OXFORD INSIGHTS

Regulators, industry bodies and professional bodies: their role in data assurance

February 2022

Contents

1. Executive summary

This report is part of a range of projects commissioned by the Open Data Institute (ODI) as part of its programme of work on data assurance. This programme supports Mission 1 of the UK's *National Data Strategy* to 'unlock the value of data across the economy',[1] and more specifically the aim set out in the *Mission 1 Policy Framework* to 'promote the development and use of good data standards so that data is held, processed and shared according to the FAIR (findable, accessible, interoperable, reusable) principles'.[2]

In this research, we explored the role of regulators, industry bodies and professional bodies in creating the conditions for the trustworthy sharing and reuse of data within individual sectors as well as across sectors. Data assurance plays an important role in this. The ODI defines data assurance as 'the process, or set of processes that increase confidence that data will meet a specific need, and that organisations collecting, accessing, using and sharing data are doing so in trustworthy ways'.[3] This helps to make data more accessible and usable. It also ensures that rights such as data protection and intellectual property (IP) are protected. In turn, this improves trust in the data ecosystem, enabling increased data-sharing.

We focused on regulators, industry bodies and professional bodies because they all play a key role in their data ecosystems: both directly, as organisations accessing, using and sharing data themselves, and indirectly, as bodies which influence the data practices of other organisations. They all conduct work to define principles, influence regulation and best practices, develop and enforce professional norms and ethics, take action to censure those that breach rules, and build skills through training. These bodies, however, have different powers and levers available to them. Using both 'hard' and 'soft' powers, all three types of organisation can support the development and adoption of data assurance schemes.

[1] 2019, Department for Digital, Culture, Media and Sport (DCMS), 'National Data Strategy', https://www.gov.uk/guidance/national-data-strategy

[2] 2021, DCMS, 'National Data Strategy Mission 1 Policy Framework: Unlocking the value of data across the economy', https://www.gov.uk/government/publications/national-data-strategy-mission-1-policy-framework-unlocking-the-value-of-data-across-the-economy/national-data-strategy-mission-1-policy-framework-unlocking-the-value-of-data-across-the-economy

[3] 2021, Open Data Institute (ODI), 'How does data assurance increase confidence in data?', https://theodi.org/article/how-does-data-assurance-increase-confidence-in-data/

1.1 Key findings

Our research highlighted opportunities and challenges for the adoption of data assurance schemes across different sectors. The UK's rapidly changing policy context for the access, use and sharing of data presents an opportunity to support data assurance activities but also poses challenges for the adoption of data assurance schemes relying on 'harder' powers provided for in existing legislation.[4] With a potential loosening of the UK's regulatory framework, it is likely that there will be a greater role for data assurance in building trust. Data assurance will help in maintaining confidence in data practices and supporting data flows between organisations, industries and governments.[5]

All three types of organisation have existing powers and levers they can use to help facilitate the trustworthy sharing, use and reuse of data. As different sectors are at different stages of data and digital maturity, however, some sectors and organisations may need support in designing and adopting data assurance schemes.

1.2 Suggestions for regulators, industry bodies and professional bodies

We have identified suggestions for these three types of organisations to help create conditions for the trustworthy sharing, use and reuse of data.

There are opportunities for all three types of organisation to:

- Lead by example by modelling industry-leading practices with the data they collect, hold and steward. Regulators often collect and hold data for their sector, industry bodies can steward data, and professional bodies collect and hold membership and other relevant data. With the data they regularly collect, all three bodies can, in different ways, model best practices for data assurance.
- Integrate data assurance into their existing activities and services. Regulators can incorporate it into existing regulatory interventions, whilst professional bodies and industry bodies can offer training and professional development programmes on trustworthy data practices. The Food Standards Agency's (FSA) use of data assurance in food labelling, for example, demonstrates that data assurance can be built into multiple regulatory interventions.

[4] 2021, ODI, 'Data: a new direction, Open Data Institute response', https://docs.google.com/document/d/1DUN51AR57gDUS3Ck2hg2zmuYIGBq0s9ikx837ws7wh8/edit#heading=h.f22474fe7t8c

[5] 2021, ODI, 'Data: a new direction, Open Data Institute response', https://docs.google.com/document/d/1DUN51AR57gDUS3Ck2hg2zmuYIGBq0s9ikx837ws7wh8/edit#heading=h.f22474fe7t8c

- Share and collaborate to build trust through developing standards, guidance or best practice, and by collaborating with researchers addressing key themes in machine learning datasets. Through forums like the UK Regulators Network (UKRN), for example, regulators can share successful approaches to introducing new data assurance schemes with each other. At the same time, industry bodies can convene members to address sector challenges through data assurance schemes.
- Conduct reviews to inform how to best improve trust in data and data practices through new data assurance schemes. This can help to understand where trust is lacking in the ecosystem.

