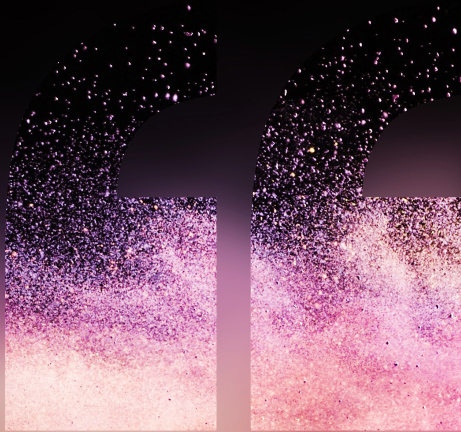
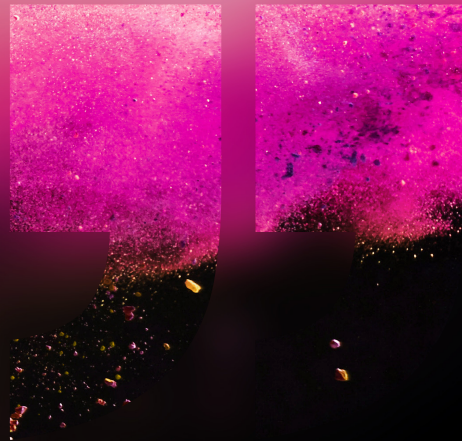


**C L I F F O R D**

**C H A N C E**



**MOVING FORWARD  
ON EXPLAINABLE AI –  
NEW GUIDANCE FROM  
THE UK ICO AND  
TURING INSTITUTE**



**– THOUGHT LEADERSHIP**

JULY 2020



## MOVING FORWARD ON EXPLAINABLE AI – NEW GUIDANCE FROM THE UK ICO AND TURING INSTITUTE

Artificial intelligence is being extensively used across sectors and around the world, but without comprehensive legal frameworks and appropriate governance and compliance programmes in place, regulators are starting to fill the gap.

### Key issues

- Explainability in relation to AI means being able to give full and clear explanations of any decisions made by or with the assistance of AI systems.
- In the UK, the General Data Protection Regulation, the Data Protection Act 2018 and the Equality Act 2010 impose certain requirements for explainability, which have been a challenge to apply practically for businesses using AI.
- Additional explainability requirements may apply depending on sector (e.g. for regulated financial institutions) and deployment use case.
- In May 2020, the UK Information Commissioner's Office and the Alan Turing Institute jointly published detailed guidance on AI explainability following an industry consultation.
- The guidance provides comprehensive and practical advice on explaining decisions made by AI, with a focus on meeting both existing legal requirements and technical and governance best practice.
- Context is key – one explanation will rarely be fit to fulfil all requirements or suitable for all audiences..

One area of focus for regulators and a problem for organisations has been the need for greater transparency and “explainability” for artificial intelligence (AI) systems. As a result, a range of academic, industry and government initiatives have sought to give practical context to new legal requirements and help organisations to counter accusations that AI systems are opaque or act as a “black box”. The latest of these, is guidance from Project explAIn, a collaboration between the UK Information Commissioner's Office (ICO) and the Alan Turing Institute, published in May 2020 following an industry consultation. This aims to provide practical advice on explaining decisions made by AI systems, in a manner that meets legal requirements, as well as technical and governance best practice. Here we consider some of the existing legal requirements for explainability in the UK and explore the key takeaways from this guidance.

### What is explainability?

In brief, having broad oversight of AI systems, and any decisions made by or with the assistance of AI, and being able to give full and clear explanations of each stage of the process to relevant stakeholders.

### Overview of key legal requirements

Currently, the regulatory frameworks closest to AI explainability in the UK are the General Data Protection Regulation (GDPR) and UK Data Protection Act, 2018 (DPA 2018). While both are technology agnostic, they include a number of provisions covering automated

decision-making more broadly (which includes AI-assisted decisions). These apply where such systems use personal data (such as to train machine learning models), or create personal data (which includes making any decisions about individuals using personal data).

Where the GDPR and DPA 2018 apply, organisations must meet a number of requirements, starting with the seven key data protection principles. Of these, fairness, transparency and accountability are particularly relevant in the context of AI. Additionally, individuals have the right to:

- be informed where solely automated decisions (such as those using AI systems) are made which produce legal or ‘similarly significant’ effects for them;
- be provided ‘meaningful information’ about the logic involved in such decisions;
- know the significance and envisaged consequences for the individual; and
- object to their data being processed in this manner (including for profiling).

In addition, Article 22 of the GDPR gives individuals the specific right not to be subject to solely automated decisions that produce legal or similarly significant effects (with limited exceptions). The Article 29 Working Party's [guidelines on automated decision-making and profiling](#), endorsed by the European Data Protection Board, provide helpful additional guidance on these aspects, including what to consider when assessing decisions producing legal or similarly significant effects on individuals.

