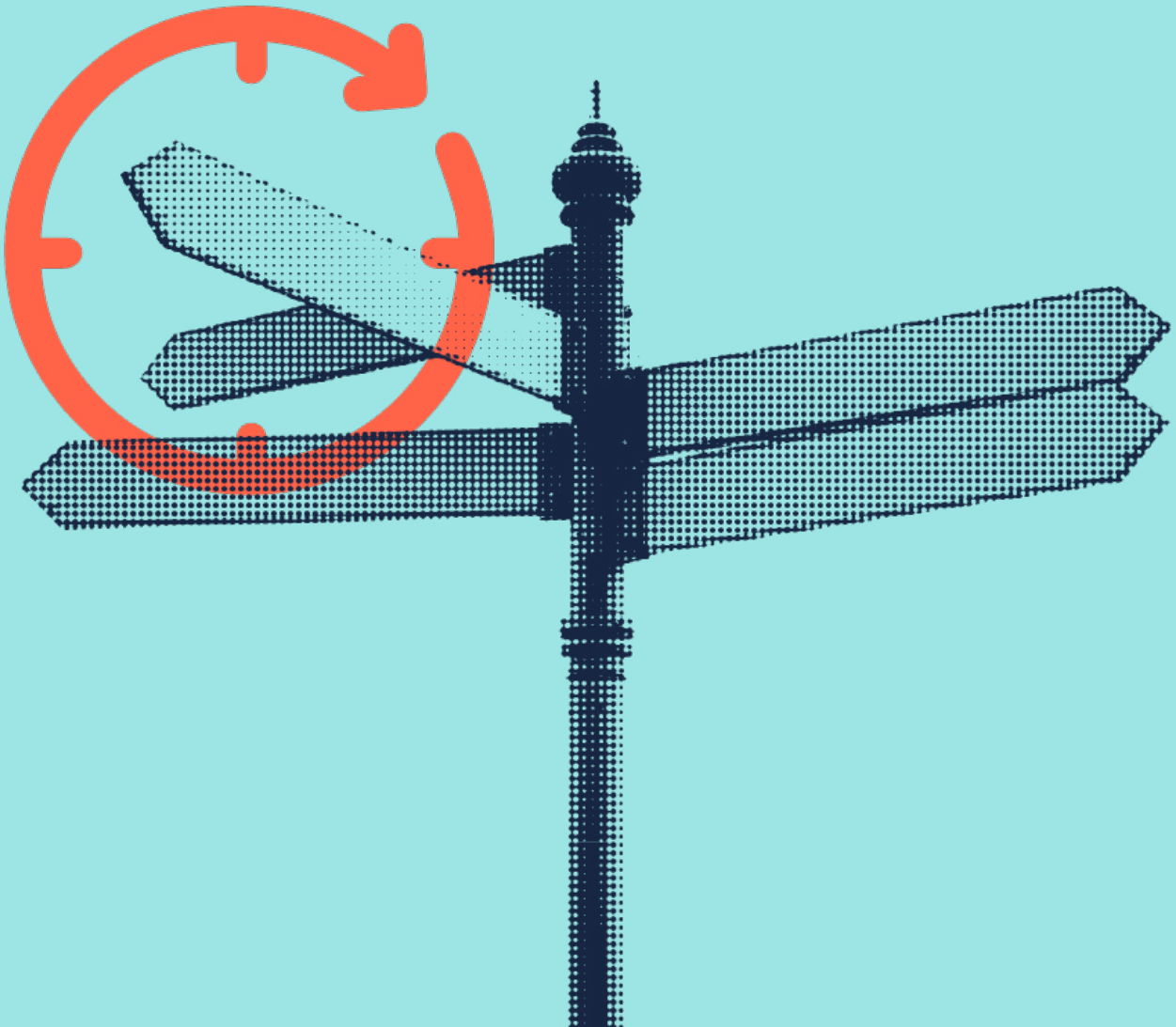
Industry Temperature Check

Barriers and Enablers to AI Assurance

December 2022



Centre for
Data Ethics
and Innovation

Table of contents

Overview

AI assurance - mechanisms to assess and communicate reliable evidence about the trustworthiness of AI systems - has an important role to play in helping to achieve the government's ambitions for a risk based, pro-growth approach to AI governance, as set out in the National AI Strategy.

Providing a toolbox of assurance mechanisms for use with AI - such as technical and governance standards, impact assessments, and possibly in the longer-term, certification - will enable greater adoption of AI and data-driven technologies, while supporting organisations to innovate responsibly. The Roadmap to an Effective AI Assurance Ecosystem, developed by the Centre for Data Ethics and Innovation (CDEI), sets out a path to building an effective and mature ecosystem of AI assurance services in the UK.

To support delivery of the Roadmap, the CDEI launched its AI Assurance Programme. In its first year, the focus of the programme has been to gain a better understanding of current levels of industry engagement with AI assurance, to best focus our efforts on areas with the highest potential for impact.