2. Background

This report is part of a broad range of activities commissioned by the Open Data Institute (ODI) to explore how to enable the conditions for trustworthy data sharing across a range of different data ecosystems, supporting Mission 1 of the *National Data Strategy*.[6]

As part of this work, the ODI is exploring the role of data assurance in increasing confidence in how data is being collected, used and shared.[7] By helping organisations to assess, build and demonstrate trust in data practices and data, they expect confidence in reuse and sharing of data to increase, whilst economic value is created.[8]

This project was carried out by Oxford Insights (OI), on behalf of the ODI,[9] between October 2021 and February 2022. The lead authors were Leigh Dodds and Kate Iida (OI), with thanks to Dr. Deborah Yates (ODI), Dr. Mahlet Zimeta (ODI), Ed Evans (ODI), Matt Davies (ODI), Mahad Alassow (ODI), Ellen Goodman (ODI) and Pablo Fuentes (OI) for comments. See the Methodology section for more background about how the work was carried out.

The project focused on the current and future role of regulators, industry and professional bodies in data assurance across the range of sectors and cross-sector ecosystems. The goal was to explore the direct and indirect mechanisms by which they support data assurance activities, their impacts and potential challenges and opportunities.

[6] 2021, DCMS, 'National Data Strategy Mission 1 Policy Framework: Unlocking the value of data across the economy', https://www.gov.uk/government/publications/national-data-strategy-mission-1-policy-framework-unlocking-the-value-of-data-across-the-economy/national-data-strategy-mission-1-policy-framework-unlocking-the-value-of-data-across-the-economy

[7] 2021, ODI, 'Assurance, trust, confidence - what does it all mean for data?', https://theodi.org/article/assurance-trust-confidence-what-does-it-all-mean-for-data/

[8] 2021, Frontier Economics for the ODI, 'Economic impact of trust in data ecosystems', https://theodi.org/article/the-economic-impact-of-trust-in-data-ecosystems-frontier-economics-for-the-odi-report/

[9] 2021, ODI, ' Call for proposals: Role of regulators, professional bodies and industry bodies in creating the conditions for trustworthy sharing and reuse of data', https://theodi.org/article/call-for-proposals-role-of-regulators-professional-bodies-and-industry-bodies-in-creating-the-conditions-for-trustworthy-sharing-and-reuse-of-data/

3. How can data assurance enable trustworthy data sharing?

Data, when it is shared with those who need it, can help to unlock 'the value of data across the economy'.[10] It can enable better decision making, support the design and implementation of more effective policies and be used to deliver innovative products and services.

But data and data sharing can also cause harm. The use and misuse of data can cause people to lose trust in services or organisations. Fear of causing harm can lead organisations to avoid collecting, using and sharing data. Building trust and encouraging openness are essential to avoid a future where data sharing is feared or data is hoarded.[11]

A lack of trust in how data is being collected, used and shared can lead people to opt out of data sharing and organisations avoiding sharing data that might otherwise be used. Without access to data we are unable to unlock its potential benefits.

As the ODI *Trustworthy Data Stewardship Guidebook*[12] highlights, building trust in data is important to create a world where data works for everyone. 'The economic impact of trust in data ecosystems' report further demonstrated that as levels of trust increase, data sharing does as well.[13] We need to build trustworthy data ecosystems where individuals and organisations across the public, private and third-sector trust that data is flowing in ways that will maximise benefits whilst minimising harms.

The guidebook highlights that in order to achieve that goal, we need to assess, build and demonstrate trust and trustworthiness.

3.1 What is data assurance?

Third-party assurance, reviews, audits, certification and accreditation all regularly play a role in helping to evidence trust and trustworthiness across different sectors.

