.reuters.com/article/us-boohoo-suppliers-idUSKBN2BH0R4>

<sup>112</sup> The Guardian: ‘Boohoo shares bounce back after pledge to improve factory conditions’

<https://www.theguardian.com/fashion/2020/jul/09/boohoo-shares-bounce-back-after-pledge-to-improve-factory-conditions>

**Limitations of this approach:** The limitation of transparency in fashion supply chains is that, as observed in other sectors, consumers may not always access and act upon transparency information available to them, e.g., within the context of fashion, seeking out ethical fashion reports and choosing to shop ethically as a result. Without consumer and investor pressure, transparency initiatives may fail to put significant pressure on manufacturers/retailers to adjust their practices.

**Who:** Transparency in fashion supply chains tends to be oriented towards consumers and investors, with whistleblowing efforts aimed at the manufacturer/retailers themselves to generate reputational pressure from the two former groups.

### Why was transparency developed? Was it voluntary or mandated by Government?

#### Why was transparency developed?

Momentum for transparency in fashion supply chains has grown in symbiosis with greater demand for ethical and sustainable supply chains across many sectors. In terms of initially observation harm within supply chains, whistleblowing investigations amplified by the media and social media campaigns, such as in the case of the accusations facing the Boohoo Group PLC in 2020<sup>113</sup> and outcry in response to specific events, notably the 2013 Rana Plaza garment factory collapse in Bangladesh<sup>114</sup>, fuel momentum for greater transparency requirements with the intention of pressuring better industry practice.

#### Voluntary or mandated by Government?

Voluntary

### What is the key information / data that is reported?

Whilst there is an abundance of activism relating to supply chain transparency, the Fashion Transparency Index<sup>115</sup> produced by NGO Fashion Revolution offers a coherent, international approach to identifying the transparency reporting processes of the world's leading fashion brands.

The index rates brands and retailers across five key areas, receiving points for information that is publicly disclosed (either self-published or via third parties)<sup>116</sup>:

- Social and environmental policy and commitments
- Governance
- Supply chain traceability
- Know, show and fix (supply chain due diligence and remediation)
- Spotlight issues (working conditions, consumption, product/material composition and climate).

Another example of transparency reporting activity in fashion stems from the Environmental Audit Committee, including the

<sup>113</sup> The Guardian: 'Boohoo shares bounce back after pledge to improve factory conditions'

<https://www.theguardian.com/fashion/2020/jul/09/boohoo-shares-bounce-back-after-pledge-to-improve-factory-conditions>

<sup>114</sup> Fashion Revolution 'Why Transparency Matters' <https://www.fashionrevolution.org/about/transparency/>

<sup>115</sup> The Fashion Transparency Index 2020 <https://www.fashionrevolution.org/about/transparency/>

<sup>116</sup> Ibid. Note: See page 13 for detailed descriptors of scoring categories.

‘Fixing Fashion: clothing consumption and sustainability report’<sup>117</sup> and the publication of fashion brands and retailer’s current sustainability actions and initiatives in addition to labour initiatives<sup>118</sup>.

It should also be noted that whistleblowing has played a major role in driving businesses to adopt more transparent approaches to their supply chains (e.g., Boohoo Group PLC, the information disclosed via the media related specifically to malpractice)<sup>119</sup>.

### What else is required to make the model work? What is the necessary context?

In the case of fashion supply chains, reputational pressure to incentivise better industry practice is unable to occur without consumers and investors first being exposed to, and made aware of, the harm. Within this context, observation and awareness of harm is dependent upon initial access to the internal workings of a manufacturer/retailer (e.g., factory access) to be able to assess harm and then having the means to disclose notification of the harm to consumers and investors,) i.e., amplifying awareness of the harm via the media and across social media channels to generate reputational pressure).

A further factor for the model to work appears to be consistency in driving awareness (i.e., supporting momentum for change to drive reputational pressure on a frequent basis). An example is Fashion Revolution’s #WHOMADEMYCLOTHES movement that was born in response to the 2013 Rana Plaza garment factory collapse<sup>120</sup>. The movement restores momentum for pressure on greater supply chain transparency on an annual basis as part of the annual ‘Fashion Revolution Week’<sup>121</sup>, annually returning focus to the supply chain transparency conversation via the hashtag that consistently trends on social media during this period.

### What were the processes and steps involved in implementing reporting?

Whilst there are mixed methods of reporting within the context of fashion supply chains, the generic trend of transparency in fashion follows the model whereby harm is observed and amplified by the media, with public exposure and engagement with the harm increasing the desire to hold businesses accountable. This awareness drives pressure for greater transparency, to which materialising via reports such as the Fashion Transparency Index<sup>122</sup> enables scrutiny and accountability of malpractice. Subsequently, this accountability in turn incentivises corrective action by a retailer/manufacturer whom, facing consumer and investor scrutiny, seek to improve their practices.

<sup>117</sup> Environmental Audit Committee: Fixing Fashion: clothing consumption and sustainability <https://publications.parliament.uk/pa/cm201719/cmselect/cmenvaud/1952/full-report.html>

<sup>118</sup> Interim Report on the Sustainability of the Fashion Industry [https://publications.parliament.uk/pa/cm201719/cmselect/cmenvaud/1148/114804.htm#\\_idTextAnchor004](https://publications.parliament.uk/pa/cm201719/cmselect/cmenvaud/1148/114804.htm#_idTextAnchor004)

<sup>119</sup> The Times: ‘Boohoo: fashion giant faces ‘slavery’ investigation’ <https://www.thetimes.co.uk/article/boohoo-fashion-giant-faces-slavery-investigation-57s3hxcth>

<sup>120</sup> Vogue ‘Why We Should Be Asking #WhoMadeMyClothes? Before Every Purchase’ <https://www.vogue.co.uk/article/who-made-my-clothes>

<sup>121</sup> Fashion Revolution: ‘Fashion Revolution Week’ <https://www.fashionrevolution.org/>

<sup>122</sup> Fashion Revolution: ‘The Fashion Transparency Index 2020’ <https://www.fashionrevolution.org/about/transparency/>

**What are the risks? Are there any unintended consequences?**

There is a potential risk that the use of whistleblowing and reputational pressure may not have the intended impact of causing companies to improve their processes but instead it could drive friction between the reporting entity (e.g., an NGO or regulator) and the company, which could hinder collaboration to improve practices moving forward, although this theory within fashion is not evidenced.

**What lessons can be learned from this model and be applied to Online Safety ?**

In application to the context of online safety, there may be an opportunity to seek a supply chain approach for content moderation. This could involve seeking transparency in the end-to-end content moderation supply chain, from policy design, human and automated intervention (including transparency in the human moderation process e.g., use of third-party moderators, welfare of moderators), takedowns and appeals processes.

The approach of an annual transparency report from Fashion Revolution based upon scoring brands and retailers on their processes additionally offers a strong example of reporting systemic approaches to transparency reporting across a wide range of providers.

## Food & Drink

**Status:** Implemented

**Principal mechanisms employed:**



**Summary:** Food and drink manufacturers provide summary nutritional information on product packaging to support consumer choice and stimulate healthier consumption habits, sometimes referred to as ‘traffic lights’.

The display of nutritional information on the front of packaging is a one-way reporting transparency mechanism for relaying information to consumers. This method of transparency is designed to provide consumers with an awareness of the nutritional value of products to inform healthier diet choices.

Research on the effectiveness of nutritional labelling in its end-goal of empowering agency for the intended audience (consumers) towards healthier consumption habits is inconclusive.

### Front-of-package (FoP) nutritional labelling

**General Overview**

The nutritional information displayed on the front of food and drinks goods, often displayed as ‘traffic lights’ are designed to inform consumers on the nutritional value of products and subsequently to incentivise healthier consumption habits.

## What does transparency help with?

The display of nutritional information on the front of packaging is a one-way reporting transparency mechanism for relaying information to consumers. However, transparency and accountability mechanisms are also incorporated into regulatory regimes when developing these FOP displays<sup>123</sup>.

This method of transparency is designed to provide consumers with an awareness of the nutritional value of products to inform healthier diet choices.

## Why and how is transparency effective? What is the intended audience for the transparency information? Do the benefits spill out beyond that audience?

Research on the effectiveness of nutritional labelling in its end-goal of empowering agency for the intended audience (consumers) towards healthier consumption habits is inconclusive.

Findings from a ComRes Diabetes UK Survey found that 87% of respondents agree that the traffic light system helps people make informed choices about food consumption<sup>124</sup> and in discussion of the Nutri-Score system, Julia & Hercberg (2017) argue for the effectiveness of this type of labelling as an effective tool in public health nutrition.<sup>125</sup> On the other hand, alternative studies such as Sacks et al. (2009) concluded there to be no discernible effect of nutritional signposting on healthier consumption habits<sup>126</sup>.

The simplified nature of the 'traffic light' system is meant to quickly provide consumers with the information that the need to inform their choices. However, contention exists over the balance of simplicity that food labelling provides. Whilst Villas-Boas (2010) suggest that food labelling is more effective when simplistic<sup>127</sup>, the findings of Ikonen et al. (2019) suggest that overly simplified labels can be misused and interpreted to demonstrate false sense of nutritional value, limiting the effectiveness of food labelling overall<sup>128</sup>.

**Who:** Ultimately, FOP nutritional labelling is designed to help consumers make more informed decisions about their food purchases and improve their health.

## Why was transparency developed? Was it voluntary or mandated by Government?

### Why was transparency developed?

"Research and modelling have also shown that even small changes to the diet can have significant impacts on individuals' health, and in reducing the costs of ill-health to the economy. FoP nutrition information was implemented to help support consumers in making healthier choices and realising those benefits"<sup>129</sup> and to ultimately improve the health of a nation.

<sup>123</sup> Jones et al. (2019), 'Front-of-pack nutrition labelling to promote healthier diets: current practice and opportunities to strengthen regulation worldwide' <https://gh.bmj.com/content/bmjgh/4/6/e001882.full.pdf>

<sup>124</sup> Diabetes UK 'Public Views on Food Labelling' – January 2018 [https://2sijwunngl41ia7ki31qqub1-wpengine.netdna-ssl.com/wp-content/uploads/2018/03/J303701\\_Diabetes-UK\\_Food-Labeling\\_Data-Tables\\_January-2018.pdf](https://2sijwunngl41ia7ki31qqub1-wpengine.netdna-ssl.com/wp-content/uploads/2018/03/J303701_Diabetes-UK_Food-Labeling_Data-Tables_January-2018.pdf)

<sup>125</sup> Julia & Hercberg (2017) 'Nutri-Score: Evidence of the effectiveness of the French front-of-pack nutrition label' [https://www.ernaehrungs-umschau.de/fileadmin/Ernaehrungs-Umschau/pdfs/pdf\\_2017/12\\_17/EU12\\_2017\\_WuF\\_Nutriscore\\_englisch.pdf](https://www.ernaehrungs-umschau.de/fileadmin/Ernaehrungs-Umschau/pdfs/pdf_2017/12_17/EU12_2017_WuF_Nutriscore_englisch.pdf)

<sup>126</sup> Hersey et al. (2009) 'Effects of front-of-package and shelf nutrition labelling systems on consumers' <https://pubmed.ncbi.nlm.nih.gov/23282247/>

<sup>127</sup> Villas-Boas (2010) 'Can information cost affect consumer choice? Nutritional labels in a supermarket experiment' [https://www.csus.edu/indiv/k/kieselk/Nutritional%20label%20experiment\\_IJIO.pdf](https://www.csus.edu/indiv/k/kieselk/Nutritional%20label%20experiment_IJIO.pdf)

<sup>128</sup> Ikonen et al. (2019) 'Consumer effects of front-of-package nutrition labelling: an interdisciplinary meta-analysis' <https://link.springer.com/article/10.1007/s11747-019-00663-9>

<sup>129</sup> Food Standards Agency: 'Guide to creating a front of pack (FoP) nutrition label for pre-packed products sold through retail outlets' [https://www.food.gov.uk/sites/default/files/media/document/fop-guidance\\_0.pdf](https://www.food.gov.uk/sites/default/files/media/document/fop-guidance_0.pdf)

## Voluntary or mandated?

FoP labelling is currently voluntary in the UK and has been widely adopted. However, there are some infamous case studies where consumer pressure has caused brands to uptake despite initial reluctances (see Kellogg's backtracking on the decisions to not use traffic lights<sup>130</sup>)

## What is the key information / data that is reported?

A Front of Pack label developed in accordance with the food.gov guidance will contain<sup>131</sup>:

- Information on the energy value in kilojoules (kJ) and kilocalories (kcal) per 100g/ml and in a specified portion of the product.
- Information on the amounts in grams of fat, saturated fat ("saturates"), (total) sugars and salt in grams, in a specified portion of the product.
- Portion size information expressed in a way that is easily recognisable by, and meaningful to the consumer. For example, ¼ of a pie or 1 burger.
- % RI information based on the amount of each nutrient and energy value in a portion of the food.
- Colour coding of the nutrient content of the food

## What else is required to make the model work? What is the necessary context?

The introduction of FoP nutritional information came alongside the production of the Eatwell Guide and the education of consumers about their nutrition. This two-sided approach of increasing consumer engagement with their nutrition and the importance of this was a major driver in improving the impact on consumer health. "Public Health England encourages organisations and individuals to use the Eatwell Guide to make sure everyone receives consistent messages about the balance of foods in a healthy diet"<sup>132</sup>.

Using a consistent and coordinated set of messaging will ensure that the information consumers receive is amplified and consolidated and will help prevent confusion.

## What were the processes and steps involved in implementing reporting?

From 14 May to 6 August 2012 the UK Government ran a joint consultation on Front of Pack (FoP) nutrition labelling to explore how greater consistency and clarity on FoP labelling might be achieved<sup>133</sup>. This consultation ensured that a range of views were expressed on their preferred formats to ensure that the developed transparency was as impactful to consumers as possible.

Furthermore, in 2016, the Department of Health and Social Care took action to build on the evidence base, and commissioned

<sup>130</sup> BBC: 'Kellogg's gives in on government's 'traffic light' labels' <https://www.bbc.co.uk/news/business-46373342>

<sup>131</sup> Food Standards Agency, Guide to creating a front of pack (FoP) nutrition label for pre-packed products sold through retail outlets [https://www.food.gov.uk/sites/default/files/media/document/fop-guidance\\_0.pdf](https://www.food.gov.uk/sites/default/files/media/document/fop-guidance_0.pdf)

<sup>132</sup> GOV.UK: 'The Eatwell Guide', <https://www.gov.uk/government/publications/the-eatwell-guide>

<sup>133</sup> Food Standards Agency, 'Front of pack Nutrition Labelling: Joint Response to Consultation' [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/216997/response-nutrition-labelling-consultation.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/216997/response-nutrition-labelling-consultation.pdf)

Kantar Worldpanel to carry out research exploring consumers' understanding and use of FOPNL in the UK <sup>134</sup>.

This approach of engaging with a broad range of players at the start of the process as well as continually learning and adapting throughout were crucial to the implementation of impactful and beneficial reporting.

### **What are the risks? Are there any unintended consequences?**

“Policymakers must decide much more than whether to apply ‘stars’, ‘traffic lights’ or ‘stop signs’. The substance of effective regulation must contain strategic regulatory objectives, clear specifications for displaying the label on pack, a valid scoring mechanism and a justified scope for including foods” <sup>135</sup>. As such, there have been concerns amongst certain food producers that this form of labelling, if not made flexible, will have negative consequences on the sales of their food (e.g., Olive Oil). Italy claims that traffic-light labelling could unfairly target Mediterranean food, which is known to be one of the healthiest diets in the world and could steer consumers away from making positive health choices <sup>136</sup>.

### **What lessons can be learned from this model and be applied to Online Safety?**

Two clear lessons can be learned from the FoP transparency approach to be applied to Online Safety:

- Having a two-sided approach to educating consumers on the benefits of health whilst improving the mechanisms for them to control this was particularly effective. For Online Safety, improving consumer knowledge of how to be safe online could therefore be combined with transparency reports on companies and the risks that platforms pose – this will allow them to make informed decisions on their online habits.
- Uniform taxonomy and imagery are crucial. Ensuring that the language used to communicate with consumers is uniform across category (e.g., confectionary, ready meals, drinks etc.) was crucial to ensuring their understanding and use. Online Safety transparency reports should take this onboard and use common language cross-platforms to ensure ease for consumers.