Since the publication of the Roadmap, the CDEI has facilitated a series of events with stakeholders. Our ***Industry Temperature Check: Barriers and Enablers to AI Assurance*** summarises key findings from these activities, which included: a series of Ministerial roundtables, the CDEI x techUK AI assurance symposium, semi-structured interviews, and an online survey, reflecting the views of diverse stakeholders across sectors.

In addition to this, we have chosen to examine in more detail three sectors that face distinct challenges from increased AI adoption: HR and recruitment, finance, and connected and automated vehicles (CAV). This is to ensure that we capture a breadth of concerns and incentives for implementing AI assurance across the economy.

This publication identifies **industry barriers and enablers to engaging with AI assurance,** to identify **potential practical interventions to support increased uptake and adoption** of AI assurance techniques and standards. The report is broken into four sections, the first focusing on cross-sectoral findings, with the following sections focusing on sector-specific findings from HR and recruitment, finance, and CAV.

The findings illustrated in this paper will inform the continued development of the CDEI's AI assurance Programme and inform our practical interventions to support the development of a thriving AI assurance ecosystem.

General findings:

Key themes

Key themes

Over the past year, the CDEI has engaged with AI developers, AI assurance service providers, and industry executives from startups, SMEs, and multinationals across a variety of sectors, to gauge familairity and engagement with AI assurance and identify priority areas for supporting the development of a world-leading AI assurance ecosystem in the UK. This exercise has identified a number of key themes, outlined below:

AI assurance as part of wider risk management

AI assurance was often contextualised by participants as an important element of a **wider organisational risk management framework.** Participants reported **developing or expanding existing risk management frameworks to include AI-related risks.**

These include **technical risks,** which can be mitigated by modifications to model design or input data, and **governance risks,** which can be mitigated by changes to organisational policies or processes.

Industry support for a proportionate approach to assurance

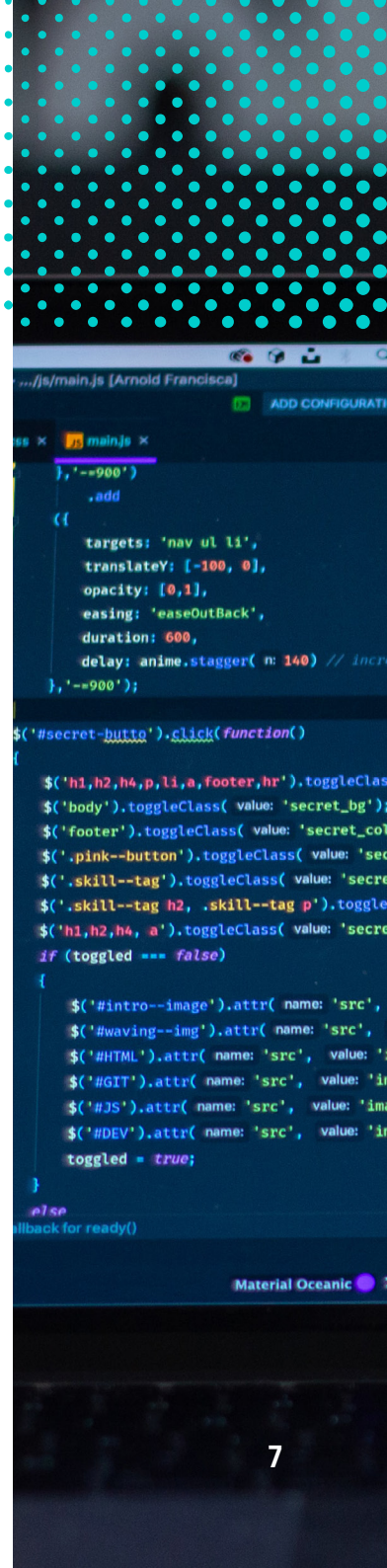
Participants emphasised that selecting an assurance technique to evaluate a system is **dependent on the context in which the system is deployed.** They noted that the appropriate technique may be determined by a range of factors, including **lifecycle stage, risk category, risk level, sector, use case, and legal/ regulatory requirements**. Typically, participants supported a **proportionate approach to assurance,** in which low-risk sectors/use cases utilise less formal assurance techniques (e.g. impact assessment), while high risk industries/use cases utilise a **combination** of assurance techniques (e.g. impact assessment, as well as performance testing, conformity assessment and/ or validation).

Industry desire for third-party certification/accreditation

Many participants felt the use of **third-party tools and services - including cloud-based and software-as-a-service (SaaS) assurance platforms - was preferable to using internal services,** as they provide an impartial perspective. However, there are **concerns around the consistency and robustness of third-party assurance** services. Participants expressed a desire for **certification or accreditation schemes as a means of recognising and demonstrating the credibility and quality** of third-party assurance service providers.

Standards to support AI assurance techniques

Many participants referenced using standards developed by standards development organisations (SDOs) alongside other assurance techniques. Some adopted **standards directly, while others used them as a point of reference for what they should be assuring for** (e.g. explainability and robustness) and then developed their own methods for achieving these aims. Organisations which did not use technical standards reported that this is, in part, because the **standards landscape is complex and difficult to navigate.**



Prevalence of impact assessments

The most frequently cited assurance techniques were impact assessments. Impact assessments **pose questions to identify potential ethical and societal impacts of an AI system**, and may focus on design and development processes, or wider organisational processes. **Participants noted that** impact assessments are often the **initial stage of an assurance engagement,** used to identify what additional measures may be required to assure the system.