[10] 2021, DCMS, 'National Data Strategy Mission 1 Policy Framework: Unlocking the value of data across the economy', https://www.gov.uk/government/publications/national-data-strategy-mission-1-policy-framework-unlocking-the-value-of-data-across-the-economy/national-data-strategy-mission-1-policy-framework-unlocking-the-value-of-data-across-the-economy

[11] ODI, 'Our theory of change', https://theodi.org/about-the-odi/our-vision-and-manifesto/our-theory-of-change/

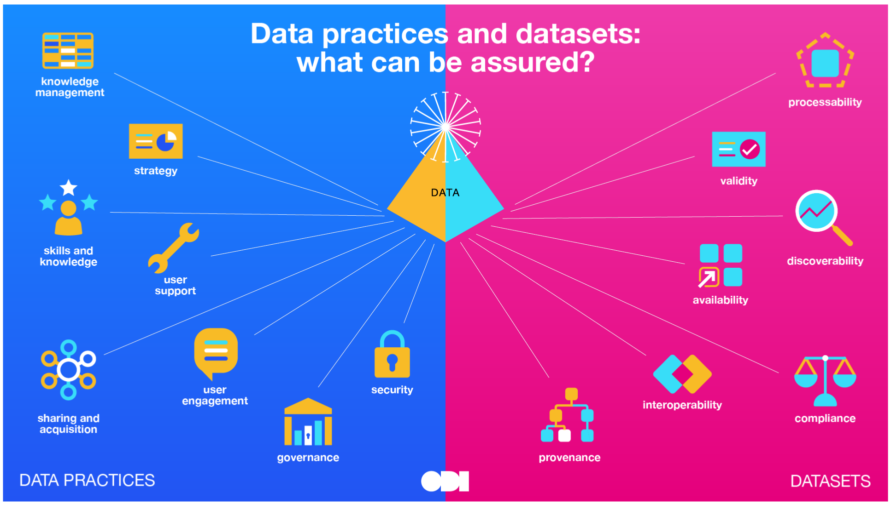
[12] 2021, ODI, 'Trustworthy Data Stewardship Guidebook - Overview', https://open-data-institute.gitbook.io/p22-trustworthy-data-stewardship-guidebook/-MW92wuAXMrYPE7sgA-M/introduction/overview

[13] 2021, Frontier Economics for the ODI, 'The economic impact of trust in data ecosystems', https://theodi.org/article/the-economic-impact-of-trust-in-data-ecosystems-frontier-economics-for-the-odi-report/

We use different types of mechanisms to help to build trust and demonstrate the trustworthiness of people, processes, products and organisations. For example, medical registers list doctors that are approved to practice medicine and consumer products undergo safety testing.

These same processes can also be applied to help to build confidence in how data is being collected, accessed, used and shared.[14] Recent research conducted by Frontier Economics on behalf of the ODI suggest that there may be a growing market for products and services in this area.[15]

3.2 What can be assured?



'Data practices and datasets: what can be assured' by Open Data Institute, used a CC-BY-SA licence.

The ODI have suggested that assurance can be applied to:

- **Data** — e.g. an individual dataset.
- **Data practices** — e.g. the processes by which data is collected or managed.

Assurance of data and data practices involves applying a range of different **data assurance activities**.

Assurance of data focuses on data or a dataset as an artefact. It includes activities like:

- Checking the quality of individual data points to assess their accuracy.

[14] 2021, ODI, 'How does data assurance increase confidence in data?', https://theodi.org/article/how-does-data-assurance-increase-confidence-in-data/
[15] 2021, Frontier Economics and the ODI, 'Review of the UK business to business data assurance market', https://theodi.org/wp-content/uploads/2021/07/PRS-ODI-Data-assurance-FINAL_23072021.pdf

- Checking the structure of a dataset against a standard to confirm that it is well-formed.
- Reviewing a dataset to ensure it does not contain personal data.

Assurance of data practices explores the broader context around how data is governed. It includes activities like:

- Reviewing the lifecycle of how data is collected, managed and shared to confirm that data is managed legally, securely and ethically.
- Assessing the decision making for how data is shared, to ensure that it is shared responsibly and in ways that will minimise harms.
- Ensuring that those involved in data practices have the necessary skills and knowledge to work with data.

3.3 The spectrum of data assurance activities

Assurance processes such as audits, certification or accreditation schemes are usually more formally defined. They will be described and supported by well-defined standards and involve multiple layers of review and assessment; not just of the process itself but the people and organisations involved in it.

There is a legal requirement in the UK, as part of the Companies Act 2006, that public companies must have their accounts audited. This helps to assure stakeholders that the company is being operated and governed responsibly. These audits are a form of financial assurance.