<sup>134</sup> Food Standards Agency: ‘Building on the success of front-of pack nutrition labelling in the UK: a public consultation’

[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/905096/front-of-pack-labelling-consultation-document-english.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/905096/front-of-pack-labelling-consultation-document-english.pdf)

<sup>135</sup> Jones et al. (2019), ‘Front-of-pack nutrition labelling to promote healthier diets: current practice and opportunities to strengthen regulation worldwide’ <https://gh.bmj.com/content/bmjgh/4/6/e001882.full.pdf>

<sup>136</sup> The Guardian: ‘Italy claims ‘traffic-light’ labelling unfair on Mediterranean food’ <https://www.theguardian.com/world/2013/oct/21/italy-traffic-light-food-labels-unfair>

## Media, Entertainment & Film

**Status:** Implemented

**Principal mechanisms employed:**



**Summary:** The British Board of Film Classification (BBFC), seeks to help children and families choose what content to consume by providing them with the guidance on the nature/suitability of content.

Transparency is in the form of clear reporting of content (ratings and ‘certification’) within media formats to educate consumers and empower them to make conscious decisions with their viewing habits.

Transparency in content certification is well established and trusted due to the BBFC’s reputation and the long-term relationship between the BBFC and the content producers/distributors. The one-way output of a single measure is especially effective at clearly engaging with consumers and helping inform their behaviour and choice.

British Board of Film Classification (BBFC)	
<b>General Overview</b>	The British Board of Film Classification (BBFC), seeks to help children and families choose what content to consume by providing them with the guidance they need to help them choose what’s right for them and avoid what’s not <sup>137</sup> .
<b>What does transparency help with?</b>	<p><b>Transparency mechanisms:</b> The clear reporting of content within media formats to educate consumers and empower them to make conscious decisions with their viewing habits.</p> <p><b>What transparency helps with:</b></p> <p>Their mission is to<sup>138</sup>:</p> <ul style="list-style-type: none"><li>• Protect the public, and especially children, from content which might raise harm risks</li><li>• Empower the public, especially parents, to make informed viewing choices</li><li>• Recognise and respect adult freedom of choice within the law</li><li>• Respond to and reflect changing social attitudes towards media content through proactive public consultation and research</li><li>• Provide a cost-effective, efficient classification service within our statutory remit</li><li>• Work in partnership with the industry to develop innovative service models to provide content advice which support emerging media delivery systems</li></ul>

<sup>137</sup> BBFC: ‘About us’ <https://www.bbfc.co.uk/about-us>

<sup>138</sup> BBFC: ‘Our Mission’ <https://www.bbfc.co.uk/about-us/who-we-are/our-mission>

- Provide an effective service to enforcement agencies

**Why and how is transparency effective? What is the intended audience for the transparency information? Do the benefits spill out beyond that audience?**

**Why transparency is effective:** Transparency is crucial in keeping consumers safe and ensuring that audiences can choose content that is appropriate for them. By being well established and trusted, the one-way output of a single measure is especially effective at clearly engaging with consumers and helping inform their behaviour (Interview 16).

**How transparency is effective:** Transparency in this instance is effective due to the long-standing relationship that the BBFC has with consumers and the level of trust that has been built up over this time. There is also an extensive consultation and research process that sits behind their gradings that ensures they remain in line with consumer expectations.

**Why was transparency developed? Was it voluntary or mandated by Government?**

**Why was transparency developed?**

In the 1980s, when Parliament passed the Video Recordings Act 1984 (VRA). This made it law that subject to certain exemptions, video recordings offered for sale or hire in the UK must be classified by an authority appointed by the Secretary of State. This is when classifying films and videos plays a far greater role in the BBFCs work than censorship, and in 1985 the President and Vice Presidents of the BBFC were appointed to apply the new test for video of 'suitability for viewing in the home'<sup>139</sup>.

**Is it voluntary or mandated?**

Mandated

**What is the key information / data that is reported?**

The age certificate and content included.

**What else is required to make the model work? What is the necessary context?**

There is a large amount of collaboration and research which is conducted behind the scenes to ensure validity in the eyes of consumers. In order to keep their guiding standards up to date with shifting consumer perceptions the BBFC carries out primary research with consumers, using this knowledge to understand what matters to families and they are one of the few regulatory bodies to do so.

**What were the processes and steps involved in implementing reporting?**

The BBFC takes a number of steps to create their ratings:

- Use 'general considerations' – context, tone & impact, theme to assess classification of content.

<sup>139</sup> BBFC: 'BBFC History' <https://www.bbfc.co.uk/education/university-students/bbfc-history>

- Conducting extensive consumer research (focus groups, large scale online surveys, interviews) to understand tolerance and acceptance of classification guidelines and assess tolerance of different features in films to classify ratings (e.g. tolerance of nudity, violence).
- Surveys into family viewing decisions<sup>140</sup>
- Developing a deep understanding of consumer tolerance.

**What are the risks? Are there any unintended consequences?**

There is a risk/unintended consequence that age-ratings (e.g., 18+ in gaming) inadvertently make the content more desirable to younger audiences through the promise of ‘adults only’ material e.g. bad language, nudity, and violence.

**What lessons can be learned from this model and be applied to Online Safety?**

An interesting idea raised by this case study is the role in which consumers are consulted when creating guidelines. The BBFC understands primary research to be a crucial aspect of their guideline creation however very few other transparency mechanisms take this into account. Furthermore, simple outputs that are easy to understand and very visual help consumers to quickly understand risk with little-to-no prior knowledge of the process.

<sup>140</sup> BBFC: ‘Survey into family viewing decisions September 2020’  
<https://darkroom.bbfc.co.uk/original/a9eb359c3c6c06f60371cba1b1e28587:adb5b3d7016e0e66093f1c7544572b02/survey-into-family-viewing-decisions-sept-2020.pdf>

## Financial Services

**Status:** Implemented

**Principal mechanisms employed:**



**Summary:** The FCA uses a range of transparency mechanisms including: Self two-way transparency, market transparency, and performance and behavioural information.

Transparency is designed to hold firms to account in in their products and operations, whilst also being able to empower consumers into making educated choices. The FCA see transparency as two-way interaction and use their own transparency to ensure the efficacy of their practices and provide clarity on the actions that they take.

The FCA noted that whilst the immediate benefits of transparency may not be clear, ensuring there is information available is critical to the functioning of our economy and preventing rogue actors.

Transparency requires three pathways in order to be effective: transparency of the FCA themselves; information the FCA could release about firms, individuals, markets to inform consumers; and, information the FCA could require firms to release about their products and about other aspects their performance and behaviour.

Financial Conduct Authority (FCA)	
<b>General Overview</b>	The Financial Conduct Authority alone regulates 58,000 firms <sup>141</sup> as just one piece of the puzzle of the estimated 300 million pages of financial regulatory reporting that is published annually worldwide <sup>142</sup> .
<b>What does transparency help with?</b>	<p><b>Transparency mechanisms:</b> The FCA uses a range of transparency mechanisms including: Self two-way transparency, Market Transparency, and Performance and Behavioural information (more details about these below).</p> <p><b>What transparency helps with:</b> Transparency is designed to hold firms to account in in their products and operations, whilst also being able to empower consumers into making educated choices about where to take their custom. The FCA themselves also see transparency as two-way interaction and so use their own transparency to ensure the efficacy of their practices and provide clarity on the actions that they take.</p>

<sup>141</sup> PA Consulting: 'Digital Regulatory Reporting' <https://www2.paconsulting.com/rs/526-HZE-833/images/DRR-Report-Sept-2020.pdf>

<sup>142</sup> Ibid.

**Why transparency is effective:** With more transparency comes greater agency and empowerment of consumers, and a heightened ability for businesses to be held accountable. The FCA in Interview 9 noted that whilst the immediate benefits of transparency may not be clear, ensuring there is information available is critical to the functioning of our economy and preventing rogue actors.

**How transparency is effective:**

Transparency requires three pathways in order to be effective. It's important that the FCA can ensure transparency throughout the whole system to avoid a weak link<sup>143</sup>:

**Why and how is transparency effective?  
What is the intended audience for the transparency information?  
Do the benefits spill out beyond that audience?**

- Transparency of the FCA themselves: This, through performance and two-way reporting is crucial to ensuring the efficacy of the transparency they demand of business. It also allows for the external world can hold them to account should failings be revealed.
- Information the FCA could release about firms, individuals, markets (disclosure as a regulatory tool): information released to inform consumers provides an incentive for firms to change their behaviour in beneficial ways.
- Information the FCA could require firms to release about their products and about other aspects their performance and behaviour: they make new rules, so firms disclose information about their behaviour and product performance, to allow market participants to make informed decisions and to better judge, either directly or via intermediaries, which product is most appropriate for their needs.

**Who:** Transparency helps empower consumers but also allows critical assessment of the FCA processes by external stakeholders.

**Why was transparency developed? Was it voluntary or mandated by Government?**

**Why was transparency developed?**

Transparency was developed as the FCA believes it will help consumers make more informed choices whilst also changing consumer or firm behaviour in ways that help us achieve our statutory objective. It was also implemented to allow for regulatory transparency and provide external stakeholders with the information hold the FCA to account.

**Voluntary or Mandated?**

Parliament has given them the statutory powers to allow the publication of information, and if publishing that information helps the FCA achieve their objectives, then it is legitimate for them to do so<sup>144</sup>.

<sup>143</sup>FCA: 'FCA Transparency Framework' <https://www.fca.org.uk/publication/feedback/fca-transparency-framework.pdf>

<sup>144</sup>FCA: 'Transparency as a Regulatory Tool' (2008), <https://www.fca.org.uk/publication/discussion/fsa-dp08-03.pdf>

**What is the key information / mechanisms of the transparency?**

- Online systems for submission and storage of reporting data (including transparency reports) – Since 2002, the FCA have operated an online collection system GABRIEL and are currently undergoing updating this system to an improved platform- RegData
- ‘Dear CEO’ Letters – Letters to firms which openly update and inform on regulatory activity, updates to processes and requests for consultations (regulator transparency)
- Encouraging good practice via the publication of research and consultation papers – Several FCA publications on consumer vulnerability led to the creation of the ‘UK Finance (formerly BBA) Financial Services Vulnerability Taskforce’
- Educating consumers on harm with the ‘ScamSmart’ tool – A digital tool to inform consumers on the dangers of pensions and investment scams
- Complaints Data – Mandated disclosure of complaints data for firms with over 500 complaints, suggested to encourage consumer choice and encouraging better practice and competition between firms
- Horizon Scanning for future and emerging risks – Carrying out research such as the annual ‘Financial Lives Survey’ to assess current trends in financial harms and vulnerabilities to better inform practice.
- Innovation Focus<sup>145</sup> - Development of initiatives to promote two-way conversations between the FCA and innovation in industry, including a regulatory sandbox and dedicated advice unit.

See table in annex for further detail <sup>146</sup>

**What else is required to make the model work? What is the necessary context?**

In addition to conducting extensive research into transparency, the FCA adopt an honest and open approach to improving their processes, openly requesting feedback, consultation, and debate to inform their decision-making. This open and iterative approach to developing transparency mechanisms may be useful in the context of online safety regulation whereby it is largely a new challenge requiring the testing of different solutions.

They also set out in clear terms what they must, can and cannot disclose under FoIA and the Financial Services and Markets Act 2000 (FSMA).

**What were the processes and steps involved in implementing reporting?**

N/A

<sup>145</sup>FCA: ‘FCA Innovation – fintech, regtech and innovative businesses’ <https://www.fca.org.uk/firms/innovation>

<sup>146</sup>FCA : ‘Transparency as a Regulatory Tool’ <https://www.fca.org.uk/publication/discussion/fsa-dp08-03.pdf>

**What are the risks? Are there any unintended consequences?**

In Interview 9, issues were raised as to the efficacy of the transparency regimes. It is widely accepted that transparency is required in the industry, but this occurs at huge costs and with little evidence of the positive impacts it actually has.

**What lessons can be learned from this model and be applied to Online Safety?**

Disclosure alone is not sufficient to provide transparency. Information should be clear enough to be processed by consumers in order to have an impact on making products, firms, or markets more transparent.

Excessive or overly complex information undermines the benefits of providing information to all consumers.

Consumers will not necessarily act on receipt of simple information i.e. consumer behaviour may be inconsistent with the logic of efficiency that assumes consumers will always seek to maximise product value.

Behavioural economics (understanding consumer behaviour and rationale) may inform the design of disclosure methods, but the evidence is limited.

Transparency will always remain limited unless the disclosure allows end-consumers to understand information available.

FCA face multiple organisational challenges:

- Astonishing costs to make changes to information requirements, due to the complexity and constraints (gates and checks needed, security entry points etc) and the race to keep up with technology.
- The need to invest in data capability/people to keep pace.
- The need to prioritise better what you need given the complexity and proliferation of data.
- For PRIIPs the need to build better systems overall. Needs to be in some way more flexible. Firms build in a costly way, FCA want to build it in a flexible way but firms have no incentives to do it flexibly. Even for the bigger firms they regulate, disclosure and reg reporting is a complex thing for them to figure out.
- The challenge of standardising definitions (e.g., even 'profit' has different meanings to different organisations)

## Utilities

**Status:** Implemented

### Principal mechanisms employed:



Education  
(Consumers)



Performance  
Information



Performance /  
Behavioural  
Accountability

**Summary:** Ofgem enforces the frequent reporting of supplier performance and complaints information (monthly/quarterly).

Information regarding a supplier's performance or complaints is then made public both on supplier websites and being published by Ofgem themselves and included in performance tables for easy comparison of performance between suppliers. This transparency assists Ofgem in understanding problems consumers face in the market but also imposes reputational pressure on consumers performing poorly compared to competitors to encourage service improvement.

### Ofgem Complaints Reporting

#### General Overview

The Authority's primary duty when carrying out its functions is to protect the interests of consumers, present and future, wherever appropriate, by promoting effective competition between persons engaged in, or in commercial activities connected with, the shipping, transportation or supply of gas conveyed through pipes and the generation, transmission, distribution or supply of electricity or the provisions or use of electricity interconnectors<sup>147</sup>.

As part of this remit, Ofgem enforces the reporting of complaints data to them on a monthly and quarterly basis<sup>148</sup>, publicly on supplier websites and the transparency of energy company profitability<sup>149</sup>.

This case study shall predominantly focus on Ofgem's complaints reporting.

#### What does transparency help with?

**Transparency mechanisms:** Suppliers are required to submit performance information for complaints data to Ofgem on a monthly and quarterly basis. They must also publish domestic complaints data on their websites, including their 'top 5' reasons for complaints and the measures they are taking to improve how they handle customer complaints<sup>150</sup>.

**What transparency helps with:** Transparency helps Ofgem better understand the problems consumers may face in the retail energy market and the actions we could take to protect their interests as well as empowering consumers to choose suppliers

<sup>147</sup> Sustainable Development Commission: 'Lost in Transmission?' <https://research-repository.st-andrews.ac.uk/bitstream/handle/10023/2344/sdc-2007-lost-in-transmission.pdf?sequence=1&isAllowed=y>

<sup>148</sup> Ofgem: 'Compare supplier performance on complaints', <https://www.ofgem.gov.uk/consumers/energy-supplier-comparison-data/compare-supplier-performance-complaints>

<sup>149</sup> Ofgem: 'Transparency Consultation', [https://www.ofgem.gov.uk/sites/default/files/docs/2014/10/transparency\\_consultation\\_0.pdf](https://www.ofgem.gov.uk/sites/default/files/docs/2014/10/transparency_consultation_0.pdf)

<sup>150</sup> Ofgem: 'Compare Supplier Performance Complaints', <https://www.ofgem.gov.uk/consumers/energy-supplier-comparison-data/compare-supplier-performance-complaints>

based on these statistics. Transparency also promoted performance accountability and could lead to reputational pressure should suppliers fall below the standards expected by consumers.

**Why and how is transparency effective?  
What is the intended audience for the transparency information?  
Do the benefits spill out beyond that audience?**

**Why is transparency effective:** Greater transparency on complaints and complaints handling can provide an important incentive to suppliers to improve their performance as it allows consumers to hold them accountable and make active decisions on their suppliers based on complaint performance. It also aims to educate customers about suppliers' performance and thereby enable customers to engage more effectively in the market<sup>151</sup>.

**How is transparency effective:** The monthly requirements placed on businesses to submit their complaints data is crucial to keeping the transparency reporting effective. Being able to show changes in complaints will allow consumers to make informed decisions on current performance as opposed to out-dated yearly reports. The domestic complaints that are posted on provider websites also create accountability and educates consumers as to the top issues had.

**Who:** Predominantly in place to educate consumers and empower them to make informed choices with their supplier?

**Why was transparency developed? Was it voluntary or mandated by Government?**

**Why was transparency developed?**

Transparency was developed to protect the interests of consumers and hold energy suppliers accountable to the standards that are expected of them.