Assurance creates competitive advantage

There was a strongly held view among participants that AI assurance can **provide organisations with 'a competitive edge'**, through **building customer trust** and **managing reputational risk.** On one hand, using assurance techniques to evaluate AI systems can build trust in **consumer-facing AI systems** by demonstrating adherence to ethical values (fairness, transparency etc.) and/or relevant regulation/legislation. On the other hand, using assurance techniques can also help identify and mitigate AI-related risks to **manage reputational risks** and avoid negative publicity. This helps to mitigate greater commercial risks, in which high-profile failures could lead to reduced customer trust and adoption of AI systems.

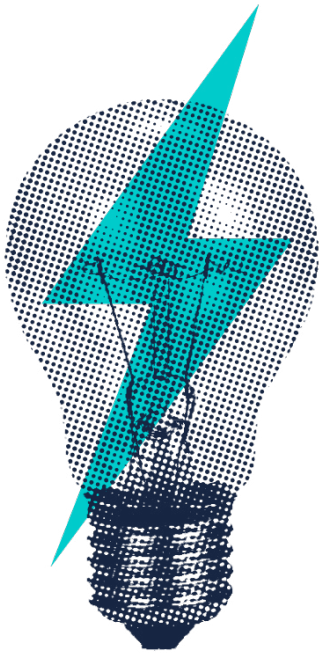
Regulatory compliance is a key driver of assurance

The need to **comply with relevant existing and future regulation and legislation to demonstrate best practice and avoid penalties** was identified by participants as a key driver of AI assurance, and is likely to drive **both the adoption of existing assurance techniques as well as the development of new assurance techniques and standards.** Organisations will need to comply with both existing legislation like the UK General Data Protection Regulation (UK GDPR) as well as anticipated future regulatory frameworks like the regime to be outlined in the UK's forthcoming White Paper on AI Regulation.

Participants were eager to understand how different assurance techniques could help organisations to demonstrate implementation and integration of the proposed regulatory principles set out in the UK Government's July 2022 Establishing a pro-innovation approach to regulating AI paper, highlighting how regulation can also drive requirements for AI assurance. Participants were also mindful of the EU AI Act, which they anticipated is **likely to mandate certain assurance activities** such as risk management frameworks and conformity assessments, acting as a direct driver of AI assurance.

Barriers to AI assurance

Over the past year, the CDEI has worked with industry to identify key barriers and enablers to engaging with AI assurance techniques and standards. **The table below summarises common barriers faced by participants across sectors.**



Barrier type	Description
Workforce barriers	**Lack of knowledge/skills:** In many organisations that design and develop AI systems, staff either don't know what unique risks AI poses, or can identify these risks but don't have the skills or knowledge to effectively mitigate them. Alternatively, in organisations that procure AI systems, teams are often unaware that they may be required to monitor, prove or assess the performance of these systems over time.
Organisational barriers	**Lack of buy-in from senior management teams:** Many senior decision-makers (e.g. managing directors/partners) are often unaware of the concept of responsible AI, and as a result don't prioritise or fund measures to support the development of responsible systems. Participants noted the importance of making a strong business case for AI assurance, which can help senior leaders demonstrate return on investment (ROI) in responsible AI. **Lack of resources:** Lack of financial resources is a common barrier to AI assurance, specifically for the adoption of standards developed by standards development organisations (SDOs). Many SDO-developed standards are labour intensive and costly, requiring considerable time and effort to adopt. This is a particularly big barrier for small and medium-sized enterprises (SMEs), who typically have more limited resources to devote to responsible AI development.

Barrier type	Description
Operational/ market barriers	**Lack of standardised/unified approach:** Due to the newness of the AI assurance market, there is fragmentation and a lack of consistency across AI assurance service providers. For example, each provider may use different metrics and/or measurement techniques for assessing an AI system. This multitude of approaches makes it difficult for organisations to determine which assurance service provider or technique is best suited to assess their AI systems.
Governance barriers	**Regulatory uncertainty:** There is considerable hesitancy to invest resources in adopting assurance techniques or standards that may be irrelevant or incompatible with future regulatory requirements. There is a need for more clarity around which standards/assurance techniques may support compliance with future regulatory requirements.

Interventions

Repositories and guidance

Participants expressed a clear desire for 'concrete and operational guidance' for AI assurance, particularly to aid in the **identification of relevant techniques and standards.** The recently launched AI Standards Hub was cited as an example of an initiative that will help industry stakeholders to identify the most relevant standards for their context. There is a similar appetite for **repositories, libraries, or knowledge hubs which showcase AI assurance techniques, frameworks and principles**, and highlight their relevance for a specific sector.