The process by which a financial audit is carried out will conform to an auditing standard. Businesses that carry out audits must be licensed to do so.[16] This helps to build confidence in the outputs of the audits, and that the organisations and practitioners performing them are reputable and have the necessary skills and experience to do so.

The UK Accreditation Service refers to these formal types of mechanism as standardisation, conformity assessment, measurement and accreditation.[17]

Formal methods of assurance are often referred to as 'quality infrastructure'.[18] Data assurance can be viewed as building the quality infrastructure that helps us to assess, build and develop the data infrastructure[19] and practices that enable us to sustain trustworthy data ecosystems.

Data assurance activities can also be classified based on how rigorously they are defined and applied.

[16] UK Government, ‘Become a registered auditor’, https://www.gov.uk/become-a-registered-auditor
[17] United Kingdom Accreditation Service (UKAS), 'UK Quality Infrastructure', https://www.ukas.com/ukqi/
[18] 2020, BEIS, 'The UK's National Quality Infrastructure', https://www.gov.uk/guidance/the-uks-national-quality-infrastructure
[19] ODI, 'What is data infrastructure', https://theodi.org/topic/data-infrastructure/

What is assured?	Less formal	More formal
Data	Guidance on use of modelling and design of data Provision of data documentation or data licenses	Quality control Validation against agreed standards
Data practices	Developing skills Building professional, social or organisational norms. Developing shared principles and best practices	Conducting audits of data processes and governance Certification of organisational processes against predefined standards

Less formal data assurance activities such as developing and applying norms and principles might help to guide and reinforce trustworthy behaviours. But their application may only be loosely defined and the results are not verifiable. In contrast, more formal data assurance activities typically follow well-defined processes and produce verifiable outputs.

The rigour of more formal approaches to assurance, and as a result, the increased likelihood of producing verifiable evidence of trustworthy data practices (for example, in the form of a certificate, formal verification, or entry of a person or organisation into an official register), may mean that these forms of assurance are more effective at building trust.

In its guide on AI assurance, the Centre for Data Ethics and Innovation (CDEI) also identifies a range of different techniques for assuring systems[20] suggesting more formal mechanisms create more certainty.[21]

[20] 2021, Centre for Data Ethics and Innovation (CDEI), 'Techniques for assuring AI systems', https://cdeiuk.github.io/ai-assurance-guide/techniques
[21] 2021, CDEI, 'The roadmap to an effective AI assurance ecosystem - extended version', https://www.gov.uk/government/publications/the-roadmap-to-an-effective-ai-assurance-ecosystem/the-roadmap-to-an-effective-ai-assurance-ecosystem-extended-version

3.4 Data assurance schemes

Data assurance activities have overlaps and will often be used in combination. For example, carrying out an assessment of the quality and provenance of a dataset might require:

- Quality control of the dataset — assurance of data.
- Reviewing the processes by which the data was originally collected, how it is managed and also the decision making processes that guide its sharing — assurance of data practices.

A specific **data assurance scheme** involves the application of a range of data assurance activities to help to build trust in how some specific types of data are being accessed, used and shared.

See Case study 1 for an illustrative example that demonstrates how different assurance activities can be combined to perform assurance around IT security.

A trustworthy ecosystem will involve the application of a variety of data assurance activities and schemes to help ensure a culture of good practice guides how data is being used. The importance of applying different approaches based on context is highlighted in a recent ODI report.[22]

Case study 1: Cyber Essentials scheme

What the intervention is:
Cyber Essentials is a UK government-backed scheme created to help protect organisations against cyber attacks.[23] While the scheme has broad application, IT security is a fundamental part in ensuring that data is safely and securely accessed, used and shared.

How it works:
The National Cybersecurity Centre oversees the Cyber Essentials scheme.[24] IASME, an organisation that specialises in helping businesses improve their cybersecurity, carries out delivery of the scheme.[25]

To obtain a Cyber Essentials certificate, businesses carry out a self-assessment using a standard questionnaire. This is then approved by a designated 'Certification Body' (see

[22] 2020, ODI, 'Demonstrating and assessing trustworthiness when sharing data', https://theodi.org/article/demonstrating-and-assessing-trustworthiness-when-sharing-data/
[23] National Cyber Security Centre (NCSC), 'About Cyber Essentials', https://www.ncsc.gov.uk/cyberessentials/overview
[24] NCSC, 'The National Cyber Security Centre', https://www.ncsc.gov.uk/
[25] IASME Consortium, 'Cyber Essentials: the benefits of certification', https://iasme.co.uk/cyber-essentials/

below). Cyber Essentials Plus adds an additional layer of assurance by requiring that a business's systems are independently tested by an auditor working for an Certification Body.