**Voluntary or mandated?**

An amendment to the Gas and Electricity Acts by virtue of Schedule 5 of the Consumers, Estate Agents and Redress Act 2007 (CEARA) requires the Authority to collect information from licence holders relating to complaints handling standards. However, suppliers have been voluntarily submitting customer complaints data to us and Citizens Advice (including Citizen Advice's predecessors) since the inception of the Complaint Handling Regulation Standards Regulations (CHR) from 1 October 2008<sup>152</sup>.

**What is the key information / data that is reported?**

Ofgem, Citizens Advice and the Ombudsman publish complaints statistics. The separate statistics show how many complaints suppliers receive, how many Citizens Advice handle, weighted by the seriousness of the complaint, and how many are accepted by the Ombudsman after failing to be resolved by the supplier. This helps to show a full picture of the customer's complaint from start to finish.

<sup>151</sup> Littlechild, S. (2020): 'An Overall Customer Satisfaction score for GB energy suppliers' <https://www.repository.cam.ac.uk/bitstream/handle/1810/314746/cwpe2090.pdf?sequence=1&isAllowed=y>

<sup>152</sup> Ofgem: 'Transparency Consultation', [https://www.ofgem.gov.uk/sites/default/files/docs/2014/10/transparency\\_consultation\\_0.pdf](https://www.ofgem.gov.uk/sites/default/files/docs/2014/10/transparency_consultation_0.pdf)

**What else is required to make the model work? What is the necessary context?**

Constant assessment is crucial to ensuring on-going relevance and efficacy. In October-December 2015, Ofgem reviewed suppliers' direct domestic complaints data jointly with Citizens Advice. This review found some differences in suppliers' approaches to reporting some statistics, which indicated the need to improve consistency of complaints data recording across suppliers, as well as reduce discrepancies between the complaint information submitted to us and that published on suppliers' webpages. This led to large scale consultation and the production of a set of guidelines which clearly outline key definitions and how reporting should take place<sup>153</sup>.

**What were the processes and steps involved in implementing reporting?**

Ofgem have introduced a mandatory common format that all active energy suppliers must use to report their domestic and micro-business complaints information to them. This ensures comparability of all data received and helps ensure the data remains useful to consumers.

**What are the risks? Are there any unintended consequences?**

None observed.

**What lessons can be learned from this model and be applied to Online Safety ?**

It has been flagged across several interviews that a key challenge to regulating Online Safety is the challenge of taxonomy and lack of consistency across platforms. In this example, Ofgem created clear guidelines for all suppliers to ensure that 'complaints' were defined universally and that the format of the data was standardised and comparable.

It's also important to note that regular assessment of the transparency regime is required to constantly assess its efficacy in achieving the designated targets.

<sup>153</sup>Ofgem: 'Guidance to submitting customer complaints data', <https://www.ofgem.gov.uk/publications-and-updates/guidance-submitting-customer-complaints-data>

## Applying lessons from legislation

The following section outlines several domestic and international approaches towards transparency and online safety legislation, both implemented and provisional, and evaluates successes and potential shortfalls.

### Australia – Online Safety Act

**Status:** Implemented

**Principal mechanisms employed:**



**Summary:** The Australian eSafety Commissioner was established as an independent statutory office under the Enhancing Online Safety Act 2015. Transparency in this context focuses around two key areas: Transparency reporting on both a voluntary and mandated level for platforms based on their size and ‘safety by design’ principles to promote safe operations from the onset of a business’s operation (core to their business model).

### Australian eSafety Commissioner<sup>154</sup>

#### General Overview

The Australian eSafety Commissioner was established as an independent statutory office under the Enhancing Online Safety Act 2015. Initially, these functions primarily related to enhancing online safety for Australian children however, in 2017, the Act was amended to expand the Commissioner’s remit to promoting and enhancing online safety for all Australians.

#### What does transparency help with?

**What Transparency Mechanisms:** In this case study there are two key types of transparency that will be looked into:

1. Transparency reporting on both a voluntary and mandated level for platforms based on their size which seeks to increase corporate accountability.
2. In-design of safety and transparency principles from the start of a business’s operations through the ‘safety by design’ principles.

**What Transparency Helps With:** To improve the online safety of all Australians and reduce cyberbullying, image-based abuse, and illegal and harmful online content

#### Why and how is transparency effective? What is the intended audience for the transparency information?

**Why transparency is effective:** Transparency is being used to hold platforms accountable for the content that they may host/ distribute and to educate consumers about the measures in place to protect them. By enforcing transparency reporting for larger platforms, the Commissioner leads the way for consumers and other stakeholders to judge their performance in moderating and

<sup>154</sup> eSafety, Our legislative functions, <https://www.esafety.gov.au/about-us/who-we-are/our-legislative-functions>

## Do the benefits spill out beyond that audience?

removing content aligned to key social issues. The safety by design principles are effective by incorporating best practices from the start in start-up companies and ensures that as they upscale so too does their corporate responsibility.

**How transparency is effective:** The two-tier ranking approach makes transparency reporting a legal commitment for large social media platforms (and other services determined by the Minister for Communications) whilst having a voluntary process for smaller, less risky platforms. This split ensures that the largest platforms are still held to high standards of transparency reporting by law whilst giving smaller platforms to opportunity to be involved without consequence. For smaller platforms, whilst voluntary, this still becomes effective due to the publicly available list of those who have registered and creates an informal reputational pressure model to promote uptake.

Furthermore, through their Safety by Design initiative, transparency is being embedded into organisations to ensure visibility of user-safety policies, annual assessments and safety considerations are present from the start<sup>155</sup>. It is theorised that by incorporating values and practices from the start, transparency and safety will continue to be present as the company grows.

**Who:** The intended audience for transparency is the general public but emphasis is placed on the requirement to 'engage with a wide user-base, including experts and key stakeholders on the development, interpretation and application of safety standards' to ensure maximum benefits for the widest audience.<sup>156</sup>

## Why was transparency developed? Was it voluntary or mandated by Government?

### Why was transparency developed?

The transparency was developed to initially enhance online safety for Australian children but was expanded to protect all Australians online.

### Voluntary or mandated?

Whilst the eSafety Commissioner has powers relating to the removal and escalation of all types of harmful content online, the transparency reporting is a combination of voluntary and compulsory dependent on the platform.<sup>157</sup>

## What is the key information / data that is reported?

As part of the safety by design, transparency and accountability is built through:

- Embed user safety considerations, training and practices into the role, functions and working practices of individuals who work on products or services
- Ensure visibility of user-safety policies
- Openly engage with a wide userbase on the development of standards and their effectiveness

<sup>155</sup> eSafety Commissioner: 'Safety by Design', <https://www.esafety.gov.au/about-us/safety-by-design>

<sup>156</sup> eSafety Commissioner: 'Social Media Tier Scheme', <https://www.esafety.gov.au/about-us/who-we-are/social-media-tier-scheme>

<sup>157</sup> eSafety Commissioner: 'What we can investigate', <https://www.esafety.gov.au/report/illegal-harmful-content/what-we-can-investigate>

	<ul style="list-style-type: none"> <li>• Publish annual assessments of reported abuses and meaningful analysis</li> </ul>
<p><b>What else is required to make the model work? What is the necessary context?</b></p>	<p>The Commissioner has a number of measures that they can take in the event of a platform not complying with the regulations. Civil penalties can be applied and there is the potential for non-compliance to be raised to a criminal offence.<sup>158</sup></p> <p>This therefore strikes the balance between enforced transparency, voluntary transparency and building transparency from the ground up.</p>
<p><b>What were the processes and steps involved in implementing reporting?</b></p>	<p>eSafety engages with the community around particular issues and sometimes consults directly with the public about their activities and what they should focus on. They also work with social media services who are operating under the Tier scheme to ensure their compliance.</p>
<p><b>What are the risks? Are there any unintended consequences?</b></p>	<p>None observed.</p>
<p><b>What lessons can be learned from this model and be applied to Online Safety?</b></p>	<p>As this model relates directly to Online Safety there are some significant learnings that can be taken. Notably, building in transparency from the ground up through their Safety by Design initiative has significant benefits and removes the burden of safety from the end user. The balance of enforced and voluntary transparency allows for this model to be scalable for different business sizes.</p>

<sup>158</sup> eSafety Commissioner: 'The actions we can take', <https://www.esafety.gov.au/report/illegal-harmful-content/the-actions-we-can-take>

## France – Avia Law

**Status:** This law was implemented in March 2019 but was subsequently declared unconstitutional in June 2019.

### Principal mechanisms employed:



1-Way  
Reporting



Performance /  
Behavioural  
Accountability

**Summary:** The French Avia Law was drafted in March 2019 but was later declared unconstitutional due to infringements of freedom of speech. The main criticisms related to the introduction of near instantaneous content takedown requirements. Transparency was also due to be implemented via the publication of a list of “banned” services online to create reputational pressure of non-compliant online service providers.

### French Avia Law<sup>159</sup>

#### General Overview

In March 2019, the Law on countering Online Hatred, commonly known as the Avia Law was drafted. The Avia law was adopted in May despite being severely criticised from a range of parties including the European Commission, the Czech Republic, digital rights organisations and LGBTQI+, feminist and antiracist organisations.

On 18 June, the French Constitutional Council declared the main provisions of the Avia law unconstitutional, due to certain provisions infringing “on freedom of speech and communication”.<sup>160</sup>

#### What does transparency help with?

**Transparency Mechanisms:** The introduction of “quasi-instantaneous take down requirements and increasing liability and sanctions where these new “notice and take down” rules are not complied with”<sup>161</sup>. Transparency was further created through the administrative authority maintaining a list of online services that have been subject to blocking requests and making these available to online advertisers. If a site is on the list of “banned” services, then any advertisers still in relation to these will be made public at least once a year as a form of behavioural accountability and reputational pressure.

**What Transparency Helps With:** The aim of the French Avia Law was to stem all online hate and extremist content online.

<sup>159</sup> Schuler and Znaty (2020): ‘New law to fight online hate speech to reshape notice, take down and liability rules in France’, <https://www.taylorwessing.com/en/insights-and-events/insights/2020/05/new-law-to-fight-online-hate-speech-in-france>

<sup>160</sup> EDRi: ‘French Avia law declared unconstitutional: what does this teach us at EU level?’, <https://edri.org/our-work/french-avia-law-declared-unconstitutional-what-does-this-teach-us-at-eu-level/>

<sup>161</sup> Schuler and Znaty (2020): ‘New law to fight online hate speech to reshape notice, take down and liability rules in France’, <https://www.taylorwessing.com/en/insights-and-events/insights/2020/05/new-law-to-fight-online-hate-speech-in-france>

**Why and how is transparency effective? What is the intended audience for the transparency information? Do the benefits spill out beyond that audience?**

**Why was transparency expected to be effective:** Transparency was hoped to be effective by the reporting of online services who did not conform to the strict takedown requirements of the law. The public-facing nature of these one-way reports was to be used as a form of corporate accountability to encourage compliance due to the reputational damage that being in the report could deliver.

**How was transparency expected to be effective:** By using two layers of transparency, one to advertisers and one to consumers, it was hoped that this mechanism would drive online service providers to abide by the rules through the financial threat of withdrawn advertising, and for advertiser cooperation through the treat of public scrutiny.

*Please note, as this legislation was not fully implemented the context of these questions now focuses on expectation. Where the law failed is addressed further down the table.*

**Who:** Whilst the law was designed to ensure consumer safety through regulation, the transparency reports were designed to change advertiser behaviour and name & shame those who interacted with sites under blocking orders.

**Why was transparency developed? Was it voluntary or mandated by Government?**

**Why was transparency developed?**

The introduction of such strict legislation was triggered by the beheading of schoolteacher Samuel Paty and the increasing prevalence of online hate speech and extremism.

**Voluntary or mandated?**

It was mandated by the Government and involved the introduction of a strict, compulsory framework.

**What is the key information / data that is reported?**

The processes in place to limit content ultimately led to platforms overzealously blocking perfectly legal content to avoid falling foul the regulations. They covered the following:

- Notice-and-action system – any user can flag manifestly illegal content and the notified online service provider is required to remove it within 24 hours
- Reduction of the intermediary’s deadline to remove illegal terrorist content and child sexual abuse material from 24 hours to one hour after the receipt of a notification
- In addition to the 24 hour take down requirement, online platforms and search engines falling in the scope of the new law will also be required to fight the dissemination of illegal content, in particular by complying with the guidance and deliberations of the French Audio-visual Council administrative body (CSA), for the performance of the following obligations:
  - to set up a standardised and easily accessible notification system for hate contents
  - to make available clear, detailed and accessible information describing the terms and conditions for content moderation and to inform users under

	<p>fifteen years of age (or legal guardians) when registering for a service involving the processing of personal data about the civic and responsible use of said service and the risks incurred in the event of the dissemination of hateful content</p> <ul style="list-style-type: none"> <li>○ to acknowledge receipt without delay of the notifications they receive</li> <li>○ to inform the author of the notification of its follow-up and the motives of the take down decision (within 24 hours when the content is withdrawn or made inaccessible and within seven days otherwise)</li> <li>○ to implement internal recourse mechanisms allowing the author of the notification or the author of the withdrawn content to challenge a decision with which they disagree</li> <li>○ to inform public authorities of hate content notified to them</li> <li>○ to implement the appropriate technological and human means to process notifications as promptly as possible and to provide a report on the means employed and their performance</li> <li>○ to designate a contact person responsible for receiving requests from the judicial authorities and the CSA.</li> </ul>
<p><b>What else is required to make the model work? What is the necessary context?</b></p>	<p>Quasi-instantaneous take-down requirements</p> <p>Increased liability and sanctions</p> <p>Enhanced cooperation obligations for online platforms to fight the dissemination of illegal content</p>
<p><b>What were the processes and steps involved in implementing reporting?</b></p>	<p>Heavy regulation and sanctions were used to ensure compliance.</p>
<p><b>What are the risks? Are there any unintended consequences?</b></p>	<p>The reasons for why the Avia Law was declined.</p> <ul style="list-style-type: none"> <li>• Disagreement within the legal measures in the law that were used to limit manifestly illegal content:</li> <li>• Notice-and-action system – any user can flag manifestly illegal content and the notified online service provider is required to remove it within 24 hours</li> <li>• Reduction of the intermediary’s deadline to remove illegal terrorist content and child sexual abuse material from 24 hours to one hour after the receipt of a notification</li> <li>• All transparency obligations in relation to unconstitutional removal measures</li> <li>• The power given to the French High Audio-visual Council with an oversight mandate to monitor the implementation of obligations</li> </ul>

- Politicians voted against the bill, due to a fear of giving too much censorship power to private companies, for example Google and Facebook.

The lack of Judicial involvement in the decision to determine whether content is legal or not and incentives (strict deadlines and heavy sanctions) to overzealously block perfectly legal speech declared as a clear breach of the French constitution.

Only the removal of manifestly illegal content can be ordered without a judge's prior authorisation, which requires some degree of analysis.

**What lessons can be learned from this model and be applied to Online Safety ?**

The Avia Law is a good example of how overly draconian measures with strict deadlines and heavy sanctions are unlikely to create a positive transparency model. It should therefore be assessed as to how strict regulation on illegal content can coexist with freedoms of expression in any future transparency models. These should consider how content is flagged and deleted but also how much censorship power should sit with the platform versus the state.

## European Union - Mandatory Non-Financial Disclosures

**Status:** Implemented

**Principal mechanisms employed:**



**Summary:** The EU Non-Financial Reporting Directive outlines legal requirements for companies to disclose information relating to social and environmental challenges, “EU rules require large companies to publish regular reports on the social and environmental impacts of their activities”<sup>162</sup>.

### Mandatory Non-Financial Disclosures (Directive 2014/95/EU)

#### General Overview

Under EU Directive 2014/95/EU (Non-Financial Reporting Directive), certain large companies are mandated to report information regarding social and environmental challenges and how their operations impact them. The rules apply to ‘large public-interest companies’ with an excess of 500 employees, covering an estimated 11,700 companies within the EU across a range of sectors. The requirements move beyond environmental information and include social matters such as treatment of employees, diversity, human rights as well as anti-corruption and bribery<sup>163</sup>.

#### What does transparency help with?

**Transparency Mechanism:** Companies are required to disclose ‘relevant, useful information’ (through reporting) necessary to glean understanding of their development, performance, position, and the overall impact of their business activity rather than providing exhaustive detailed reports that are difficult to understand. The directive aims to provide the flexibility for companies to disclose information in a manner which they consider to be most useful<sup>164</sup>.

Whilst as it stands 2014/95/EU does not require the audit of information, the recent Corporate Sustainability Reporting Directive (CRSD) adopted in April 2021 would amend the current directive and require the auditing of disclosed information going forward.