Support for SMEs

Participants from SMEs reported having limited resources and/or relevant expertise within their organisations to dedicate to AI assurance. Therefore standards, assurance techniques and related guidance **need to be clear, concise, and presented in a way that SMEs can implement in a financially and time efficient manner.** Participants suggested that access to **a library of free tools would be a useful resource** to support SMEs to identify and use relevant AI assurance techniques and standards, within tight budgets and timescales. Additionally, participants suggested that **mechanisms for SMEs to partner with other organisations and/or academia** could also help to bolster limited internal expertise and resources**.**

Communication across disciplines

It is important to **build common language and understanding** across the diverse set of stakeholder communities and disciplines involved in AI assurance. Participants expressed a need for **assurance techniques and evaluation frameworks that are comprehensible to non-technical staff** as well as **established definitions for foundational concepts** like 'fairness' and 'explainability'.

Clear link between regulation and assurance

Across our engagements, participants expressed that communicating the link between how **AI assurance activities may support compliance with relevant regulatory frameworks will be a key motivator for industry engagement**. To this end, there is considerable demand from industry for resources from regulators that communicate this link and provide clear guidance to help organisations understand what is required of them to demonstrate compliance. This is of particular importance for **organisations that operate internationally, with a need to understand the similarities and differences between regulatory requirements across different jurisdictions.** Participants suggested that this could be achieved by mapping key international regulatory requirements and frameworks for AI assurance to identify areas where more consistency or collaboration is needed. It was suggested that this could also be achieved through international cooperation and collaboration in SDOs to ensure the alignment of national and international standards objectives.

Sector-specific findings:

Overview

Overview

The following sections summarise our analysis of key barriers and enablers to AI Assurance within sectors. **We've adopted a decentralised, context-based approach to align with the UK's approach to AI regulation**. As outlined in the National AI Strategy and reiterated in the government's policy paper Establishing a pro-innovation approach to regulating AI, the UK will adopt a context-based, pro-innovation approach to regulating AI, in which regulators will set sector-specific guidelines for compliance with six proposed principles for AI regulation. These are:

- Ensure that AI is used **safely**
- Ensure that AI is **technically secure and functions as designed**
- Make sure that AI is appropriately **transparent and explainable**
- Embed considerations of **fairness** into AI
- **Define legal persons' responsibility** for AI governance
- Clarify **routes to redress** or contestability.

AI raises unique risks depending on its context of use. As such, the risks and appropriate regulatory responses must be considered in the relevant context.

Different sectors have varying levels of readiness and skill for the implementation and governance of AI. The following sections explore current engagement with AI assurance, as well as key barriers and interventions to encourage the uptake of AI assurance across three sectors:

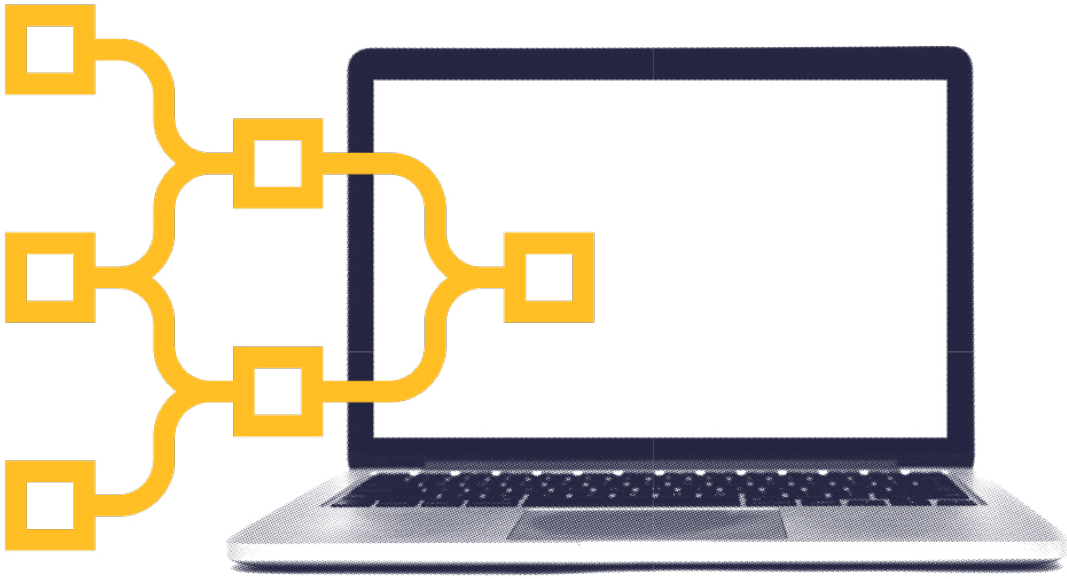
- **HR and recruitment**
- **Finance**
- **Connected and automated vehicles (CAV).**

These sectors were selected based on the breadth of risks introduced by increased AI adoption. **Each sector faces distinct risks posed by AI and is subsequently likely to require different types of AI assurance to manage these risks and support compliance with the wider regulatory framework.** For example, the use of AI in HR and recruitment may introduce risks of discriminatory bias, requiring system fairness. The use of AI in finance may introduce risks of cyberattack and financial fraud requiring technical security and robustness. Finally the use of AI in CAV may introduce risks to human life, requiring safety and routes to redress.