Businesses that have been certified can be found in a public register,[26] allowing their customers or partners to independently check their status.

IASME provides a readiness toolkit to help organisations think through their cyber security processes, so they can implement the necessary changes to gain certification. The list of essential requirements that must be met in order to achieve certification is also publicly available.[27] These requirements are jointly developed by NCSC, IASME and revised based on feedback from assessors and applicants.

A 'Certification Body' is any organisation with the authority to issue a certification. To become a body an organisation has to meet certain criteria. This includes gaining some specification certifications of their own and completing training by IASME.[28] IASME maintains a register of certification bodies.[29]

This level of accreditation helps to ensure that certifications are carried out consistently and correctly, whilst also providing applicants with a list of qualified companies that they can approach for help and support in gaining certification.

Some government contracts require that a business has Cyber Essentials certification. This creates an incentive for businesses to become compliant. A review of the impact of Cyber Essentials on UK organisations found that the scheme has an overall positive influence.[30] Organisations certified under the scheme are more likely to be aware of cybersecurity risks, confident that they can be protected from cyber attacks, and likely to conduct further efforts beyond becoming certified to guard against potential threats.

Key insights:
As noted above, technical security is fundamental to good data governance. Cyber Essentials and similar schemes are relevant to building trust in data ecosystems.

This case study provides a useful illustration of how public sector bodies can act to create a marketplace around certification and accreditation:

[26] IASME, 'Search for a Certificate', https://iasme.co.uk/certified-organisations/
[27] NCSC, 'Resources', https://www.ncsc.gov.uk/cyberessentials/resources
[28] IASME, 'Become an Assessor', https://iasme.co.uk/become-an-assessor/
[29] IASME, 'Find an IASME Certification Body', https://iasme.co.uk/certification-bodies/
[30] 2020, NCSC, 'Review of Cyber Essentials influence on cyber security attitudes and behaviours in UK organisations', https://www.ncsc.gov.uk/information/setting-baseline-ce-prior-to-iasme

- A common standard (here defined as the set of IT requirements required to achieve CyberEssentials) is used as the basis for an assessment process.
- The result of that process is a public register of certificates.
- The assessment process itself is assured by requiring those carrying out the review to be accredited.
- Anyone seeking certification can find an accredited organisation.
- An independent body provides oversight of the development of the standard, the certification and accreditation processes.

Creating a marketplace for auditors allows the scheme to scale to meet the demands of applicants. Incentives, such as procurement requirements, can drive adoption of a scheme independent of legal or regulatory requirements. See *Commercial opportunities* on page 40 for a further discussion of the opportunities offered by creating marketplaces around assurance schemes.

Assurance mechanisms used:

- Development of common technical standards.
- Certification of compliance.
- Accreditation and training of assessors.
- Public registers of certified and accredited organisations.

3.5 Who is involved in data assurance?

The ODI's theory of change[31] highlights the role of a range of organisations in helping to build a more trustworthy data ecosystem. In addition to the organisations that are stewarding data themselves or using it to make decisions, there are a variety of other organisations including regulators, researchers, and advocacy organisations with a role to play.

These other actors help to create policies and legislative frameworks that guide how data is accessed, used and shared; provide guidance, support and influence to encourage and enable others to behave in certain ways; or develop the tools and technologies that support their activities.

These same actors also have a role to play in data assurance.

Quality infrastructure relies on a variety of different types of Standards Development Organisations. In the UK this includes the British Standards Institution (BSI), Office for Product Safety and Standards, National Physical Laboratory and the United Kingdom Accreditation Service.[32]

[31] ODI, 'Our theory of change', https://theodi.org/about-the-odi/our-vision-and-manifesto/our-theory-of-change/
[32] UKAS, 'UK Quality Infrastructure', https://www.ukas.com/ukqi/

Standards Development Organisations like the BSI help to develop and maintain the standards that describe how a product or service should operate (core standards), the standard processes by which a data-enabled product or service might be certified against that standard, and standard training and assessments that are used to accredit organisations to complete these certifications.