<sup>162</sup> EU: Corporate sustainability reporting [https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting\\_en#disclosed](https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en#disclosed)

<sup>163</sup> Ibid.

<sup>164</sup> European Commission, 2017: Information from European Union Institutions, Bodies, Offices and Agencies ‘Guidelines on non-financial reporting’ [https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52017XC0705\(01\)&from=EN](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52017XC0705(01)&from=EN)

	<p><b>What Transparency Helps With:</b> The purpose of the directive is to increase transparency of companies activity to increase company resilience, performance and increase trust amongst stakeholders<sup>165</sup>. The <b>accompanying guidelines</b> aim to assist companies in compliance, specifically noting that ‘significant efforts’ to avoid undue administrative burdens, boilerplate disclosures or box-ticking has been made<sup>166</sup> to reduce the burden on companies and increase the usefulness of information to disclosed to relevant stakeholders.</p>
<p><b>Why and how is transparency effective? What is the intended audience for the transparency information? Do the benefits spill out beyond that audience?</b></p>	<p><b>Why is transparency effective:</b> Clear, accurate and accessible information in theory makes it easier for stakeholders to understand information disclosed whilst flexible reporting guidelines reduce the administrative burden on companies to comply with the directive.</p> <p>However, the quality of disclosure has been criticised for a lack of consistency in the quality, accessibility, and overall compliance<sup>167</sup>.</p> <p><b>Who:</b> Company stakeholders, including investors and consumers.</p>
<p><b>Why was transparency developed? Was it voluntary or mandated by Government?</b></p>	<p><b>Why was transparency developed?</b> Came into force in December 2014, replacing Directive 2013/23/EU (2) in recognition of the need to improve the consistency of transparency information across all member states with regards to the disclosure of non-financial information<sup>168</sup>.</p> <p><b>Voluntary or mandated?</b></p> <p>Mandated</p>
<p><b>What is the key information / data that is reported?</b></p>	<p>The directive outlines that eligible undertakings “should prepare a non-financial statement containing information relating to at least <b>environmental matters, social and employee-related matters, respect for human rights, anti-corruption, and bribery matters</b>. Such statement should include a description of the <b>policies, outcomes and risks</b> related to those matters and should be included in the management report of the undertaking concerned. The non-financial statement should also include information on the <b>due diligence processes</b> implemented by the undertaking, also regarding, where relevant and proportionate, its supply and subcontracting chains, in order to identify, prevent and mitigate existing and potential adverse impacts”<sup>169</sup>.</p>
<p><b>What else is required to make the model work? What is the necessary context?</b></p>	<p>Unknown.</p>

<sup>165</sup> Ibid.

<sup>166</sup> Ibid.

<sup>167</sup> Allen & Overy, 2020 ‘Review of the Non-Financial Reporting Directive: towards an EU-wide ESG reporting standard’ <https://www.allenoverly.com/en-gb/global/news-and-insights/publications/review-of-the-non-financial-reporting-directive>

<sup>168</sup> EU, 2014: Directive 2014/95/EU of the European Parliament and of the Council <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014L0095&from=EN>

<sup>169</sup> Ibid.

<p><b>What were the processes and steps involved in implementing reporting?</b></p>	<p>Came into force in late 2014 having been initially proposed in 2013 – See ‘Proposal for a Directive of the European Parliament and of the Council amending Council Directives 78/660/EEC and 83/349/EEC as regards disclosure of non-financial and diversity information by certain large companies and groups’<sup>170</sup> for more information.</p>
<p><b>What are the risks? Are there any unintended consequences?</b></p>	<p>Offering too much flexibility within a framework can lead to ambiguity in specific disclosure requirements risks impeding the consistency and quality of information disclosed. This has been a criticism of the efficacy of Directive 2014/95/EU<sup>171</sup>.</p>
<p><b>What lessons can be learned from this model and be applied to Online Safety?</b></p>	<p>Disclosure requirements can be tailored for different organisation size which can reduce overburdening smaller organisations. Additionally, where flexible disclosure requirements are provided, efforts may still be made to ensure that information disclosed still meets expected outcomes in terms of type and quality of information.</p>

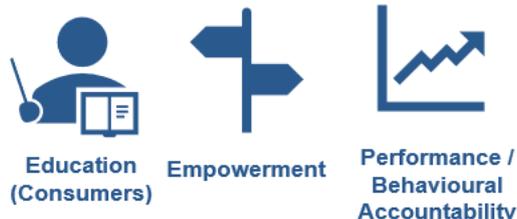
<sup>170</sup> European Commission, 2013: ‘Proposal for a Directive of the European Parliament and of the Council amending Council Directives 78/660/EEC and 83/349/EEC as regards disclosure of non-financial and diversity information by certain large companies and groups’ <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52013PC0207&from=EN>

<sup>171</sup> Allen & Overy, 2020 ‘Review of the Non-Financial Reporting Directive: towards an EU-wide ESG reporting standard’ <https://www.allenoverly.com/en-gb/global/news-and-insights/publications/review-of-the-non-financial-reporting-directive>

## European Union – General Data Protection Regulation (GDPR)

**Status:** Implemented

**Principal mechanisms employed:**



**Summary:** The General Data Protection Regulation, created in 2016, is a regulation in EU law on data protection and privacy in the European Union and the European Economic Area. Transparency in this initiative is designed to disclose to users, information on the collection and usage of their personal data. Transparency is implemented via clear labelling on data usage, for example when users access a site with notifications, and the empowerment of consumers of the right to access and erase personal data held by organisations.

<b>General Data Protection Regulation (GDPR)</b>	
<b>General Overview</b>	The General Data Protection Regulation, created in 2016, is a regulation in EU law on data protection and privacy in the European Union and the European Economic Area. The regulation lays down rules relating to the protection of natural persons with regard to the processing of personal data and rules relating to the free movement of personal data as well as protection the fundamental rights and freedoms in relation to the right to the protection of personal data. <sup>172</sup>
<b>What does transparency help with?</b>	<p><b>Transparency Mechanisms:</b> One of the core aspects within GDPR is the right to be informed. In this context, transparency refers to the reporting of data held on an individual and how it is to be used. This could be through emails, when accessing a site or clearly labelling in terms and conditions. Transparency is further bolstered by the right of consumers to actively request the information held on them by a business.</p> <p><b>What Transparency Helps With:</b> It is designed to clearly keep an individual informed about the collection and use of their personal data. Other aspects of GDPR are improving consumer empowerment to access the data held on them and their rights to erasure.</p>
<b>Why and how is transparency effective?</b> <b>What is the intended audience for the transparency information?</b> <b>Do the benefits spill out beyond that audience?</b>	<p><b>Why is transparency effective:</b> GDPR gives consumers the education to understand how their data is collected, stored, and used. The resulting empowerment and ability to hold platforms accountable for their data was not previously held and forced a fundamental change in the operations of businesses and their data handling practices. This form of transparency is hoped to be effective by providing consumers with the power to choose who uses their data and ultimately alter their online behaviours to align to this.</p>

<sup>172</sup> Intersoft Consulting: 'Article 1 GDPR' <https://gdpr-info.eu/art-1-gdpr/>

**How is transparency effective:** This form of transparency is effective as it creates an environment where consumers have the power to hold platforms accountable for the data that they collect on them. Transparency is effective due to the legislative backing of the GDPR requirements. All companies being forced to abide by the regulations ensures greater consumer confidence as to how their data is being used and how they are being tracked online.

**Who:** Whilst the intention was to provide greater protection of consumer data, and a greater transparency of who is holding it, businesses have also seen benefits of the strict regulations as there is improved consumer confidence for companies seen to be GDPR compliant.

### Why was transparency developed? Was it voluntary or mandated by Government?

#### Why was transparency developed?

The EU says GDPR was designed to "harmonise" data privacy laws across all its members countries as well as providing greater protection and rights to individuals. GDPR was also created to alter how businesses and other organisations can handle the information of those that interact with them <sup>173</sup>.

#### Voluntary or mandated?

Mandated

### What is the key information / data that is reported?

Whilst GDPR encompasses the right to be informed, of access, rectification, erasure, restrict processing, data portability and the right to object, the right to be informed reflects how transparency reports are being implemented for Online Safety . As part of the right to be informed, companies must conform to the following<sup>174</sup>:

- Individuals have the right to be informed about the collection and use of their personal data. This is a key transparency requirement under the UK GDPR.
- You must provide individuals with information including: your purposes for processing their personal data, your retention periods for that personal data, and who it will be shared with. We call this 'privacy information'.
- You must provide privacy information to individuals at the time you collect their personal data from them.
- If you obtain personal data from other sources, you must provide individuals with privacy information within a reasonable period of obtaining the data and no later than one month.
- There are a few circumstances when you do not need to provide people with privacy information, such as if an individual already has the information or if it would involve a disproportionate effort to provide it to them.

<sup>173</sup> Wired: 'What is GDPR? The summary guide to GDPR compliance in the UK' <https://www.wired.co.uk/article/what-is-gdpr-uk-eu-legislation-compliance-summary-fines-2018>.

<sup>174</sup> ICO: 'Right to be informed', <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/right-to-be-informed/>

- The information you provide to people must be concise, transparent, intelligible, easily accessible, and it must use clear and plain language.
- You must regularly review, and where necessary, update your privacy information. You must bring any new uses of an individual's personal data to their attention before you start the processing.

**What else is required to make the model work? What is the necessary context?**

There are two key factors that have led to the successful roll-out of GDPR that should be considered:

Firstly, the backing of the EU and enforced nature of the legislation ensures all businesses comply. The clearly defined actions that business have to take, alongside the potential for large-scale fines are crucial to forcing company compliance.

Secondly, each nation covered by GDPR was given a level of autonomy in managing how it is regulated. This flexibility ensures market-level approaches could be taken and further ensure greater compliance. As an example, France took a much stronger, punitive stance whereas in the UK there was a greater hand-holding of businesses as part of a collaborative approach.

**What were the processes and steps involved in implementing reporting?**

Clear instructions to ensure compliance stated that is a business collect information about individuals for any reason other than their own personal, family or household purposes, they need to comply.

The UK data protection regime is set out in the DPA 2018, along with the UK GDPR. It takes a flexible, risk-based approach which puts the onus on the business to think about and justify how and why they use data.

The ICO regulates data protection in the UK. They offer advice and guidance, promote good practice, carry out audits, consider complaints, monitor compliance and take enforcement action where appropriate.<sup>175</sup>

**What are the risks? Are there any unintended consequences?**

Two issues which arose from GDPR were the costs of implementation and the impact on consumers that overregulation and caution had:

1. The cost of achieving compliance for many organisations reached thousands of euros which accumulated quickly for smaller businesses.
2. Consumers all experienced issues with double verification, excessive emails requesting their consent and incessant pop-ups when browsing online.

**What lessons can be learned from this model and be applied to Online Safety ?**

In order for a model to be effective, it is clear that there needs to be some level of Government mandate that makes the measures compulsory for at least the largest of businesses. Giving nations flexibility in how to manage the rollout of GDPR showed the value that being adaptable has. Being clear with what's expected, and

<sup>175</sup> ICO: 'Some Basic Concepts', <https://ico.org.uk/for-organisations/guide-to-data-protection/introduction-to-data-protection/some-basic-concepts/>

the repercussions for non-compliance is key but also being able to help businesses through the initial stages is also important.

## European Union – Digital Services Act

**Status:** Not yet implemented

**Principal mechanisms employed:**



1-Way  
Reporting



Performance /  
Behavioural  
Accountability

**Summary:** The European Union’s proposed Digital Services Act aims to create an EU-wide framework to bridge the information asymmetries between online intermediaries and users. Under this initiative digital service providers will be obligated to disclose, via reporting, a range of information including on content takedown, advertising, and the use of algorithms.

### Digital Services Act

#### General Overview

The EU Digital Services Act aims to provide a common set of rules to create a safer and more open digital space. The proposal for this legislation was published on 15 December 2020 as part of the Digital Services Act package (Digital Services Act and Digital Markets Act). Agreement on a final text is needed before regulations will be adopted, and it may take years before the rules are adopted, implemented, and enforced.

“The DSA adopts a graduated approach to online platforms, with different rules applying for different platform ‘sizes’. Only very large online platforms (VLOPs) will be subject to the full scope of the proposed Regulation. VLOPs “provide their services to a number of average monthly active recipients of the service in the Union equal to or higher than 45 million”<sup>176</sup>

#### What does transparency help with?

**Transparency Mechanisms:** The EU Digital Services Act aims to create an EU-wide framework on:

- the handling of illegal or potentially harmful content online,
- the liability of online intermediaries for third party content,
- the protection of users’ fundamental rights online

As part of this, digital service providers will have obligations to disclose key information to regulators.

**What Transparency Helps With:** It seeks to bridge the information asymmetries between online intermediaries and their users.

<sup>176</sup> Wagner and Jassen, 2021, “A first impression of regulatory powers in the Digital Services Act”, <https://verfassungsblog.de/regulatory-powers-dsa/>

**Why and how is transparency effective?  
What is the intended audience for the transparency information?  
Do the benefits spill out beyond that audience?**

**Why transparency is expected to be effective:** All digital service providers will have obligations to disclose to regulators and users how algorithms work, how decisions to remove content are taken, and on the way that advertisers target users. This one-way occurrence reporting and performance information hopes to be effective in bridging the information asymmetry that currently exists between online platforms, key stakeholders, and users, and ultimately create new standards of accountability.

**How transparency is expected to be effective:** This reporting is hoped to be effective by providing regulators with the information that they need to ensure platforms are acting in the public interest and in line with their codes of practice. It was noted in interview 16 that having insight into algorithms, decisions to remove content, and advertising targeting is critical for not only holding platforms accountable but also understanding the scale of current and future problems. Transparency in the instance of the DSA is therefore expected to be effective by holding platforms accountable for the handling of illegal or potentially harmful content online; the liability of online intermediaries for third party content; and, the protection of users' fundamental rights online.

*N.B. Currently this is only proposed legislation and so the above are hypothesised conclusions.*

**Who:** The goal, once implemented, for this legislation is to benefit the general public by driving greater platform accountability.

**Why was transparency developed? Was it voluntary or mandated by Government?**

**Why was transparency developed?**

The Digital Services Act aims to make platforms responsible for hosted context by updating the EU's legal framework on illegal content, transparent advertising, and disinformation.

It was also created to fill the "regulatory gap" that exists in the absence of no dedicated platform regulator in the EU which could oversee and enforce aspects such as content moderations and transparency <sup>177</sup>.

**Voluntary or mandated?**

The EU have a history of strong rhetoric against big tech giants and much of the legislation will focus on ensuring the rules of the EU single market are applied online, and to address anti-competitive behaviour by big tech firms. It will therefore be mandated by government and not voluntary.

**What is the key information / data that is reported?**

All digital service providers will have obligations to disclose to regulators and users how algorithms work, how decisions to remove content are taken, and the way advertisers target users.

Annual reports on content moderation should be published including the following information:

- Amount of content taken down and average takedown time (by member state and by type of illegal activity)
- Proactive content moderation initiatives (measures that affect the availability, visibility, and accessibility of information)

<sup>177</sup> Wagner and Ferro (2020), 'Governance of Digitalization in Europe' [https://benwagner.org/wp-content/plugins/zotpress/lib/request/request.dl.php?api\\_user\\_id=2346531&dlkey=8KUHYZ5M&content\\_type=application/pdf](https://benwagner.org/wp-content/plugins/zotpress/lib/request/request.dl.php?api_user_id=2346531&dlkey=8KUHYZ5M&content_type=application/pdf)

- Number of complaints received, basis of complaints, decisions taken (including reversals, and average time taken)
- Number of disputes submitted to the out-of-court dispute settlement bodies
- Number of suspensions imposed, and reasons for suspension
- Any use of automatic content moderation

Online platforms should also publish information on monthly active users every six months, communicate with, and respond to, requests from the Digital Services Coordinator.

**What else is required to make the model work? What is the necessary context?**

Online platforms will need to establish new processes for:

- Internal complaint-handling systems, engage with out-of-court dispute settlement bodies to resolve disputes with users, suspend repeat infringers.
- EU based legal representation
- Dedicated compliance officer(s) for the DSA
- Undergoing annual independent audits
- Giving access to data so authorities and academic researchers can monitor compliance

The EU will establish the following mechanisms to govern transparency:

- Supervisory authorities, the European Commission, and the European Board for Digital Services (group of coordinators to assist with harmonisation of the DSA)
- Fines for non-compliance (up to 6% of the annual income)
- Manage the independent audit process

**What were the processes and steps involved in implementing reporting?**

Due to the Digital Services Act being official legislation, it will need to be passed through and approved by the EU and its member states.