Summary view across sectors

While the barriers to AI assurance are unique for each sector, our analysis found that a large number of barriers were common across the sectors we examined. Some familiar risks such as lack of knowledge/skills, and lack of awareness of assurance techniques and technical standards were particularly prominent across sectors.

Respondents were asked to select the three most pressing barriers to engaging with AI assurance techniques and standards in their sector. This means that for sectors in which a barrier was not identified, this barrier may still exist - but is of less importance in the view of respondents. The lack of identified barriers in the HR and recruitment sector may also reflect a smaller sample size of respondents.



Barrier	HR & Recruitment	Finance	CAV
Lack of knowledge/skills	🔴	🔴	🔴
Lack of guidance	---	🔵	🔵
Lack of awareness of available assurance techniques	🔵	🔴	🟣
Lack of awareness of available standards	---	🔴	🔵
Lack of signposted best practice	---	🔴	🔴
Lack of demand (both internal/external)	🔴	🔵	🔴
Difficult to choose an appropriate technique/standard	🟣	🔵	🟣
Financial cost of standards	---	---	🟣
Lack of international interoperability	🟣	🟣	🟣
Regulatory uncertainty	---	🔴	🟣
Lack of mechanisms to recognise assurance efforts	---	🟣	🔴
Complexity of standards landscape	---	🟣	🟣

🟣 Low Priority

🔵 Medium Priority

🔴 High Priority

--- Barrier not identified

HR and recruitment:

Overview

Overview

In the HR and recruitment sector, AI and data-driven systems are applied across a range of functions within the recruitment life cycle. These include:

- **Sourcing:** Attracting high-quality candidates to employment opportunities using targeted advertising and recommendation systems, automated CV matching, AI chatbots, and multi-database candidate sourcing.
- **Screening:** Assessing and sifting potential candidates using CV screening and evaluation software, as well as game-based assessments and psychometric testing.
- **Interview:** Supporting recruiters in the interview stages using video screening software (including voice and emotion expression recognition), automated 'asynchronous' video interviews, and automated transcription.
- **Selection:** Analysing large amounts of data to execute automated background checks, and providing recommendations on salary offers a candidate is likely to accept.

Why do we need AI assurance in this sector?

Adopting AI-enabled tools in the HR and recruitment sector offers the automation and simplification of existing processes, promising greater efficiency, scalability, and consistency. However, these technologies also pose novel risks. These risks include:

- Discriminatory job advertising/targeting
- Perpetuating existing biases and creating new barriers
- Digital exclusion
- Low accuracy and ease of explanation of recruitment tool models
- Failure of consent models for data collection, use, sharing, and retention
- Lack of specific regulatory framework and limited specific, dedicated regulator resourcing, driving compliance challenges and costs
- Potential implications for data processing and data protection laws.

AI assurance techniques can play a vital role in managing these risks and building trust. They can help organisations ensure compliance with norms and principles of responsible innovation beyond regulatory compliance alone. Unlike other sectors, the use of AI in HR and recruitment is not likely to result in risks to physical safety. Instead, AI assurance can help to **maximise the benefits of these technologies and mitigate potential rights-based harms** such as those that arise from a lack of fairness**.**

Barriers to AI assurance

The table below summarises key barriers to using AI assurance techniques and technical standards in the HR and recruitment sector, as identified by individuals working in this sector.

Barrier	Low Priority	Medium Priority	High Priority
Lack of knowledge/skills			●
Lack of guidance			
Lack of awareness of available assurance techniques		●	
Lack of awareness of available technical standards			
Lack of signposted best practice			
Lack of demand (both internal/external)			●
Difficult to choose an appropriate technique/standard	●		
Financial cost of standards			
Lack of international interoperability	●		
Regulatory uncertainty			
Lack of mechanisms to recognise assurance efforts			
Complexity of standards landscape			
Relevant standards not available			

Key Barriers

Lack of resources/skills

Many HR and recruitment organisations procure AI-enabled tools from **third party providers**. As such, these organisations often have **limited in-house AI expertise, and may assume that requisite checks and balances have been performed by the supplier.** Many participants reported not knowing that they may be required to monitor and assess the performance of these systems over time.

Lack of demand

In the HR and recruitment sector, recognition of the benefits of AI assurance remains limited, despite high-risk use cases. There is a subsequent lack of both **internal and external demand for assurance** to evaluate these systems (i.e. from senior leaders and end-users/customers, respectively).

Regulatory uncertainty

While some sectors have designated regulatory bodies (e.g. the Financial Conduct Authority in Finance), there is **no dedicated regulator for HR and recruitment**. As such, whilst more than one regulator has a role in this space, there is limited dedicated, specific regulatory resourcing to provide guidance to support the compliance of AI-enabled tools with future AI regulation.