Data assurance will also require the development of a range of open standards[33] that, as noted in the previous section, will provide the basis for validating and assessing data and data practices. The UK government recently announced an initiative to drive development and adoption of standards for AI. Given that AI is a data-enabled process, it seems like that this initiative will also look at standards for data.[34]

Researchers are currently exploring ways to carry out assurance of data. This includes standardised labelling of data and AI (e.g. The Data Nutrition label,[35] Model Cards[36]), reviewing the quality of standardised training datasets[37] and developing frameworks for algorithmic auditing that encompass aspects of data collection and use.[38]

For this report we are focusing on the role of regulators, industry and professional bodies in supporting the development and adoption of data assurance activities and schemes.

[33] 2019, ODI, Types of open standards for data, https://standards.theodi.org/introduction/types-of-open-standards-for-data/

[34] 2022, DCMS and Office for AI, 'New UK initiative to shape global standards for Artificial Intelligence', https://www.gov.uk/government/news/new-uk-initiative-to-shape-global-standards-for-artificial-intelligence

[35]Data Nutrition, 'The Data Nutrition Project', https://datanutrition.org/

[36] 2019, Margaret Mitchell et al, 'Model Cards for Model Reporting', https://doi.org/10.1145/3287560.3287596

[37] 2021, Inioluwa Deborah Raji, Genevieve Fried, 'About Face: A Survey of Facial Recognition Evaluation', https://arxiv.org/abs/2102.00813

[38] 2020, Inioluwa Deborah Raji et al, 'Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing', https://arxiv.org/abs/2001.00973

4. How are regulators, industry bodies and professional bodies involved in data assurance?

4.1 Role and activities

For the purpose of this research we have looked at the following three types of organisations: regulators; industry bodies; and professional bodies.

Regulators are public organisations or agencies set up by government to oversee activities in specific sectors or across sectors. In the UK, they are typically accountable to Parliament.[39] The purpose of regulators is to 'protect and benefit people, businesses and the environment and to support economic growth'.[40] They may impose requirements, conditions or restrictions, set standards, and enforce compliance.[41] Some regulators, like the Office for Statistics Regulation (OSR) and the Information Commissioner's Office (ICO) focus specifically on the regulation of data. Other regulators work in cross-sector ecosystems while others have sector-specific scopes.

Industry bodies, also known as trade associations, are organisations made up of businesses that all operate in a specific industry.[42] Industry bodies provide a unified voice for their members, often linking them with regulators and policymakers. They can help share and develop best practices, organise events and training, and carry out research amongst other activities.[43]

Professional bodies are organisations whose individual members all practice a profession or occupation. The organisation has the power to oversee the knowledge, skills, conduct and practice of that profession.[44] Some professional bodies are empowered by regulators or legislation, whereas for others, the powers they hold come from the permission or agreement of their members. They are also one type of organisation that a regulator may delegate responsibility to in order to deliver on desired outcomes.

Regulators often work with industry bodies and professional bodies to build trust and carry out assurance of different kinds.[45] In regulated professions, there are a range of bodies involved in

[39] 2017, National Audit Office (NAO), 'A Short Guide to Regulation',https://www.nao.org.uk/wp-content/uploads/2017/09/A-Short-Guide-to-Regulation.pdf
[40] 2017, National Audit Office (NAO), 'A Short Guide to Regulation',https://www.nao.org.uk/wp-content/uploads/2017/09/A-Short-Guide-to-Regulation.pdf
[41] 2021, Safeopedia, 'Regulatory Body', https://www.safeopedia.com/definition/625/regulatory-body
[42] Cambridge Dictionary, 'Industry Association', https://dictionary.cambridge.org/dictionary/english/industry-association
[43] 2006, Mobile Marketing Association (MMA), 'Benefits of industry associations', https://www.mmaglobal.com/articles/benefits-industry-associations
[44] Science Council, 'Our definition of a professional body', https://sciencecouncil.org/about-science/our-definition-of-a-professional-body/
[45] 2017, NAO, 'A Short Guide to Regulation', https://www.nao.org.uk/wp-content/uploads/2017/09/A-Short-Guide-to-Regulation.pdf

training and certifying people working in those sectors. The General Medical Council (GMC), for example, has worked with NHS Education for Scotland and the Royal College of Physicians and Surgeons of Glasgow to help support professional development for doctors.[46]

Scope and powers

Taken broadly, all three types of organisation have powers in common and also differences in scope and authority. All three types of organisation conduct work to:

- Define principles, such as codes of practice.
- Influence and shape regulation and best practices.
- Develop and enforce professional norms and ethics.
- Take action to censure organisations and individuals that breach rules.
- Build skills through professional development and training.