**What are the risks? Are there any unintended consequences?**

A stringent transparency framework could potentially shut out smaller media companies that may struggle to comply with requirements, thereby further strengthening big tech firms (GDPR can be cited as a recent example of this risk). With that said, the proposal notes exemption to some obligations for SME's (small and medium sized enterprises) which could in fact encourage innovation and growth for smaller operations.

There are also a number of potential socio-political risks to regulating freedom of expression. The legislation arguably follows the principle of 'delete first, think later': as soon as anyone flags any content as potentially illegal, online platforms become liable and are required to remove or disable access to the content. This heavy-handed approach could lead to expeditious deletion of content rather than nuanced development of content moderation mechanisms (note the French Avia Law as arguably an example of the failure of heavy-handed legislation).

The Act doesn't discuss the details of what content should be counted as illegal and states that companies must carry out their

own risk assessments about how their platforms could be used for illegal content sharing. This could result in a fragmented approach to mitigate the risks of OH.

Finally, the regulation proposed in the DSA has been noted to be highly centralised with the European Commission putting itself forward as the sole regulator. This has raised questions around the independence and impartiality of regulators and has raised concerns amongst many that there should be an independent, decentralised regulator<sup>178179</sup>

**What lessons can be learned from this model and be applied to Online Safety ?**

It has been noted by several of the interviews that any approach by Ofcom should use consistent core definitions and actions with larger governing bodies.

---

<sup>178</sup> Wagner and Jassen, 2021, 'A first impression of regulatory powers in the Digital Services Act', <https://verfassungsblog.de/regulatory-powers-dsa/>

<sup>179</sup> Wagner and Ferro (2020), 'Governance of Digitalization in Europe' [https://benwagner.org/wp-content/plugins/zotpress/lib/request/request.dl.php?api\\_user\\_id=2346531&dlkey=8KUHYZ5M&content\\_type=application/pdf](https://benwagner.org/wp-content/plugins/zotpress/lib/request/request.dl.php?api_user_id=2346531&dlkey=8KUHYZ5M&content_type=application/pdf)

## UK Modern Slavery Act

**Status:** Implemented

**Principal mechanisms employed:**



Education  
(Consumers)



Performance  
Information



1-Way  
Reporting



Performance /  
Behavioural  
Accountability

**Summary:** The Modern Slavery Act requires many businesses to disclose a 'slavery and human trafficking statement'.

One-way occurrence and performance transparency reporting is implemented with a goal of educating consumers and subsequently driving accountability and reputational pressure for those who fail to comply or are not effective in tackling forced labour and slavery.

Transparency reduces the knowledge gap between large companies and the public to benefit governments and civil society.

### UK Modern Slavery Act

#### General Overview

The Modern Slavery Act is a piece of legislation which sets out a range of measures on how modern slavery and human trafficking should be dealt with in the UK. Section 54 entitled 'Transparency in supply chains' impacts the corporate sector and came into force on 29th October 2015. This requires many businesses to disclose a 'slavery and human trafficking statement'.<sup>180</sup>

In January 2021, it was announced by the UK government that fines for businesses who do not comply with their transparency obligations were going to be introduced to address existing failures to do so.<sup>181</sup>

#### What does transparency help with?

**Transparency mechanisms:** One-way occurrence and performance transparency reporting is implemented with a goal of educating consumers and subsequently driving accountability and reputational pressure for those who fail to comply or are not effective in tackling forced labour and slavery.

**What Transparency Helps With:** Aimed at tackling forced labour and slavery in UK businesses and their supply chains, it seeks to enhance protection and support for victims as well as adding transparency requirement for businesses to show that modern slavery is not taking place in their companies or their supply chains.<sup>182</sup>

<sup>180</sup> PWC, The Modern Slavery Act, <https://www.pwc.co.uk/services/sustainability-climate-change/sustainability-strategy-support/supply-chain/the-modern-slavery-act.html>

<sup>181</sup> Elson, Modern slavery transparency failings to lead to UK fines, <https://www.pinsentmasons.com/out-law/news/modern-slavery-transparency-failings-to-lead-to-uk-fines>

<sup>182</sup> Bloomfield and LeBaron, 2018, The UK Modern Slavery Act: Transparency through Disclosure in Global Governance, <https://www.e-ir.info/pdf/75729>

**Why and how is transparency effective?  
What is the intended audience for the transparency information?  
Do the benefits spill out beyond that audience?**

**Why is transparency effective:** Previously, corporations are afforded high levels of secrecy with respect to their organisational and purchasing practices as “their governance power is exercised through private law mechanisms like commercial contracts (Cutler and Dietz 2017<sup>183</sup>) and enforced through privatised industry-led mechanisms like ethical auditing and certification schemes (LeBaron et al 2017a<sup>184</sup>)”<sup>185</sup>. Transparency is effective through the education and empowerment of consumers which, it is hoped, will hold businesses accountable in the roles to tackle modern slavery.

**How is transparency effective:** Transparency reduces the knowledge gap between large companies and the public to benefit governments and civil society. Prior to the Act, it was largely advocacy groups who provided the public with the information they needed to make informed choices through their consumption decisions and therefore was logical to ensure a clear path of information.

It is hoped that the transparency will be effective in driving businesses to change their practices off the back of consumer pressures.

**Who:** Predominantly seeking to empower consumers with the information to choose where they spend based off this data, it was hoped to create a reputational mechanism of regulation.

**Why was transparency developed? Was it voluntary or mandated by Government?**

**Why was transparency developed?**

The Act was established in 2015 with the intention of eradicating modern slavery in the UK, following a 2013 Government estimate that there are between 10,000 and 13,000 potential victims of modern slavery in the country <sup>186</sup>.

**Voluntary or mandated?**

No penalties can be imposed for non-compliance under the Act currently, though it is open to the UK government to seek court injunctions against businesses that fail to meet their obligations under the legislation.

**What is the key information / data that is reported?**

The Modern Slavery Act sets out corporate transparency requirements applicable to organisations with a turnover of £36 million or more, including that of subsidiaries, that carry on a business or a part of a business in the UK; and that provide goods or services. This set to also extent to UK public bodies with a budget of over £36 million.

Current Home Office guidance ‘recommends’ that organisations include the following areas in a modern slavery statement for the financial year<sup>187</sup>:

- organisational structure and supply chains

<sup>183</sup> Cutler, A. C. and Dietz, T. (eds) (2017) ‘The Politics of Private Transnational Governance by Contract’. New York: Routledge

<sup>184</sup> LeBaron, G., & Rühmkorf, A. (2017a). ‘Steering CSR Through Home State Regulation: A Comparison of the Impact of the UK Bribery Act and Modern Slavery Act on Global Supply Chain Governance’. *Global Policy*, 8, 15–28.

<sup>185</sup> Bloomfield and LeBaron, 2018, ‘The UK Modern Slavery Act: Transparency through Disclosure in Global Governance’, <https://www.e-ir.info/pdf/75729>

<sup>186</sup> Bloomfield and LeBaron, 2018, ‘The UK Modern Slavery Act: Transparency through Disclosure in Global Governance’, <https://www.e-ir.info/pdf/75729>

<sup>187</sup> Deakins, O. ‘UK Government Proposes Tighter Reporting Requirements Under Modern Slavery Act, but No Action on Enforcement Measures’ (2020) <https://www.lexology.com/library/detail.aspx?g=74958926-7f47-4190-a5d7-f49242f0c635>

- policies on modern slavery and human trafficking
- due diligence processes in relation to slavery and human trafficking
- risk assessment and management in relation to modern slavery
- actions taken to prevent slavery and human trafficking in their businesses or supply chains, measured against performance indicators, where possible
- details of staff training undertaken on slavery and human trafficking risks.

Bloomfield and LeBaron (2018) identified two major failings with the Act: there were no checks of compliance, and the Act did not require companies to report on the most important issues (such as ethical certifications and private auditing practices).

“A number of NGOs and other organisations such as the Ethical Trading Initiative and the British Institute of Human Rights have been monitoring compliance reporting requirements. However, it has been noted to be very weak, with no formal legal sanctions other than civil proceedings involving injunctions in the High Court, unlikely to impact significantly on profitability. The government’s view is that reputational pressure, with the financial impact this may have, might be adequate to persuade companies to take effective action, a view not widely shared. Early experience confirms feelings that the provision is inadequate”<sup>188</sup>

**Why the model did not work? What is the necessary context?**

“As the Modern Slavery Act became operational, the government committed to reviewing the National Referral Mechanism, the system by which the claims of those alleging to be victims of modern slavery were assessed (see National Crime Agency, n.d.). The NRM had been widely criticised, including by a consortium of NGOs which argued, inter alia, that the NRM was racist, with those from countries outside the EEA (most of whom were Black or from other minority ethnic groups) standing only one quarter of the chance of having their claims accepted as those from within the EU”<sup>189</sup>

The data collection, recording and analysis within the UK was noted to be extremely deficient from the outset of the legislation. A separate crime recording category of modern slavery was introduced long after the implementation of the Modern Slavery Act and it was noted that, less than one third of the total cases known to the NRM were logged in police records.

There were also large issues in the failure by businesses to develop effective training for the new measures. Craig (2016) noted that a year on from the implementation of the Modern Slavery Act, many organisations were still yet to understand or implement any of the provisions required.

<sup>188</sup> Craig, G. (2016), ‘The UK’s Modern Slavery Legislation: An Early Assessment of Progress’, <https://www.cogitatiopress.com/socialinclusion/article/view/833/833>

<sup>189</sup> Craig, G. (2016), ‘The UK’s Modern Slavery Legislation: An Early Assessment of Progress’, <https://www.cogitatiopress.com/socialinclusion/article/view/833/833>

**What are the risks? Are there any unintended consequences?**

It was noted by Bloomfield and LeBaron (2018) that “many Modern Slavery Act statements focus on: philanthropic activities; bilateral partnerships with NGOs to address social issues; policies about forced labour (e.g., codes of conduct) and private initiatives to ‘monitor’ compliance (e.g., private auditing); and the ethical certifications and voluntary multi-stakeholder initiatives a company subscribes to”<sup>190</sup> which were not effective. It should therefore be noted that transparency reports need to focus on the key measures and not purely be created to satisfy a legislative tick box.

**What lessons can be learned from this model and be applied to Online Safety ?**

It is important to have processes in place to understand the efficacy of current approaches rather than reporting on measures which have not been proven to be effective. Standards should therefore be established for universal indicators which all companies should be forced to report.

There should also be strong measures of enforcement as, in this instance, the lack of penalties for non-compliance led to widespread poor practices from corporations. It is also proven that name-and-shame mechanisms in this instance were not enough to change consumer behaviours and impact businesses.

---

<sup>190</sup> Bloomfield and LeBaron, 2018, ‘The UK Modern Slavery Act: Transparency through Disclosure in Global Governance’, <https://www.e-ir.info/pdf/75729>

## Applying lessons from other approaches

Whilst not necessarily demonstrating example usage cases of transparency, the following sections outline further approaches identified during the research process and have been included for reference.

### Australia – ACMA Annual Compliance Priorities

**Status:** Implemented

**Principal mechanisms employed:**



**Summary:** The Australian Communications and Media Authority (ACMA) are an independent Commonwealth statutory authority who regulate communications and media services in Australia. ACMA produce annual compliance reports demonstrating **regulatory transparency** on their regulatory activity and additionally adapt their regulatory focuses annually based upon ‘annual priorities’ to meet changes in technology and shifts in public interests and concerns.

### ACMA Annual Compliance Priorities<sup>191</sup>

#### General Overview

The Australian Communications and Media Authority (ACMA) are an independent Commonwealth statutory authority who regulate communications and media services in Australia. They have an Authority that makes decisions and an executive team that oversees their work.

#### What does transparency help with?

**Transparency Mechanisms:** Annual compliance reports to provide regulatory transparency on their regulatory activity and highlight areas of public interest or concern.

**What Transparency Helps With:** ACMA’s annual compliance priorities are implemented across the organisation and span beyond transparency reporting alone. It gives the ACMA the ability to refocus regulatory activity in parallel with emerging areas of concern. This allows a degree of flexibility which can see reporting requirements pivoting to meet current and emerging challenges.

#### Why and how is transparency effective? What is the intended audience for the transparency information? Do the benefits spill out beyond that audience?

**Why transparency is effective:** With the ACMA providing self-reported transparency reports, it ensures that they too can be held to account for the efficacy of their transparency practices.

**How transparency is effective:** Annual compliance reports allow for a yearly assessment of their performance and behaviours but also allows for flexibility and adaptability. A regular reporting timescale allows for comparisons to be made with previous reports and assess the changing efficacy of different measures and the emergence of new challenges.

<sup>191</sup> ACMA: ‘Compliance Priorities 2020-21’ <https://www.acma.gov.au/consultations/2019-11/compliance-priorities-2020-2021-consultation#outcome>

**Who:** This reporting benefits not only the ACMA, through allowing self-assessment, but the communications and media services who they regulate. This will ultimately also benefit the end consumer.

**What were the processes and steps involved in implementing reporting?**

In shaping annual priorities, ACMA undergo public consultations to understand public opinion on current policy priorities and any emerging issues or concerns. Specifically, ACMA requests the public consider the following questions<sup>192</sup>:

- What are the matters of significant public interest or concern?
- What are the potential and actual causes of harm to consumers?
- What are the high-level risks of non-compliance, including from technological developments?
- What are the emerging issues where we can encourage compliant behaviour, deter non-compliance, or boost public confidence?
- What are the technological or market developments that test the effectiveness of the regulatory framework?
- In what specific areas can we clarify the scope and reach of the law?

**What are the risks? Are there any unintended consequences?**

There is a potential risk that by shaping annual priorities based alone on challenges or concerns that are known to the public that efforts tackling lesser known or unknown challenges are reduced.

**What lessons can be learned from this model and be applied to Online Safety?**

Given the fast pace at which platforms have demonstrated they can develop and scale, having a flexible framework with regulatory transparency and self-assessment will allow Ofcom to be more adaptable with their focus as new threats emerge and old ones are tackled.

---

<sup>192</sup> Ibid.

## The Santa Clara Principles

**Status:** Guiding principles and therefore not implemented.

**Principal mechanisms employed:**



Education  
(Consumers)



1-Way  
Reporting



Performance /  
Behavioural  
Accountability

**Summary:** The Santa Clara Principles were created in 2018 and outline a framework for the minimum levels of transparency and accountability in content moderation.

<b>The Santa Clara Principles<sup>193</sup></b>	
<b>General Overview</b>	Created on February 2nd, 2018, at the first Content Moderation at Scale conference in Santa Clara, CA, The Santa Clara Principles are recognised by many to serve as a starting point for outlining the minimum levels of transparency and accountability that can form the basis of future dialogue.
<b>What does transparency help with?</b>	<p><b>Transparency mechanisms:</b> Transparency reports and how the provision of data by platforms can be improved.</p> <p><b>What Transparency Helps With:</b> Transparency in the Santa Clara Principles seeks to hold platforms accountable and ensure meaningful due process to impacted speakers and better ensure that the enforcement of their content guidelines is fair, unbiased, proportional, and respectful of users' rights. It also seeks to educate users as to the process of content moderation and removals.</p>
<b>Why and how is transparency effective? What is the intended audience for the transparency information? Do the benefits spill out beyond that audience?</b>	<p><b>Why is transparency effective:</b> Transparency reporting is effective as it holds platforms accountable and ensures that meaningful due process is followed for enforcing content guidelines. It also ensures that the measures taken are unbiased, proportional, and respectful of user rights.</p> <p><b>How is transparency effective:</b> The effectiveness relies on companies being open about the numbers of posts removed, accounts suspended and the wider context of these numbers to help everyone understand the issues. It is also hoped that should platforms be honest and open about their processes regarding content reports, takedowns and enforcement of guidelines, the pressure of consumers and corporate accountability that comes from being open will ensure best practices are followed.</p> <p><b>Who:</b> The individual users as the principles are about enhancing the experience of those reporting and/ or being reported as well as the general population who see that due process will be followed.</p>

<sup>193</sup> The Santa Clara Principles: <https://santaclaraprinciples.org/>

<p><b>Why was transparency developed? Was it voluntary or mandated by Government?</b></p>	<p><b>Why was transparency developed?</b></p> <p>The principles were created in response to the poor standards of existing transparency reporting.</p> <p><b>Voluntary or mandated?</b></p> <p>These are voluntary principles drawn up by academics to provide guidelines as to how companies should perform, moderate and report.</p>
<p><b>What is the key information / data that is reported?</b></p>	<ol style="list-style-type: none"> <li>1. Numbers: Companies should publish the numbers of posts removed and accounts permanently or temporarily suspended due to violations of their content guidelines. Data should include numbers of discrete posts and accounts that are flagged, suspended, and removed across category, source, and location.</li> <li>2. Notice: Companies should provide notice to each user whose content is taken down or account is suspended about the reasons for the removal or suspension. Companies should provide detailed guidance about what is prohibited, along with an explanation of how automated detection is used across each category of content.</li> <li>3. Appeal: Companies should provide a meaningful opportunity for timely appeal of any content removal or account suspension. This should include human review with the opportunity to present additional information that will be considered during the review. Notification of the results and a statement of reasoning must be provided.</li> </ol>
<p><b>What else is required to make the model work? What is the necessary context?</b></p>	<p>In order for this model to work, there is an assumption that the provision of transparency will in turn lead to accountability. Whilst at an individual level, these principles ensure the correct process is followed once content has been observed, it is the reporting of these processes to the general population that is hoped to cause change. As it is a public-facing accountability scheme, there is a hope that consumers and civic bodies will act if correct process are not being followed or reported.</p>
<p><b>What were the processes and steps involved in implementing reporting?</b></p>	<p>Whilst there was no enforcement of transparency reporting, in creating the principles there was an extensive engagement of organisations, advocates and academic experts to allow for a wide coverage of inputs.</p>
<p><b>What are the risks? Are there any unintended consequences?</b></p>	<p>With these principles only being guidelines, the risk is that there will be minimal uptake by platforms and legislators.</p>
<p><b>What lessons can be learned from this model and be applied to Online Safety ?</b></p>	<p>The Santa Clara process involved those with the particular view of supporting the right to free expression online and did not include industry consultation. Thinking forward to how Ofcom can learn from this, there should be a focus on the broadest level of consultation and incorporating voluntary guidelines with enforced action.</p>

## The Transparency Reporting Toolkit: Content Takedown Reporting

**Status:** This is a recommendatory model and has little evidence of implementation.

**Principal mechanisms employed:**



Education  
(Industry)



Education  
(Consumers)



1-Way  
Reporting



Performance /  
Behavioural  
Accountability

**Summary:** Developed by thinktank New America, The Transparency Reporting Toolkit surveys a range of telecommunications companies reporting guiding best practices on how reporting can be improved in future.