Interventions

Sector-specific guidance

HR and recruitment organisations procuring AI systems and data-driven tools from third party suppliers noted that they require additional **guidance on their responsibility for assuring these systems.** The CDEI has published Data-driven tools in recruitment guidance with the Recruitment and Employment Confederation (REC), which provides HR and recruitment organisations procuring AI systems with a **series of recommendations on assurance good practice across the procurement and deployment life cycle** (e.g. before purchasing, during purchasing, before use, during and after use). The CDEI will continue to promote this existing guidance within the sector, and gauge industry appetite for further interventions.

Demonstrate value-add by assurance

Many organisations reported not being aware of the potential benefits of using AI assurance to evaluate their AI systems. There is, therefore, a need to raise awareness about the benefits of using AI assurance techniques, particularly for senior leaders and decision-makers, **to demonstrate the ROI** of investing in AI assurance to measure, evaluate, and communicate the trustworthiness of AI systems. **Benefits may include contributing towards legal and regulatory compliance** to avoid penalties and enforcement action, ensuring the **safety and security of high-risk systems**, **increasing consumer trust, and demonstrating adherence to organisational values** (e.g. responsible innovation, ethical AI, as well as wider environmental, social and corporate governance (ESG)). The CDEI's AI Assurance Guide outlines high level benefits, and the CDEI will continue to work closely with industry to signal the value of AI assurance mechanisms.

Additional regulatory clarity

The HR and recruitment sector falls across the remit of multiple regulatory bodies, such as the Equality and Human Rights Council (EHRC) and Information Commissioner's Office (ICO). As such, there is a particular need for **more certainty around regulatory responsibility in this sector, to determine who will provide guidance on future AI regulation.** Over the coming months, the government will be working closely with a variety of regulators to understand how the proposed regulatory principles will work in practice, including in the context of regulatory overlaps.

Finance

Overview

Overview

AI and data-driven approaches are being applied across a range of functions in the financial services sector. These include:

- **Fraud detection and anti-money laundering:** Analysing patterns in financial transfer data to detect money laundering and potential cases of fraud (e.g. by identifying unusual spending activities).
- **Customer interactions**: Automating client interactions and increasing the efficiency of routine decisions (e.g. credit ratings, loan applications) through the use of chatbots and voice assistants.
- **Risk management:** Analysing large volumes of data to identify, predict and manage potential risks (e.g. credit risk, insurance pricing and asset management).
- **Compliance:** Financial firms can utilise machine learning techniques to comply more accurately and efficiently with regulatory requirements.

These approaches can also be used as part of supervisory technology to support regulators in monitoring compliance.

Why do we need AI assurance in this sector?

The use of AI-enabled systems in finance offers a wealth of benefits to society. These systems have the potential to better detect economic crime, increase cybersecurity, facilitate more thorough risk assessments, enable fintech innovation, and provide increased access to finance products. However, these technologies also pose unique risks. These risks include:

- Discriminatory bias in financial decisions
- Digital exclusion
- Consumer disempowerment
- Lack of insurability
- Data monopolies/concentration

- Potential implications for data processing and data protection laws.

Being a highly regulated industry, AI assurance in the finance sector may help organisations to **demonstrate that their AI systems comply with relevant regulation, in order to build customer trust and avoid penalties** for non-compliance. In addition, it may help to mitigate a range of potential financial and rights-based harms including **fairness and privacy for consumers**.

Barriers to AI assurance

The table below summarises key barriers to using AI assurance techniques and technical standards in the finance sector, as identified by individuals working in this sector.

Barrier	Low Priority	Medium Priority	High Priority
Lack of knowledge/skills			●
Lack of guidance		●	
Lack of awareness of available assurance techniques			●
Lack of awareness of available technical standards			●
Lack of signposted best practice			●
Lack of demand (both internal/external)		●	
Difficult to choose an appropriate technique/standard		●	
Financial cost of standards	●		
Lack of international interoperability	●		
Regulatory uncertainty			●
Lack of mechanisms to recognise assurance efforts	●		
Complexity of standards landscape	●		
Relevant standards not available			●

Key Barriers

Lack of knowledge/skills

Many organisations lack the adequate resources and skills to carry out AI assurance, as the number of people with relevant AI assurance training or skills remains relatively low. Many respondents reported a considerable need **to upskill staff on concepts around assurance, ethical AI, and responsible innovation.**

Lack of awareness of available assurance techniques

Due to the nascency of the AI assurance ecosystem, many organisations **don't know what assurance techniques exist**, where to look for assurance techniques, or which technique(s) should be used to evaluate a particular system.

Lack of signposted best practice

The finance sector has **several governance mechanisms already in place to support AI assurance**. Notably, the sector has a designated regulator, the FCA, with significant experience ensuring compliance with relevant regulation, and some financial regulation that already covers the use of AI-enabled systems. Moreover, many organisations have governance processes in place to support non-AI related assurance practices - most commonly, financial audit. However, particpants reported **lack of clarity around how to use or adapt exisitng governance frameworks to address novel AI-related risks.**

Interventions

Learning and development to increase awareness of AI assurance

To address the lack of knowledge and awareness of responsible AI, participants noted that there is a need to **increase learning and development opportunities** in the AI assurance space. **Resources for upskilling are already being offered by thought leaders in the AI assurance domain**, with room for expansion. For example, the UK AI Standards Hub currently provides free access to the CDEI/ Alan Turing Institute (ATI) Introduction to AI Assurance e-learning module. In future, the government will continue to promote resources and training material hosted on the UK AI standards Hub, that aim to demonstrate the utility of AI assurance techniques and standards.