These bodies, however, have different powers and levers available to them. Regulators have statutory powers to enforce standards, and often also take on accreditation or oversight roles. In general, industry bodies and professional bodies have 'softer' powers than regulators, and can enact changes by encouraging or requiring members to comply with membership rules and by-laws, or by offering training and certification. When professional bodies oversee regulated professions, they may, however, have 'harder' powers granted by Parliament, such as the GMC.[47]

The exact powers and scope of operation varies between different organisations within the same category. The ICO, for example, has broad regulatory powers concerning information rights across many sectors, while the Financial Conduct Authority (FCA), a sector-specific regulator, has the power to regulate businesses in the finance industry, supervise firms and set standards, but does not have the same cross-sector regulatory powers as the ICO. Some industry bodies have been given the power to oversee the implementation of standards, while others exist to advocate for their members and can do so through lobbying for favourable policies. Some professional bodies have powers mandated by Parliament to regulate a profession, while others bring together practitioners in a sector or across sectors but do not have the same government-mandated authority. The powers of these bodies, therefore, may be granted by legislation or via the approval of an organisation's members.

All three of these bodies are involved in general assurance activities, such as:

- Developing standards and processes for common activities.
- Developing standards for audit or review.
- Carrying out certification of individuals and organisations.

[46]2018, Professional Standards Authority for Health and Social Care, 'Professional healthcare in the UK explained (part 2)', https://www.professionalstandards.org.uk/news-and-blog/blog/detail/blog/2018/04/25/professional-healthcare-regulation-in-the-uk-explained-(part-2)

[47] General Medical Council, 'About us', https://www.gmc-uk.org/about

- Providing training, e.g. as part of continuing professional development.
- Developing and maintaining registers of individuals and organisations, as part of implementing licensing requirements.
- Developing and sharing guidance and best practices.
- Conducting audits to ensure quality.

These categories of intervention can also be applied to data assurance.

Assurance of data may be just an extension of the existing activities these bodies undertake: for example by carrying out training on data protection as part of a broader professional development programme, or developing a data standard to support statutory reporting of data.[48] Regulators, industry bodies and professional bodies, however, may not see themselves as having any data-specific responsibilities or opportunities, aside from required compliance with personal data protection. The collection and use of data, however, underpins many of these activities and so there are opportunities to apply data assurance activities to help make data more accessible and usable. See Case study 3 for further information.

Regulators, industry bodies and professional bodies also collect, steward and use data themselves. They have an opportunity to directly apply data assurance activities to their own work and to act as exemplars for others in their sector. In the report *Mapping data in the UK government*, the ODI makes a similar recommendation about the value of bodies with influence leading by example to demonstrate data assurance.[49] See the table below for past and current examples of regulators, industry bodies and professional bodies involved in data assurance activities.

[48] See subsection 'Data assurance in statutory reporting' on page 19 for more information.

[49] 2021, ODI, 'Mapping data in the UK Government', https://docs.google.com/document/d/1QNn71vOlDHJZnoFPr_QXZzcsQLwldprR1WeA1vjawl0/edit#

4.2 Examples of regulators, industry bodies, and professional bodies supporting data assurance

Regulators, industry bodies and professional bodies are involved in a wide range of existing data assurance schemes and activities across multiple sectors and in cross-sector data ecosystems. A number of these can help to highlight examples of existing approaches to data assurance, and surface new opportunities.

The following table provides a summary of some past and current examples of data assurance schemes and activities.[50] Most of the interventions listed here fall into the 'more formal' category, with the exception of training, guidance and best practices. Some interventions, such as regulatory or data sharing sandboxes or certification, may be 'more formal' or 'less formal' interventions based on the context in which they are applied. Case study 1 demonstrates an example of certification used as a formal intervention, and Case study 5 demonstrates an example of certification as a 'less formal' intervention.