### The Transparency Reporting Toolkit: Content Takedown Reporting<sup>194</sup>

#### General Overview

This toolkit surveys how 35 global internet and telecommunications companies report on six categories of content takedowns and offers a set of guiding best practices on how their reporting can be improved going forward, with a focus on making them clearer, more detailed, and more standardised across companies. It is the latest in a series of toolkits that started in 2016 when OTI and Harvard University's Berkman Klein Center for Internet & Society released the first Transparency Reporting Toolkit<sup>195</sup>, which focused on reporting about government demands for user data

#### What does transparency help with?

**Transparency mechanisms:** One-way occurrence reporting on content takedowns.

**What Transparency Helps with:** Transparency reporting on content takedowns in this instance is seen as critically important in holding companies accountable for their roles as gatekeepers of online speech, and helps the public identify where they think the companies are doing too much - or not enough - to address content issues on their platforms and networks.

#### Why and how is transparency effective? What is the intended audience for the transparency information? Do the benefits spill out beyond that audience?

**Why is transparency effective:** Transparency reporting educates consumers and legislators on the scale of issues and what is being done to resolve them. It is therefore the intention that in this instance transparency will help increase the accountability of platforms and allow the public and other stakeholders to judge whether companies are doing enough to address key social issues.

**How is transparency effective:** It is hoped that ensuring the correct, actionable, and understandable information is reported will ensure greater accountability. It is particularly important that different audiences can access and understand information as, if

<sup>194</sup> Singh & Bankston, 2018: 'The Transparency Reporting Toolkit: Content Takedown Reporting' <https://www.newamerica.org/oti/reports/transparency-reporting-toolkit-content-takedown-reporting/>

<sup>195</sup> Woolery, Budish and Bankston: 'The Transparency Reporting Toolkit - Guide and Template', <https://www.newamerica.org/oti/policy-papers/transparency-reporting-toolkit-reporting-guide-and-template/>

not, then the data presented will have little impact and the effectiveness of the transparency itself will be minimal. To address this, the creation of general best practices seeks to ensure all the relevant information is reported for the relevant audiences – ranging from simple for consumers through to meta-data for academics and legislators – and ensuring common taxonomy across platforms to allow for easy comparability.

**Who:** Transparency reporting, whilst benefiting consumers, also offers benefits to the companies themselves, including helping to build trust with their users and policymakers. For companies, this is an opportunity to highlight to users and policymakers the responses they have implemented and communicate the size and complexity of the problems they are addressing, as well as the impact they have had thus far.

**Why was transparency developed? Was it voluntary or mandated by Government?**

**Why was transparency developed?**

There is an acknowledgement that transparency currently exists, however this toolkit looks to provide guidance for how this can be developed to ensure maximum value for all stakeholders. There is a great deal of variance in the approaches, styles, and scope of data that companies are using in their reporting, making it hard to combine or compare data in a meaningful way.

**Voluntary or mandated?**

These recommendations, which are not legally binding, reflect an understanding of current best practices and seeks to help business develop their transparency reporting.

**What is the key information / data that is reported?**

- Reports should be specific to the type of demands received and provide specific numbers for each
- Reports should break down demands by country
- Reports should include the categories of objectionable content targeted by demands
- Reports should include the products targeted by demands
- The specific government agencies/ parties that submitted demands
- Specifying which laws pertain to specific demands
- Reporting on the number of accounts and items specified in demands
- The number of accounts and items impacted by demands
- How the company responds to demands

**What else is required to make the model work? What is the necessary context?**

Issuing regular reports on clearly and consistently delineated reporting periods as, regardless of the publication schedule a company adopts, there should be a clearly defined time period covered by each report, and subsequent reports should cover the same length of time to allow for easy comparison.

**What were the processes and steps involved in implementing reporting?**

Whilst the principles are there to act as guidelines, and therefore not forced implementation, there were extensive consultations with a wide range of companies and civil society experts alongside the work of the Ranking Digital Rights project at New America, which developed a broad set of indicators for judging how well companies perform when it comes to protecting human rights.

**What are the risks? Are there any unintended consequences?**

No unintended consequences observed in the literature. As has been flagged in several interviews however, voluntary initiatives for transparency seldom work.

**What lessons can be learned from this model and be applied to Online Safety?**

The extensive consultation period, considering a broad range of opinions, is crucial to creating transparency for online safety. Clear guidelines to platforms for the information that they should provide is critical to avoid grey areas and room for inconsistencies.

## Summary of Principal Mechanisms Employed

Case Studies	Principal Mechanisms Employed													
	Engagement	Exposure	Education (Industry)	Education (Consumers)	Empowerment	Performance and behavioural accountability	Investment	Confidentiality	Regulator Transparency	Audit	One-way occurrence reporting	Two-way reporting and feedback	Performance reporting	Researcher Access
Advertising	✓	✓		✓	✓	✓			✓			✓		
Asset & Wealth Management			✓	✓	✓	✓	✓				✓	✓		
Aviation	✓		✓		✓			✓				✓		
Data Services (The 5 Safes Framework)										✓				✓
European Pollutant Release and Transfer Register (E-PRTR)				✓	✓	✓					✓			
Fashion Textiles & Clothing				✓	✓	✓					✓		✓	
Food & Drink	✓			✓	✓						✓			
Media, Entertainment & Film	✓			✓	✓						✓			
Financial Services				✓	✓	✓						✓		
UK Modern Slavery Act				✓		✓					✓		✓	
Ofgem Complaints Reporting (Utilities)				✓		✓							✓	
EU Mandatory Non Financial Disclosures				✓		✓				✓	✓			
The Santa Clara Principles				✓		✓					✓			
The Transparency Reporting Toolkit: Content Takedown Reporting			✓	✓		✓					✓			
Australian - Online Safety Act			✓			✓					✓		✓	
Australian Communications and Media Authority (ACMA)						✓			✓			✓		
French AVIA Law						✓					✓			
German Network Enforcement Act (NetzDG)				✓	✓	✓					✓			
EU General Data Protection Regulation (GDPR)				✓	✓	✓								
EU Digital Services Act						✓					✓			

Figure 4: Summary of the principal transparency mechanisms employed in each regulated sector

# Opportunities for Ofcom

## Looking Forward - Adaptive Regulation

The speed of innovation in information and communications technology poses a challenge for online safety regulation in driving the continuous iteration of regulatory approaches in order for regulation to be able to evolve alongside the very technology it seeks to control.

To meet this challenge, the development of a transparency regime could consider placing adaptability and flexibility core to its design and functionality. Through focusing on the capability to pre-empt regulatory risks and challenges emerging from the sector and the flexibility to frequently update the focus of transparency initiatives accordingly, a regulator could iteratively adjust their regime to ensure disclosure continues to meet changing demands for transparency in society as technology trends, ebb, and flow.

**INTERVIEW 13** Having an adaptable framework through which you can demand transparency from platforms is something definitely to be considered.

To achieve this, a regulator may seek to consider their own agility and how their existing organisational structures and processes impact their flexibility to adapt to meet new regulatory demands at pace. On reflection of challenges to transparency in online safety regulation, the same factors that pose significant challenge (international scale, pace of innovation, skills & information asymmetry) also represent opportunities for a regulator in seeking international collaboration, developing adaptive and flexible reporting frameworks, and upskilling technical and policy expertise.

Additionally, to reinforce the validity of regulatory frameworks that may often adapt and evolve, a regulator may consider the value of pre-emptively gaming their initiatives prior to implementation, a process of stress-testing in seeking to identify flaws, gaps or risks in reporting requirements that could undermine how meaningful the outputs of a given transparency regime may be or any unintended impacts that could occur.

**INTERVIEW 19** It's important to game and think through the ways in which you're asking platforms for disclose certain metrics. Think about how this might be strategized by a company and the ways in which we can prevent them from gaming these metrics.

Whilst it is clear Online Safety regulation poses unique challenges, recognition of the value of adaptable regulation is shared across many sectors. For example, in automotive regulation, the Centre for Connected and Autonomous Vehicles has created an ‘outcome focused’ approach whereby regulation can be updated as automotive technology evolves<sup>196</sup>, and both the Civil Aviation Authority and the Financial Conduct Authority operate innovation hubs to promote two-way conversations and transparency between regulators and innovation in the market to pre-empt where regulation may need to adapt to new technologies. In this sense, a flexible approach acts as an enabler for a regulator to collaborate with industry on innovation. As Jonathon Evans of the Better Regulation Executive, BEIS, discusses in ‘A brave new world: agile regulation to unleash innovation’ (2021)<sup>197</sup>, ‘regulation is the gatekeeper to the market’ but it should enhance not hinder innovation. Having this adaptive and flexible approach to regulation, underpinned by organisational agility as a regulator, could allow transparency regulation to evolve in collaboration with innovation, and its associated opportunities and risks, rather than acting as a costly barrier to it.

## Themes for exploration in transparency regulation

The following section outlines 12 themes emerging primarily from the interview conversations that pose potential opportunities for further exploration. It should be noted, the following does not represent an exhaustive list of suggestions that arose throughout the research; however, it highlights some commonality in the conversations with participants and their specific focuses and interests.

**Adaptive Regulation**<sup>198</sup> – The potential value of an adaptive, flexible, and dynamic approach to online safety regulation that allows for ease of iteration.

**Improving Algorithmic Transparency**<sup>199</sup> – Improving transparency in usage of algorithms for a range of purposes including for content curation, amplification (or content blocking e.g. ‘shadow banning’<sup>200</sup>), and within content moderation.

**Innovation Sandbox**<sup>201</sup> – A controlled testing environment to promote two-way transparency between the regulator and emerging technology companies to promote innovation in a safe and compliant setting.

**International Collaboration & A ‘Beacon of hope’**<sup>202</sup> – The opportunity to collaborate internationally in standardised approaches for improving online safety (including common taxonomies) that could raise the bar for regulation and platform safety internationally beyond the UK, described as a ‘Beacon of hope’ during interview.

---

<sup>196</sup> Ibid.

<sup>197</sup> Jonathon Evans, Better Regulation Executive, BEIS: ‘A brave new world: agile regulation to unleash innovation’ <https://civilservice.blog.gov.uk/2021/01/20/a-brave-new-world-agile-regulation-to-unleash-innovation/>

<sup>198</sup> Interviews 1,3,9,11,13,14,19,20 , 2021

<sup>199</sup> Interviews 4,5,7,8,10,16,17,18,19,20,21,22, 2021

<sup>200</sup> *Shadow banning*: A method of curtailing user content appearances (posts etc.) on a social media feed without informing the user of any such intervention.

<sup>201</sup> Interview 9, Interview 21, 2021

<sup>202</sup> Interviews 2,3,13,14,21,22, 2021

**Moderation Process Transparency**<sup>203</sup> – Improving transparency in the entire content moderation process, including an organisations logic and reasoning behind setting standards (e.g. community standards), the range of content intervention methods and explanation of their proportionate usage, detail in the use of moderation algorithms, detail in the use of human moderators (considering factors such as welfare, training levels, linguistic capabilities), the relationship between automated and human moderation and details of appeals processes & policies for users.

**Obtainable & Realistic Transparency**<sup>204</sup> – Considering how transparency standards can be applied to different sized platforms. The threshold to apply different transparency requirements may apply proportionately to the level of harms experienced on a platform rather than blanketly to all platforms that sit above a certain threshold on their number of users i.e., big platforms are better resourced and could have a lot of users but low levels of harm, whilst a smaller platform could have less users, less resources but be very harmful.

**Pro-active gamification**<sup>205</sup> – Actively seeking to pre-empt the gaming of mandated transparency initiatives via testing how metrics and policies could be gamed to obfuscate malpractice by different organisations and stakeholders.

**Procedural or process-based regulation**<sup>206</sup> – Using regulation to improve transparency of an organisation’s policies and processes in improving online safety, this could involve using standardised ‘tech-agnostic’ frameworks with metrics that are applicable to a range of platforms e.g., measuring the maturity of risk assessment processes.

**Risk assessment**<sup>207</sup> – Assessing platforms’ standards and processes for risk-assessing new features to measure potential safety concerns across different user groups (e.g., children, marginalized groups).

**Skills Asymmetry**<sup>208</sup> – The idea that a lack of understanding and collaboration on managing the impact of platforms technical functionalities (e.g., algorithms,) and the information asymmetry that propagates by this lack of understanding, can be partially attributed to an asymmetry of skillsets between technical vs non-technical expertise.

**Tailoring specific transparency requirements to harms**<sup>209</sup> – Understanding the need for specificity in which methods of transparency (and enforcement of transparency) are necessary and proportionate to tackle different harms and the potential requirement for transparency and enforcement methods to differ in response to specific harms.

**Transparency beyond platforms**<sup>210</sup> – Transparency beyond tech platforms, including in law enforcement takedown requests and the processes behind moderation databases (e.g., to prevent human rights abuses being masked).

---

<sup>203</sup> Interviews 2,7,8,14,16,19,20, 2021

<sup>204</sup> Interviews 1,2,3,6,7,8,10,12,13,14,16,17,18,19,20, 2021

<sup>205</sup> Interview 1, Interview 19, 2021

<sup>206</sup> Interviews 2, 8, 16, 19, 2021

<sup>207</sup> Interview 1, Interview 2, Interview 21, 2021

<sup>208</sup> Interview 22, 2021

<sup>209</sup> Interview 1, Interview 3, 2021

<sup>210</sup> Interview 3, Interview 10, Interview 20, 2021

# Annex

## Transparency Concept to Implementation Process Diagram

The following diagram was created during the project to help ideate the process that may be considered when taking a new transparency initiative from concept through to implementation. It has been included for reference.