Repository of AI assurance techniques

Many organisations stated that they lack awareness of what assurance techniques exist to measure and evaluate their AI systems. There is a **demand for a centralised repository of AI assurance techniques** to help organisations navigate the assurance landscape. In Spring 2023, the CDEI will be publishing a portfolio of AI assurance case studies, which will **showcase a range of AI assurance techniques being used across sectors,** to provide others with a reference of assurance good practice. The **OECD. AI has also published the** Tools for Trustworthy AI report**, and is currently developing a *Tools for Trustworthy AI Framework* that will include an online database** of assurance tools for trustworthy AI.

Signposted best practice

Despite the prevalence of existing assurance techniques in the finance sector, many organisations in this space reported uncertainty around best practice for AI assurance, due to a lack of clarity regarding standards and signposted guidance. Participants suggested that the existence of mature assurance techniques in this sector may be adequate to assure against AI-related risks, and expressed some hesitancy around the role of new standards to support the implementation of future regulation. Rather, in the finance sector, coherently signposted best practice may be more useful for designating how **existing governance mechanisms (e.g. audit) can be used and/or adapted to support the responsible design and development of AI.**

The UK government's policy paper on AI regulation - published in July 2022 - indicated that it may decide to issue guidance to regulators on how to implement the proposed principles in their specific context. The government also signalled an intention to look for ways to support collaboration between regulators, including the FCA, to ensure a streamlined approach to the implementation of the regulatory principles across sectors.

Connected and automated vehicles (CAV)

Overview

Overview

A CAV operates by gathering data from vehicle sensors (including some or all of **cameras, radar and Light Detection and Ranging** (LiDAR)) and feeding that data into a series of machine learning algorithms which facilitate the vehicle behaviour. Such processes typically involve multiple algorithms, each developed and designed for a specific purpose. These may include:

- **Object detection and classification:** Analysing sensor data to detect and classify objects to determine road conditions, interpret road signs and detect other vehicles and potential obstructions.
- **Object localisation:** Interpreting sensor data to situate the CAV in its surroundings based on a learned relationship between an image and the position of objects within it. For example, AI can be used to identify a pedestrian and determine how far away that pedestrian is, based on previous data relating to how a pedestrian appears at different distances.
- **Route planning and optimisation:** Analysing and cross-referencing sensor data with connected data sources (e.g. traffic reports) to optimise the trajectory of the vehicle to reduce delays and avoid congestion on the road.
- **Automated decision making:** Analysing and interpreting sensor data to make decisions related to the driving task. For example, deciding that the CAV needs to brake if a road sign has been detected.

Why do we need AI assurance in this sector?

CAVs offer a wealth of potential benefits to society. They have the potential to eliminate human error from driving, increase productivity, and assist in traffic management, reducing congestion and pollution. However, these technologies also pose unique risks. These include:

- Road traffic collisions involving CAVs
- Discriminatory CAV/robotics safety outcomes (e.g. inability to detect certain groups of people)
- Data monopolies inhibiting innovation, particularly around CAV training and real-world data
- Greater impact of cyberattacks.

CAVs are an example of the use of AI within a safety-critical system. Whilst issues like privacy and fairness are still relevant, a key reason for conducting AI assurance is to **manage risks in order to protect the health, safety and security of the vehicle users as well as other road users**. Road vehicles are also highly regulated. Therefore, AI assurance can play an important role, alongside existing assurance processes, in **testing and demonstrating regulatory compliance.**

Barriers to AI assurance

The table below summarises key barriers to using AI assurance techniques and technical standards in the connected and automated vehicles (CAV) sector, as identified by individuals working in this sector.

Barrier	Low Priority	Medium Priority	High Priority
Lack of knowledge/skills			●
Lack of guidance		●	
Lack of awareness of available assurance techniques	●		
Lack of awareness of available technical standards			●
Lack of signposted best practice			●
Lack of demand (both internal/external)			●
Difficult to choose an appropriate technique/standard	●		
Financial cost of standards	●		
Lack of international interoperability	●		
Regulatory uncertainty	●		
Lack of mechanisms to recognise assurance efforts			●
Complexity of standards landscape	●		
Relevant standards not available			●

Key Barriers

Lack of awareness of available standards

Many organisations are **not aware of the existence of CAV-related standards**. Low awareness was reflected in respondents' selection of 'relevant standards not available' as a barrier to adoption, despite the publication of many standards in this domain.

Lack of mechanisms to recognise assurance efforts

Due to the nascency of this sector, there are **very few governance mechanisms** (i.e., certification, kite marking) to recognise whether CAV manufacturers have adopted relevant AI assurance techniques and standards. Respondents expressed a **desire for such mechanisms to demonstrate their compliance to customers.**

Lack of signposted best practice

There is **limited guidance** that advises on which AI assurance techniques to use and when. Participants reported a need for tools to **aid the selection and application of assurance techniques** at each stage of the AI lifecycle.