The UK Equality Act 2010 (**Equality Act**), which applies to organisations including government departments, service providers and employers, also gives rise to a requirement for explainability. The Equality Act prohibits discrimination on the basis of certain “protected characteristics” including race, age, disability, sexual orientation and marital status. Where the Equality Act applies and AI is used in a decision-making process, organisations must ensure that this does not result in discrimination on the basis of a protected characteristic, and be able to demonstrate this in a suitable explanation to the decision recipient in a format that they can meaningfully engage with.

Further requirements also apply to use of AI by regulated financial institutions, which impact on explainability and more broadly. The Prudential Regulation Authority and Financial Conduct Authority (FCA) have both shown their willingness to apply existing regulatory principles to such uses, including rules on Senior Management Arrangements, Systems and Controls (SYSC). For example, where robo-advice (automated investment services) are being provided and involve the processing of personal data, regulated financial institutions will need to consider how any explanation might need to meet the expectations of both data and financial regulators.

Beyond these, a range of regulations covering consumer protection, competition law, human rights, healthcare and public sector use of such systems may apply, depending on deployment use cases.

## Overview of the Project explAIIn guidance

The guidance is structured as follows:

- **A high-level overview** aimed primarily at data protection officers and compliance teams of what AI systems are, where they are commonly used, the legal requirements and obligations around explaining AI, and common types of explanations and associated actors.

- **Practical guidance** aimed primarily at technical teams on approaching explainability from the point of view of key legal requirements, and the technical options and measures available for producing and delivering explainability in context.
- Guidance aimed primarily at senior managers on **operationalising explainability** at the organisational level through governance measures such as key policies and processes.

## The types of explainability

The guidance (and much of the underlying research in this space) categorises types of AI explanations as either **process based** or **outcome based**. Process based explanations describe the system and how it has been designed and deployed keeping best practice in mind, so that the reasoning behind a specific decision is made more transparent. Outcome based explanations, as the name suggests, explain a particular decision made by an AI system, as commonly seen in the credit scoring space for example.

The guidance next identifies six ways of explaining decisions made by AI systems. These are (i) **Rationale explanation**, (ii) **Responsibility explanation**, (iii) **Data explanation**, (iv) **Fairness explanation**, (v) **Safety and performance explanation**, and (vi) **Impact explanation**. The guidance describes each of these in detail, including the purpose they serve, what they constitute, and the technical steps involved (such as pre-processing data, building and using interpretable systems, model choices, and extracting information for delivering explanations).

While these categories may seem largely self-explanatory, it is important to consider why these six types of explanation might be key and what they mean in practice. The guidance notes the importance of clarity, ease of understanding and most importantly, context. The underlying context of what AI is in use and where, who the explanation is targeted at, and what

level of information would provide them an adequate and clear explanation are key. One explanation will rarely be fit for all purposes and audiences. These criteria also resonate with the legal requirements discussed above, around providing meaningful information about the use of AI, and the corresponding rights individuals have. The guidance also importantly highlights that these transparency requirements continue to apply where there is a “human-in-the-loop”, i.e. the AI system is not ‘solely automated’.

### **Creating and governing an explainability ecosystem**

The guidance helpfully also sets out governance advice for organisational stakeholders who may not be involved in the technical implementation and delivery of AI explanations, but play a critical role nonetheless, such as senior management and compliance professionals. The successful and meaningful delivery of AI explanations entails putting in place robust, end-to-end processes across the organisation’s AI lifecycle, from initial decision-making, to product management, development, implementation, delivery, compliance and senior management oversight.

In this regard, the guidance highlights the importance of implementing policies and procedures providing governance for a business’ use of AI, including covering the functioning of these key stakeholders and enabling the delivery of AI explanations in a manner that is consistent, clear, accountable and supports organisational adoption. It also provides high level guidance on what these policies and procedures must cover. While the policy elements are similar to those organisations may implement for other governance programmes (such as data protection), the recommended documentation and processes are tailored for AI systems, and contextualise key GDPR

documentation requirements with illustrations, including (i) Documenting key decisions around using AI, (ii) How explanation types are selected, (iii) Documenting data collection and procurement for AI, (iv) Pre-processing, (v) Model selection, building, testing and monitoring, and (vi) Explanation extraction and delivery. The guidance further provides examples of how this documentation can be put in place, for example through argument-based assurance cases (described in detail in Annexe 5 of the guidance).

### **Looking ahead on explainability – what action should businesses take now?**

Project explAI and this comprehensive guidance incorporating industry input have made important advances for organisations building and using AI. Whether AI systems are used off the shelf or developed in-house, it is crucial that organisations develop and implement appropriate governance and compliance programmes, addressing explainability requirements along with other legal and ethical considerations. It is particularly important that AI governance is not considered in a silo: the interplay with existing data, human rights, ethics or other compliance frameworks (e.g. financial services regulation), is critical.