## Defining Trust, Transparency and Accountability

Term	Definitions
<b>Trust</b>	<ol style="list-style-type: none"> <li>1 Assured reliance on the character, ability, strength, or truth of someone or something (Merriam-Webster)<sup>211</sup>.</li> <li>2 To rely on the truthfulness or accuracy of [something] (Merriam-Webster)<sup>212</sup>.</li> </ol>
<b>Transparency</b>	<ol style="list-style-type: none"> <li>1 The quality or state of being transparent (Merriam-Webster)<sup>213</sup>.</li> <li>2 Something transparent, <i>especially</i> a picture (as on film) viewed by light shining through it or by projection (Merriam-Webster)<sup>214</sup>.</li> <li>3 <i>In Business</i>: transparency is the process of being open, honest, and straightforward about various company operations (Forbes, 2019)<sup>215</sup></li> </ol>

<sup>211</sup> Merriam-Webster: 'Trust' <https://www.merriam-webster.com/dictionary/trust>

<sup>212</sup> Merriam-Webster: 'Trust' <https://www.merriam-webster.com/dictionary/trust>

<sup>213</sup> Merriam-Webster: 'Transparency' <https://www.merriam-webster.com/dictionary/transparency>

<sup>214</sup> Merriam-Webster: 'Transparency' <https://www.merriam-webster.com/dictionary/transparency>

<sup>215</sup> Forbes, 2019: 'Transparency In Business: 5 Ways to Build Trust'

<https://www.forbes.com/sites/mikekappel/2019/04/03/transparency-in-business-5-ways-to-build-trust/?sh=146bdcc86149>

<p><b>Transparency (cont.)</b></p>	<p>4 <i>In Social Science</i>: capacity of outsiders to obtain valid and timely information about the activities of government or private organizations (Johnston)<sup>216</sup>.</p> <p>5 The open flow of information (Holzner in Park &amp; Blenkinsopp, 2011)<sup>217</sup>.</p> <p>6 <i>As described by participants in a Food Standards Agency transparency study – honesty</i><sup>218</sup>.</p>
<p><b>Accountability</b></p>	<p>1 The quality or state of being accountable <i>especially</i> an obligation or willingness to accept responsibility or to account for one’s actions (Merriam-Webster)<sup>219</sup>.</p>

## References

Access Now: ‘Transparency Reporting Index’ <https://www.accessnow.org/transparency-reporting-index/>

ACMA: ‘Compliance Priorities 2020-21’ <https://www.acma.gov.au/consultations/2019-11/compliance-priorities-2020-2021-consultation#outcome>

Allen & Overy, 2020 ‘Review of the Non-Financial Reporting Directive: towards an EU-wide ESG reporting standard’ <https://www.allenoverly.com/en-gb/global/news-and-insights/publications/review-of-the-non-financial-reporting-directive>

Ananny & Crawford, 2018: ‘Seeing without knowing: Limitations of the transparency ideal and its application to algorithmic accountability’

ASA: ‘About ASA and CAP’ <https://www.asa.org.uk/about-asa-and-cap.html>

ASA: ‘Ad Alerts’ <https://www.asa.org.uk/advice-and-resources/resource-library/ad-alerts.html>

ASA: ‘Influencers’ guide to making clear that ads are ads’ <https://www.asa.org.uk/resource/influencers-guide.html>

ASA: ‘National ad campaign’ <https://www.asa.org.uk/about-asa-and-cap/national-ad-campaign.html>

ASA: ‘Non-compliant online advertisers’ <https://www.asa.org.uk/codes-and-rulings/non-compliant-online-advertisers.html>

ASA: ‘Our History’ <https://www.asa.org.uk/about-asa-and-cap.html>

ASA: ‘Recognising ads: Social media and influencer marketing’ <https://www.asa.org.uk/advice-online/recognising-ads-social-media.html#3>

ASA: ‘Rulings’ <https://www.asa.org.uk/codes-and-rulings/rulings.html>

ASA: ‘Transparency’ <https://www.asa.org.uk/transparency.html>

<sup>216</sup> Johnson in Britannica: ‘Transparency’ <https://www.britannica.com/topic/transparency-government>

<sup>217</sup> Park & Blenkinsopp, 2011: ‘The roles of transparency and trust in the relationship between corruption and citizen satisfaction’ <https://journals.sagepub.com/doi/pdf/10.1177/0020852311399230>

<sup>218</sup> Food Standards Agency Transparency: Understanding public views and priorities <https://www.food.gov.uk/sites/default/files/media/document/fsatransparencyreport.pdf>

<sup>219</sup> Merriam-Webster: ‘Accountability’ <https://www.merriam-webster.com/dictionary/accountability#:~:text=%3A%20the%20quality%20or%20state%20of,actions%20public%20officials%20ackinq%20accountability>

ASA-CAP: 'Annual Report 2019' <https://www.asa.org.uk/uploads/assets/5d2b832d-8d12-4f28-9e9f5df54f7b93c1/809de510-6285-4c57-9c904b1c91f6d0ec/ASA-CAP-2019-Annual-Report.pdf>

ASA-CAP-BCAP: 'How we consult' <https://www.asa.org.uk/about-asa-and-cap/the-work-we-do/how-we-consult.html>

Baldwin, 1999: 'Understanding Regulation: Theory, Strategy and Practice'

BBC: 'Kellogg's gives in on government's 'traffic light' labels' <https://www.bbc.co.uk/news/business-46373342>

BBFC: 'About us' <https://www.bbfc.co.uk/about-us>

BBFC: 'BBFC History' <https://www.bbfc.co.uk/education/university-students/bbfc-history>

BBFC: 'Our Mission' <https://www.bbfc.co.uk/about-us/who-we-are/our-mission>

BBFC: 'Survey into family viewing decisions September 2020' <https://darkroom.bbfc.co.uk/original/a9eb359c3c6c06f60371cba1b1e28587:adb5b3d7016e0e66093f1c7544572b02/survey-into-family-viewing-decisions-sept-2020.pdf>

Birkinshaw, 2006: 'Transparency as a Human Right' [Abstract] <https://britishacademy.universitypressscholarship.com/view/10.5871/bacad/9780197263839.001.0001/upso-9780197263839-chapter-3>

BlackRock, BlackRock Earns Straight A's for Transparency by Principles for Responsible Investment, <https://www.blackrock.com/corporate/newsroom/press-releases/article/corporate-one/press-releases/blackrock-earns-straight-a-for-transparency>

BlackRock: '#BlackRock Sustainability Accounting Standards Board Disclosure' <https://www.blackrock.com/corporate/literature/continuous-disclosure-and-important-information/blackrock-2019-sasb-disclosure.pdf>

BlackRock: 'BlackRock Earns Straight A's for Transparency by Principles for Responsible Investment' <https://www.blackrock.com/corporate/newsroom/press-releases/article/corporate-one/press-releases/blackrock-earns-straight-a-for-transparency>

BlackRock: 'Delivering on our commitment to sustainability and stewardship' <https://www.blackrock.com/institutions/en-gb/solutions/sustainable-investing/our-commitment-to-stewardship>

BlackRock: 'Supplier Code of Conduct & Ethics' <https://www.blackrock.com/uk/individual/literature/publication/blackrock-supplier-code-of-conduct-and-ethics.pdf>

Bloomfield and LeBaron, 2018, 'The UK Modern Slavery Act: Transparency through Disclosure in Global Governance', <https://www.e-ir.info/pdf/75729>

Boom, 2011 Price Intransparency, Consumer Decision Making and European Consumer Law' <https://link.springer.com/content/pdf/10.1007/s10603-011-9163-8.pdf>

Brandies, 1914: 'Chapter V- What Publicity Can Do' <http://louisville.edu/law/library/special-collections/the-louis-d.-brandeis-collection/other-peoples-money-chapter-v>

CAA: 'CAA Innovation Hub' <https://www.caa.co.uk/Our-work/The-CAA-Innovation-Hub/>

CAA: 'Just Culture for the Team' <https://caainternational.com/course/just-culture-team/>

CAA: 'Occurrence Reporting' <https://www.caa.co.uk/Our-work/Make-a-report-or-complaint/MOR/Occurrence-reporting/>

Children's Research Network, 2020: 'Do children have rights in the online world and can they be enforced? An analysis of the Council of Europe's 'Guidelines on Children's Rights in the Digital Environment' from an Irish perspective.' <https://www.childrensresearchnetwork.org/knowledge/resources/article-child-rights-online-world>

Craig, G. (2016), 'The UK's Modern Slavery Legislation: An Early Assessment of Progress', <https://www.cogitatiopress.com/socialinclusion/article/view/833/833>

Cutler, A. C. and Dietz, T. (eds) (2017) 'The Politics of Private Transnational Governance by Contract'. New York: Routledge

Deakins, O. 'UK Government Proposes Tighter Reporting Requirements Under Modern Slavery Act, but No Action on Enforcement Measures' (2020) <https://www.lexology.com/library/detail.aspx?g=74958926-7f47-4190-a5d7-f49242f0c635>

Delivering a Duty of Care, NSPCC (2021) <https://www.nspcc.org.uk/globalassets/documents/online-safety/delivering-a-duty-of-care.pdf>

DfBEIS, 'Consultation on requiring mandatory climate-related financial disclosures by publicly quoted companies, large private companies and Limited Liability Partnerships (LLPs)'. [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/972422/Consultation\\_on\\_BEIS\\_mandatory\\_climate-related\\_disclosure\\_requirements.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/972422/Consultation_on_BEIS_mandatory_climate-related_disclosure_requirements.pdf)

Diabetes UK Public Views on Food Labelling – January 2018 [https://2sjiwunnql41ia7ki31qqub1-wpengine.netdna-ssl.com/wp-content/uploads/2018/03/J303701\\_Diabetes-UK\\_Food-Labelling\\_Data-Tables\\_January-2018.pdf](https://2sjiwunnql41ia7ki31qqub1-wpengine.netdna-ssl.com/wp-content/uploads/2018/03/J303701_Diabetes-UK_Food-Labelling_Data-Tables_January-2018.pdf)

Disclosure in Global Governance, <https://www.e-ir.info/pdf/75729>

Douek, 2020 'The Rise of Content Cartels' <https://knightcolumbia.org/content/the-rise-of-content-cartels>

Echikson & Knodt, 2018: 'Germany's NetzDG: A key test for combatting online hate' [https://www.counterextremism.com/sites/default/files/CEP-CEPS\\_Germany%27s%20NetzDG\\_020119.pdf](https://www.counterextremism.com/sites/default/files/CEP-CEPS_Germany%27s%20NetzDG_020119.pdf)

EDRI: 'French Avia law declared unconstitutional: what does this teach us at EU level?', <https://edri.org/our-work/french-avia-law-declared-unconstitutional-what-does-this-teach-us-at-eu-level/>

Elson, 'Modern slavery transparency failings to lead to UK fines', <https://www.pinsentmasons.com/out-law/news/modern-slavery-transparency-failings-to-lead-to-uk-fines>

Environmental Audit Committee: Fixing Fashion: clothing consumption and sustainability <https://publications.parliament.uk/pa/cm201719/cmselect/cmenvaud/1952/full-report.html>

E-PRTR: 'What does the E-PRTR cover?' <https://prtr.eea.europa.eu/#/static?cont=about>

E-PRTR: <https://prtr.eea.europa.eu/#/home>

eSafety Commissioner: 'Safety by Design' <https://www.esafety.gov.au/about-us/safety-by-design>

eSafety Commissioner: 'Social Media Tier Scheme', <https://www.esafety.gov.au/about-us/who-we-are/social-media-tier-scheme>

eSafety Commissioner: 'The actions we can take', <https://www.esafety.gov.au/report/illegal-harmful-content/the-actions-we-can-take>

eSafety Commissioner: 'What we can investigate', <https://www.esafety.gov.au/report/illegal-harmful-content/what-we-can-investigate>

eSafety, Our legislative functions, <https://www.esafety.gov.au/about-us/who-we-are/our-legislative-functions>

Etzioni, 2010: 'Is Transparency the Best Disinfectant?' <https://www.uv.es/sasece/docum/mayo-sept2010/295Transparency.pdf>

EU, 2014: Directive 2014/95/EU of the European Parliament and of the Council <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32014L0095&from=EN>

EU: Corporate sustainability reporting [https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting\\_en#disclosed](https://ec.europa.eu/info/business-economy-euro/company-reporting-and-auditing/company-reporting/corporate-sustainability-reporting_en#disclosed)

European Commission 'EFIT evaluation of Regulation (EC) No 166/2006 concerning the establishment of a European Pollutant Release and Transfer Register (E-PRTR)' <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52017SC0710&from=en>

European Commission 'Guidance Document for the implementation of the European PRTR' [https://ec.europa.eu/environment/industry/stationary/e-prtr/pdf/en\\_prtr.pdf](https://ec.europa.eu/environment/industry/stationary/e-prtr/pdf/en_prtr.pdf)

European Commission, 2013: 'Proposal for a Directive of the European Parliament and of the Council amending Council Directives 78/660/EEC and 83/349/EEC as regards disclosure of non-financial and diversity information by certain large companies and groups' <https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52013PC0207&from=EN>

European Commission, 2017: Information from European Union Institutions, Bodies, Offices and Agencies 'Guidelines on non-financial reporting' [https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52017XC0705\(01\)&from=EN](https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:52017XC0705(01)&from=EN)

European Commission: 'REFIT evaluation of Regulation (EC) No 166/2006 concerning the establishment of a European Pollutants Release and Transfer Register (E-PRTR)' <https://eur-lex.europa.eu/legal-content/EN/TXT/HTML/?uri=CELEX:52017SC0710&from=EN>

European Commission: 'The Aarhus Convention' <https://ec.europa.eu/environment/aarhus/>

European Environment Agency: 'European Pollutant Release and Transfer Register- a harmonised PRTR across Europe' [https://unece.org/fileadmin/DAM/env/pp/PRTR%20Bureau/GRT2013-Item4-1-EU-EEA\\_E-PRTR.pdf](https://unece.org/fileadmin/DAM/env/pp/PRTR%20Bureau/GRT2013-Item4-1-EU-EEA_E-PRTR.pdf)

European Environment Agency: 'European Pollutant Release and Transfer Register- a harmonised PRTR across Europe' [https://unece.org/fileadmin/DAM/env/pp/PRTR%20Bureau/GRT2013-Item4-1-EU-EEA\\_E-PRTR.pdf](https://unece.org/fileadmin/DAM/env/pp/PRTR%20Bureau/GRT2013-Item4-1-EU-EEA_E-PRTR.pdf)

European Parliament, 2021: 'Corporate due diligence and corporate accountability: European Parliament resolution of 10 March 2021 with recommendations to the Commission on corporate due diligence and corporate accountability' [https://www.europarl.europa.eu/doceo/document/TA-9-2021-0073\\_EN.pdf](https://www.europarl.europa.eu/doceo/document/TA-9-2021-0073_EN.pdf)

Fashion Revolution 'Why Transparency Matters' <https://www.fashionrevolution.org/about/transparency/>

Fashion Revolution: 'Fashion Revolution Week' <https://www.fashionrevolution.org/>

Fashion Revolution: 'The Fashion Transparency Index 2020: What do we mean by transparency?' <https://www.fashionrevolution.org/about/transparency/>

Fashion Revolution: 'The Fashion Transparency Index 2020' <https://www.fashionrevolution.org/about/transparency/>

FCA : 'Transparency as a Regulatory Tool' <https://www.fca.org.uk/publication/discussion/fsa-dp08-03.pdf>

FCA: 'Transparency as a Regulatory Tool' (2008), <https://www.fca.org.uk/publication/discussion/fsa-dp08-03.pdf>

FCA: 'FCA Innovation – fintech, regtech and innovative businesses' <https://www.fca.org.uk/firms/innovation>

FCA: 'FCA Transparency Framework' <https://www.fca.org.uk/publication/feedback/fca-transparency-framework.pdf>

Food Standards Agency, 'Front of pack Nutrition Labelling: Joint Response to Consultation' [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/216997/response-nutrition-labelling-consultation.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/216997/response-nutrition-labelling-consultation.pdf)

Food Standards Agency: 'Building on the success of front-of pack nutrition labelling in the UK: a public consultation' [https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/905096/front-of-pack-labelling-consultation-document-english.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/905096/front-of-pack-labelling-consultation-document-english.pdf)

Food Standards Agency: 'Guide to creating a front of pack (FoP) nutrition label for pre-packed products sold through retail outlets' [https://www.food.gov.uk/sites/default/files/media/document/fop-guidance\\_0.pdf](https://www.food.gov.uk/sites/default/files/media/document/fop-guidance_0.pdf)

Food Standards Agency: 'Guide to creating a front of pack (FoP) nutrition label for pre-packed products sold through retail outlets' [https://www.food.gov.uk/sites/default/files/media/document/fop-guidance\\_0.pdf](https://www.food.gov.uk/sites/default/files/media/document/fop-guidance_0.pdf)

Food Standards Agency: 'Transparency: Understanding public views and priorities' <https://www.food.gov.uk/sites/default/files/media/document/fsatransparencyreport.pdf>

Forbes, 2019: 'Transparency In Business: 5 Ways to Build Trust' <https://www.forbes.com/sites/mikekappel/2019/04/03/transparency-in-business-5-ways-to-build-trust/?sh=146bdcc86149>

FSA: 'Discussion Paper DP13/1 Transparency' <https://www.fca.org.uk/publication/discussion/fsa-dp13-01.pdf>

G.K Chesterton, 1911: 'The Sign of a Broken Sword'