Interventions

Resources to identify and/or select technical standards

The standards landscape is complex and difficult to navigate, with numerous standards development organisations (SDOs) publishing standards across multiple platforms. Participants noted that this can make it difficult for organisations to find standards that may be relevant for the design and development of CAV. Initial resources to support standards use include the **AI Standards Hub,** which includes a **repository of AI standards that can be filtered by domain,** including transport and autonomous vehicles, as well as **The Knowledge Base on Connected and Automated Driving,** which includes a live database of over 175 published standards in this sector.

Mature governance and regulatory landscape

Increasing consumer trust and demonstrating best practice is a key driver of AI assurance. Many organisations are eager to engage with AI assurance techniques and standards if this will allow them to obtain external recognition of best practice and/or compliance with relevant standards or regulation (i.e. via certification and/or kitemarking). However, in the CAV domain the **governance landscape is just emerging** with governance and regulatory mechanisms still in the early stages of development. Early efforts to shape this landscape to develop robust regulation and supporting governance mechanisms are underway. **The CDEI published** Responsible Innovation in Self-Driving Vehicles **in August 2022,** which set out a series of proposals for the trustworthy regulation and governance of self-driving vehicles, to inform the work of the Centre for Connected and Autonomous Vehicles (CCAV) as they develop primary and secondary legislation in this area.

This work will support the Department for Transport in delivering **Connected & Automated Mobility 2025: realising the benefits of self-driving vehicles**, a roadmap which commits to developing a new legislative framework that builds trust in self-driving vehicles while enabling innovation.

Examples of good practice

A robust assurance engagement is likely to involve multiple assurance techniques, used together to evaluate different aspects of an AI system across each stage of the AI lifecycle. However, there is **limited understanding of how to go about finding, using, and combining these assurance practices.** Participants expressed the **desire for practical guidance on AI assurance, to demonstrate 'what does good look like?'**. In Spring 2023, the CDEI will publish a portfolio of AI assurance case studies, to **showcase examples of how different organisations are using assurance techniques across sectors** and provide a reference of a starting point for assurance good practice.

Looking towards the future

As the AI assurance programme develops, the CDEI will be looking at how both we and others can address the barriers to AI assurance identified in this report. The below table sets out the **most prominent barriers across sectors, as well as potential interventions** to support industry in overcoming these challenges to adopt assurance techniques and standards.

Barrier Type	Description	Interventions
Workforce barriers	Lack of knowledge/skills	General L&D and sector-specific guidance • CDEI / ATI e-learning module on Introduction to AI assurance • CDEI / REC Data-driven tools in recruitment guidance
	Lack of awareness of available assurance techniques	Toolkit of AI assurance techniques • OECD database (forthcoming)
	Lack of awareness of technical standards	AI standards repository • AI Standards Hub

Barrier Type	Description	Interventions
Operational/ market barriers	Lack of demand	Demonstrate value add of assurance • CDEI AI Assurance guide
	Lack of mechanisms to recognise assurance efforts	Mature governance & regulatory landscape • CDEI/CCAV Responsible innovation in self-driving vehicles report
	Lack of signposted good practice	Examples of good practice • CDEI portfolio of AI assurance case studies (forthcoming)
Governance barriers	Regulatory uncertainty	Additional regulatory clarity • HMG AI Regulation White Paper describing government's proposals for a new pro-innovation approach to AI regulation (forthcoming) • Government collaboration with regulators and stakeholders to support the development and implementation of the forthcoming regulatory framework (ongoing).

Methodology

The Industry Temperature Check: Barriers and Enablers to AI Assurance was developed by analysing data from four industry engagement activities. These include:

- **Ministerial roundtables:** The former Minister for Technology and the Digital Economy hosted two roundtables in March/April 2022, to seek industry views on how the government can support the growth of an AI assurance ecosystem in the UK. The first roundtable featured AI developers, and the second roundtable convened AI assurance service providers.
- **CDEI x techUK AI assurance symposium:** The CDEI and techUK co-led a hybrid AI assurance symposium, which featured industry presentations, plenary discussions, and interactive workshops to identify key barriers and enablers to AI assurance across sectors.
- **Semi-structured interviews:** We led a series of semi-structured interviews with industry stakeholders, to share their views on the current AI assurance landscape. Interviewees worked in organisations ranging from start-ups to SMEs and multinationals, and offered perspectives across sectors including finance, IT, healthcare, business and management consulting, and HR and recruitment.
- **Online survey:** We launched an online survey targeted at engagement with organisations in the HR and recruitment, finance, and CAV sectors to better understand sector-specific barriers and enablers to AI assurance in these domains. We had 41 respondents across sectors.

We conducted qualitative thematic analysis on the data collected from each of these activities. This report summarises key findings from across our engagements.

100 Parliament Street
London
SW1A 2BQ

ai.assurance@cdei.gov.uk
www.gov.uk

Centre for
Data Ethics
and Innovation