The global regulatory landscape is evolving quickly, and universally, as regulators in the UK, EU and beyond narrow their focus and prioritise AI and other advanced technologies. This is highlighted for example, in the ICO’s broader technology strategy and upcoming framework on auditing AI, as well as the increased focus by regulators such as the Financial Conduct Authority and the Competition and Markets Authority on the use of AI and its implications for the areas they regulate. At the EU level, the European Commission published a white paper

‘On Artificial Intelligence – A European approach to excellence and trust’ in February 2020, promoting a risk-based common European approach to help prevent fragmentation. The white paper includes recommendations for a number of improvements to the existing EU legislative frameworks in the context of AI, including a focus on product liability, transparency, safety, consumer protection and more effective enforcement of existing EU legislation relevant for AI. More broadly, regulators the world over are beginning to turn their attention to focused regulatory issues around managing machine learning model risk, lack of transparency and anti-competitive behaviour, sending clear signals on the road ahead.

Courts too, when considering liability for AI systems or claims by data subjects that businesses have failed to comply with GDPR and the DPA 2018, will closely scrutinise both the underlying explanation of the AI and the governance around it (as such materials are likely to have to be disclosed as part of proceedings).

For organisations developing or deploying AI systems, it is crucial to carefully consider and apply the principles emerging from the global patchwork of guidance, through new or existing policies and processes. Our recent briefing on [Artificial Intelligence Risk](#) provides practical tools for assessing your AI risk and responsible AI management.

Being a part of the conversation early on (as has been the case with those that participated in the development of this guidance) is critical, and we are already seeing first movers emerge who are helping to share best practice. Organisations which are ahead of the curve in implementing a comprehensive governance and compliance programme for their use of AI, that addresses both legal requirements and ethical considerations will yield significant benefits for themselves, their customers and society at large. Those that do not, are potentially increasing their legal risk, as a higher volume of data subject access requests and claims, investigations, and enforcement action around the use of AI will inevitably follow.



## CONTACTS



**Jonathan Kewley**  
**Partner**  
T: +44 20 7006 3629  
E: jonathan.kewley  
@cliffordchance.com



**Kate Scott**  
**Partner**  
T: +44 20 7006 4442  
E: kate.scott  
@cliffordchance.com



**Laura Nixon**  
**Senior Associate**  
**Knowledge Lawyer**  
T: +44 20 7006 8385  
E: laura.nixon  
@cliffordchance.com



**Monica Freely**  
**Senior Associate**  
T: +44 20 7006 2322  
E: monica.freely  
@cliffordchance.com



**Arnav Joshi**  
**Senior Associate**  
T: +44 20 7006 1303  
E: arnav.joshi  
@cliffordchance.com



**Herbert Swaniker**  
**Lawyer**  
T: +44 20 7006 6215  
E: herbert.swaniker  
@cliffordchance.com

## OUR INTERNATIONAL NETWORK 32 OFFICES IN 22 COUNTRIES



Abu Dhabi  
Amsterdam  
Barcelona  
Beijing  
Brussels  
Bucharest  
Casablanca  
Dubai  
Düsseldorf  
Frankfurt  
Hong Kong  
Istanbul

London  
Luxembourg  
Madrid  
Milan  
Moscow  
Munich  
Newcastle  
New York  
Paris  
Perth  
Prague  
Rome

São Paulo  
Seoul  
Shanghai  
Singapore  
Sydney  
Tokyo  
Warsaw  
Washington, D.C.

Kyiv<sup>1</sup>  
Riyadh<sup>2</sup>

<sup>1</sup> Clifford Chance has a best friends relationship with Redcliffe Partners in Ukraine.

<sup>2</sup> Clifford Chance has a co-operation agreement with Abuhimed Alsheikh Alhagbani Law Firm in Riyadh

C L I F F O R D

C H A N C E

This publication does not necessarily deal with every important topic nor cover every aspect of the topics with which it deals. It is not designed to provide legal or other advice.

[www.cliffordchance.com](http://www.cliffordchance.com)

Clifford Chance, 10 Upper Bank Street,  
London, E14 5JJ

© Clifford Chance 2020

Clifford Chance LLP is a limited liability partnership registered in England and Wales under number OC323571  
Registered office: 10 Upper Bank Street,  
London, E14 5JJ

We use the word 'partner' to refer to a member of Clifford Chance LLP, or an employee or consultant with equivalent standing and qualifications.

If you do not wish to receive further information from Clifford Chance about events or legal developments which we believe may be of interest to you, please either send an email to [nomorecontact@cliffordchance.com](mailto:nomorecontact@cliffordchance.com) or contact our database administrator by post at Clifford Chance LLP, 10 Upper Bank Street, Canary Wharf, London E14 5JJ.

Abu Dhabi • Amsterdam • Barcelona  
Beijing • Brussels • Bucharest  
Casablanca • Dubai • Düsseldorf  
Frankfurt • Hong Kong • Istanbul  
London • Luxembourg • Madrid  
Milan • Moscow • Munich • Newcastle  
New York • Paris • Perth • Prague  
Rome • São Paulo • Seoul • Shanghai  
Singapore • Sydney • Tokyo • Warsaw  
Washington, D.C.

Clifford Chance has a co-operation agreement with Abuhimed Alsheikh Alhagbani Law Firm in Riyadh.

Clifford Chance has a best friends relationship with Redcliffe Partners in Ukraine.