Gaventa & McGee, 2013: 'The Impact of Transparency and Accountability Initiatives'

Global Network Initiative: 'The GNI Principles' <https://globalnetworkinitiative.org/gni-principles/>

Google: 'Transparency Report' [https://transparencyreport.google.com/about?hl=en\\_GB](https://transparencyreport.google.com/about?hl=en_GB)

GOV.UK: 'The Eatwell Guide', <https://www.gov.uk/government/publications/the-eatwell-guide>

Heldt, 2019, 'Reading between the lines and the numbers: an analysis of the first NetzDG reports' <https://policyreview.info/articles/analysis/reading-between-lines-and-numbers-analysis-first-netzdg-reports>

Hersey et al. (2009) 'Effects of front-of-package and shelf nutrition labelling systems on consumers' <https://pubmed.ncbi.nlm.nih.gov/23282247/>

Hilsdon, 2018, 'The impact of the Task Force on Climate-related Financial Disclosures', <https://www.reutersevents.com/sustainability/impact-task-force-climate-related-financial-disclosures>  
[https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment\\_data/file/905096/front-of-pack-labelling-consultation-document-english.pdf](https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/905096/front-of-pack-labelling-consultation-document-english.pdf)

Human Rights Watch, 2020: 'Joint Letter to New Executive Director, Global Internet Forum to Counter Terrorism' <https://www.hrw.org/news/2020/07/30/joint-letter-new-executive-director-global-internet-forum-counter-terrorism>

ICO: 'Right to be informed', <https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/individual-rights/right-to-be-informed/>

ICO: 'Some Basic Concepts', <https://ico.org.uk/for-organisations/guide-to-data-protection/introduction-to-data-protection/some-basic-concepts/>

Ikonen et al. (2019) Consumer effects of front-of-package nutrition labelling: an interdisciplinary meta-analysis <https://link.springer.com/article/10.1007/s11747-019-00663-9>

Independent Review into the boohoo Groups PLC'S Leicester supply chain <https://www.boohooplc.com/sites/boohoo-corp/files/final-report-open-version-24.9.2020.pdf>

Interim Report on the Sustainability of the Fashion Industry [https://publications.parliament.uk/pa/cm201719/cmselect/cmenvaud/1148/114804.htm#\\_idTextAnchor004](https://publications.parliament.uk/pa/cm201719/cmselect/cmenvaud/1148/114804.htm#_idTextAnchor004)

Intersoft Consulting: 'Article 1 GDPR' <https://gdpr-info.eu/art-1-gdpr/>

Johnson in Britannica: 'Transparency' <https://www.britannica.com/topic/transparency-government>

Jonathon Evans, Better Regulation Executive, BEIS: 'A brave new world: agile regulation to unleash innovation' <https://civilservice.blog.gov.uk/2021/01/20/a-brave-new-world-agile-regulation-to-unleash-innovation/>

Jones et al. (2019), 'Front-of-pack nutrition labelling to promote healthier diets: current practice and opportunities to strengthen regulation worldwide' <https://gh.bmj.com/content/bmjgh/4/6/e001882.full.pdf>

Julia & Herberg (2017) Nutri-Score: Evidence of the effectiveness of the French front-of-pack nutrition label [https://www.ernaehrungs-umschau.de/fileadmin/Ernaehrungs-Umschau/pdfs/pdf\\_2017/12\\_17/EU12\\_2017\\_WuF\\_Nutriscore\\_englisch.pdf](https://www.ernaehrungs-umschau.de/fileadmin/Ernaehrungs-Umschau/pdfs/pdf_2017/12_17/EU12_2017_WuF_Nutriscore_englisch.pdf)

Kim et al. (2020): 'Is Honesty the Best Policy? Examining the Role of Price and Production Transparency in Fashion Marketing' [https://www.mdpi.com/2071-1050/12/17/6800?type=check\\_update&version=1](https://www.mdpi.com/2071-1050/12/17/6800?type=check_update&version=1) LeBaron, G., & Rühmkorf, A. (2017a). 'Steering CSR Through Home State Regulation: A Comparison of the Impact of the UK Bribery Act and Modern Slavery Act on Global Supply Chain Governance'. *Global Policy*, 8, 15–28.

Littlechild, S. (2020): 'An Overall Customer Satisfaction score for GB energy suppliers' <https://www.repository.cam.ac.uk/bitstream/handle/1810/314746/cwpe2090.pdf?sequence=1&isAllowed=y>

Livingstone, 2021: 'Children's rights apply in the digital world!' <https://blogs.lse.ac.uk/mediase/2021/02/04/childrens-rights-apply-in-the-digital-world/>

Meijer in Mabillard & Zumofen, 2016: 'The complex relationship between transparency and accountability: A synthesis and contribution to existing frameworks' [https://www.researchgate.net/publication/304070388\\_The\\_complex\\_relationship\\_between\\_transparency\\_and\\_accountability\\_A\\_synthesis\\_and\\_contribution\\_to\\_existing\\_frameworks](https://www.researchgate.net/publication/304070388_The_complex_relationship_between_transparency_and_accountability_A_synthesis_and_contribution_to_existing_frameworks)

Merriam-Webster: 'Accountability' <https://www.merriam-webster.com/dictionary/accountability#:~:text=%3A%20the%20quality%20or%20state%20of,actions%20public%20officials%20lacking%20accountability>

Merriam-Webster: 'Transparency' <https://www.merriam-webster.com/dictionary/transparency>

Merriam-Webster: 'Trust' <https://www.merriam-webster.com/dictionary/trust>

Michener, 2018: 'Gauging the Impact of Transparency Policies'

Møllgaard & Overgaard, 2001: 'Market transparency competition policy' <https://www.econstor.eu/bitstream/10419/208446/1/cbs-wp2001-06.pdf>

Morgan, 2018: 'The Transparency Challenge: Making children aware of their data protection rights and the risks online' <https://www.dataprotection.ie/sites/default/files/uploads/2019-02/TransparencyChallenge.pdf>

Network Enforcement Act (Netzdurchsetzungsgesetz, NetzDG) <https://germanlawarchive.iuscomp.org/?p=1245>

Network for Business Sustainability: How Nike Turned a Supply Chain Crisis into Opportunity <https://www.nbs.net/articles/just-do-it-how-nike-turned-a-supply-chain-crisis-into-opportunity>

Office for National Statistics: 'The Five Safes- Data Privacy at ONS' <https://blog.ons.gov.uk/2017/01/27/the-five-safes-data-privacy-at-ons/>

Ofgem: 'Compare Supplier Performance Complaints', <https://www.ofgem.gov.uk/consumers/energy-supplier-comparison-data/compare-supplier-performance-complaints>

Ofgem: 'Guidance to submitting customer complaints data', <https://www.ofgem.gov.uk/publications-and-updates/guidance-submitting-customer-complaints-data>

Ofgem: 'Transparency Consultation', [https://www.ofgem.gov.uk/sites/default/files/docs/2014/10/transparency\\_consultation\\_0.pdf](https://www.ofgem.gov.uk/sites/default/files/docs/2014/10/transparency_consultation_0.pdf)

Ormerod in Ritchie, 2008: Secure access to confidential microdata: four years of the Virtual Microdata Laboratory <https://link.springer.com/content/pdf/10.1057/elmr.2008.73.pdf>

PA Consulting: 'Digital Regulatory Reporting' <https://www2.paconsulting.com/rs/526-HZE-833/images/DRR-Report-Sept-2020.pdf>

Park & Blenkinsopp, 2011: 'The roles of transparency and trust in the relationship between corruption and citizen satisfaction' <https://journals.sagepub.com/doi/pdf/10.1177/0020852311399230>

PWC, The Modern Slavery Act, <https://www.pwc.co.uk/services/sustainability-climate-change/sustainability-strategy-support/supply-chain/the-modern-slavery-act.html>

Reuters: 'Boohoo cuts number of UK suppliers' <https://www.reuters.com/article/us-boohoo-suppliers-idUSKBN2BH0R4> Ritchie, 2008: Secure access to confidential microdata: four years of the Virtual Microdata Laboratory <https://link.springer.com/content/pdf/10.1057/elmr.2008.73.pdf>

Ruche, D. 2020, Green investing 'is definitely not going to work', says ex-BlackRock executive, <https://www.theguardian.com/business/2021/mar/30/tariq-fancy-environmentally-friendly-green-investing>

SASB (Sustainability Accounting Standards Board): 'About' <https://www.sasb.org/about/>

Schuler and Znaty (2020): 'New law to fight online hate speech to reshape notice, take down and liability rules in France', <https://www.taylorwessing.com/en/insights-and-events/insights/2020/05/new-law-to-fight-online-hate-speech-in-france>

Singh & Bankston, 2018: 'The Transparency Reporting Toolkit: Content Takedown Reporting' <https://www.newamerica.org/oti/reports/transparency-reporting-toolkit-content-takedown-reporting/>

Souter, 2021: 'Inside the Digital Society': Children's rights in the digital society' <https://blogs.lse.ac.uk/mediase/2021/04/12/inside-the-digital-society-childrens-rights-in-the-digital-society/>

Sustainable Development Commission: 'Lost in Transmission?' <https://research-repository.st-andrews.ac.uk/bitstream/handle/10023/2344/sdc-2007-lost-in-transmission.pdf?sequence=1&isAllowed=y>

Suzor, Nicolas P. et al. What Do We Mean When We Talk About Transparency? Toward Meaningful Transparency in Commercial Content Moderation. International Journal of Communication, [S.I.], v. 13, p. 18, mar. 2019..

The European Parliament and Council of the European Union: 'Regulation (EC) No 166/2006 of the European Parliament and of the Council' <https://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2006:033:0001:0017:EN:PDF#page=8>

The European Pollutant Release and Transfer Register (E-PRTR): 'What is E-PRTR?' <https://ec.europa.eu/environment/industry/stationary/eper/legislation.htm>

The Fashion Transparency Index 2020 <https://www.fashionrevolution.org/about/transparency/>

The Guardian: 'Boohoo shares bounce back after pledge to improve factory conditions' <https://www.theguardian.com/fashion/2020/jul/09/boohoo-shares-bounce-back-after-pledge-to-improve-factory-conditions>

The Guardian: 'Italy claims 'traffic-light' labelling unfair on Mediterranean food' <https://www.theguardian.com/world/2013/oct/21/italy-traffic-light-food-labels-unfair>

The Santa Clara Principles: <https://santaclaraprinciples.org/>

The Times: 'Boohoo: fashion giant faces 'slavery' investigation' <https://www.thetimes.co.uk/article/boohoo-fashion-giant-faces-slavery-investigation-57s3hxcth>

Tony Blair Institute for Global Change, 2020: 'Online Harms: Bring in the Auditors' <https://institute.global/policy/online-harms-bring-auditors>

UK Data Service: 'Access to sensitive data for research: The 5 Safes' <http://blog.ukdataservice.ac.uk/access-to-sensitive-data-for-research-the-5-safes/>

UK Data Service: 'Five Safes' <https://www.ukdataservice.ac.uk/manage-data/legal-ethical/access-control/five-safes.aspx>

UK Data Service: Secure Lab 'Five Safes framework' <https://www.ukdataservice.ac.uk/use-data/secure-lab/security-philosophy>

UNECE: 'Introduction to the Kyiv Protocol on Pollutant Release and Transfer Registers' <https://unece.org/env/pp/former%20bodies-working-group-prtrs-introduction>

UNGC: 'Integrate the Principles for Responsible Investment' <https://www.unglobalcompact.org/take-action/action/responsible-investment>

Villas-Boas (2010) Can information cost affect consumer choice? Nutritional labels in a supermarket experiment [https://www.csus.edu/indiv/k/kieselk/Nutritional%20label%20experiment\\_IJIO.pdf](https://www.csus.edu/indiv/k/kieselk/Nutritional%20label%20experiment_IJIO.pdf)

Vogue 'Why We Should Be Asking #WhoMadeMyClothes? Before Every Purchase' <https://www.vogue.co.uk/article/who-made-my-clothes>

Wagner and Ferro (2020), 'Governance of Digitalization in Europe' [https://benwagner.org/wp-content/plugins/zotpress/lib/request/request.dl.php?api\\_user\\_id=2346531&dlkey=8KUHYZ5M&content\\_type=application/pdf](https://benwagner.org/wp-content/plugins/zotpress/lib/request/request.dl.php?api_user_id=2346531&dlkey=8KUHYZ5M&content_type=application/pdf)

Wagner and Jassen, 2021, "A first impression of regulatory powers in the Digital Services Act", <https://verfassungsblog.de/regulatory-powers-dsa/>

Wired: 'What is GDPR? The summary guide to GDPR compliance in the UK' <https://www.wired.co.uk/article/what-is-gdpr-uk-eu-legislation-compliance-summary-fines-2018>.

Wood, 2021: 'Online harms: why we need a systems-based approach towards internet regulation' <https://blogs.lse.ac.uk/mediase/2021/02/19/online-harms-why-we-need-a-systems-based-approach-towards-internet-regulation/>

Woolery, Budish and Bankston: 'The Transparency Reporting Toolkit - Guide and Template', <https://www.newamerica.org/oti/policy-papers/transparency-reporting-toolkit-reporting-guide-and-template/>

**Table 1**

Subject	Types of Transparency	Examples
About the FSA	Corporate Plans & Performance	<ul style="list-style-type: none"> <li>Annual Report</li> <li>Performance Account</li> <li>Annual review by statutory Panels</li> </ul>
	Policy	<ul style="list-style-type: none"> <li>Discussion and Consultation Papers</li> <li>Policy Statements</li> </ul>
	Process	<ul style="list-style-type: none"> <li>Regulatory Guides (e.g. the Enforcement Guide)</li> <li>'Doing business with the FSA' - website guides</li> </ul>
About firms/products/markets	Firms disclose (voluntary)	<ul style="list-style-type: none"> <li>Various</li> </ul>
	FSA requires firms to disclose (i.e. Regulated Information)	<ul style="list-style-type: none"> <li>Terms of Business</li> </ul>
	Public information - aggregated	<ul style="list-style-type: none"> <li>Financial Risk Outlook</li> </ul>
	Public information - firm-specific	<ul style="list-style-type: none"> <li>Type of intermediary (at proposal stage)</li> </ul>
	Information received by the FSA - aggregated	<ul style="list-style-type: none"> <li>Number of firms authorised by year and sector</li> </ul>
	Information received by the FSA - firm specific	<ul style="list-style-type: none"> <li>Comparative tables</li> <li>Persistency tables (to 2002)</li> <li>The Public Register</li> </ul>
	Information generated by the FSA - aggregated	<ul style="list-style-type: none"> <li>Feedback on thematic work</li> </ul>
	Information generated by the FSA - firm-specific	<ul style="list-style-type: none"> <li>Waivers granted</li> <li>Disciplinary record</li> <li>Undertakings under Unfair Contract Terms legislation</li> </ul>

Financial Services Authority: 'Transparency as a Regulatory Tool', Table 1: <https://www.fca.org.uk/publication/discussion/fsa-dp08-03.pdf>



## **About PA.**

We believe in the power of ingenuity to build a positive human future in a technology-driven world.

As strategies, technologies and innovation collide, we create opportunity from complexity.

Our diverse teams of experts combine innovative thinking and breakthrough use of technologies to progress further, faster. Our clients adapt and transform, and together we achieve enduring results.

An innovation and transformation consultancy, we are over 3,200 specialists in consumer, defence and security, energy and utilities, financial services, government, health and life sciences, manufacturing, and transport. Our people are strategists, innovators, designers, consultants, digital experts, scientists, engineers and technologists. We operate globally from offices across the UK, US, Europe, and the Nordics.

**PA. Bringing Ingenuity to Life.**

## **Corporate Headquarters**

10 Bressenden Place  
London  
SW1E 5DN  
+44 20 7730 9000

## **paconsulting.com**

This report has been prepared by PA Consulting Group on the basis of information supplied by the client, third parties (if appropriate) and that which is available in the public domain. No representation or warranty is given as to the achievability or reasonableness of future projections or the assumptions underlying them, targets, valuations, opinions, prospects or returns, if any, which have not been independently verified. Except where otherwise indicated, the report speaks as at the date indicated within the report.

## **All rights reserved**

**© PA Knowledge Limited 2021**

This report is confidential to the organisation named herein and may not be reproduced, stored in a retrieval system, or transmitted in any form or by any means, electronic, mechanical or otherwise, without the prior written permission of PA Consulting Group. In the event that you receive this document in error, you should return it to PA Consulting Group, 10 Bressenden Place, London, SW1E 5DN. PA Consulting Group accepts no liability whatsoever should an unauthorised recipient of this report act on its contents.