Table of data assurance schemes and activities

Intervention	What is assured?	Type of body	Description	Examples

Regulatory or data sharing sandbox	Data practices	Regulators	Secure, trusted environment to facilitate sharing of data between organisation	ICO Innovation Hub;
				FCA Digital Sandbox;
				CAA Regulatory Sandbox

Standards for data	Data, Data practices	Regulators	Development and adoption of standards that describe how data is collected, accessed, used and shared	Open Energy; **Open Banking (see Case study 4)**; Pensions dashboard data standards

[50] See section 3.2 for further exploration of types of data assurance.

Standards for audit or review	Data practices	Regulators	Development and adoption of standards or frameworks that describe how the audit, review or certification of a dataset or data practices might be carried out	ICO's AI Auditing Framework; ICO's accreditation guidelines for GDPR certification bodies

Certification	Data practices	Regulators, industry bodies, professional bodies	Providing a certification or public register of people, organisations, products that have passed through an assurance process to assess their data practices	**Cyber Essentials scheme (see Case study 1**);
				GDPR certification;
				CDMP certification (see Case study 5); DCAM certification; CDMC certification;
				Open banking (see Case study 4); BSI data protection certification

Accreditation	Data practices	Regulators, industry bodies	Accrediting organisations to offer certifications in data practices	ICO GDPR certification scheme; **Open banking (see Case study 4)**; **Cyber Essentials scheme (see Case study 1)**
Training	Data practices	Regulators, industry bodies, professional bodies	Providing training to support development of data skills	**FSA food labelling training (see Case study 3**); **CDMP training (see Case study 5)**; CDMC training; BSI data protection training
Licensing requirements	Data practices	Regulators	Requiring that people or products have to achieve a certification in data practices or have to pass through an assurance process in order to have a licence to operate	MHRA AI regulations; **Cyber Essentials scheme (see Case study 1)**; **Open banking (see Case study 4)**
Guidance, Best practices, principles	Data, Data practices	Industry bodies, professional bodies	Development of guidance, best practices and principles to help guide how data is accessed, used and shared	UKRN; EDM Council's DCAM and CDMC frameworks; PASA; Alliance for data science professional standards; FAIR data

Quality audit	Data	Regulators	Quality review of a dataset against some agreed measures	**OSR (see Case study 2)**

Data assurance in statutory reporting

Regulators frequently require businesses to publish and share data on a statutory basis, for example on gender pay disparities.[51] The requirements for statutory reporting of data define the type, frequency and detail to be reported which may be formalised as a data standard.

Regulators may also carry out assurance of data that is published within their sector, as part of statutory reporting or their broader remit. For example, the OSR assesses the quality of statistics produced by crown bodies and ensures they adhere to the code of practice for statistics. See Case study 2 for more information.

Case study 2: Compliance of national statistics

What the intervention is:
The Office for Statistics Regulation (OSR) is the regulatory arm of the UK Statistics Authority. It supports public confidence in statistics by addressing harms and ensuring that statistics serve the public good.[52] The OSR assesses the compliance of official statistics with the Code of Practice for Statistics in order to build trust in official statistics published in the UK.[53]

How it works:
The Code of Practice for Statistics is maintained by the Office for Statistics Regulation. It applies to all bodies that produce official statistics in the UK. It may also be voluntarily applied by other organisations that produce statistics. The Code of Practice is designed to ensure that official statistics are trustworthy, fit for purpose, use and easy to access.

Public bodies may request the OSR to assess data against the Code of Practice.[54] Carrying out these assessments, alongside systemic reviews and other compliance checks are all parts of the regulatory activities of the OSR.[55] The OSR may also prompt public bodies to

[51] 2020, Government Equalities Office, 'Gender pay gap reporting', https://www.gov.uk/government/collections/gender-pay-gap-reporting
[52] Office for Statistics Regulation (OSR), 'What we do', https://osr.statisticsauthority.gov.uk/what-we-do/
[53] 2022, UK Statistics Authority, 'About the code', https://code.statisticsauthority.gov.uk/the-code/
[54] 2012, UK Statistics Authority, 'How national and official statistics are assured', https://www.gov.uk/government/statistics/how-national-and-official-statistics-are-assured/how-national-and-official-statistics-are-assured#assessment-of-official-statistics
[55] OSR, 'Our Regulatory Work', https://osr.statisticsauthority.gov.uk/our-regulatory-work/

seek assessment if it is in the public interest.[56] OSR publishes guidance for statistical publishers about how the process is carried out along with material that can help them carry out a self-assessment.[57]

The OSR reviews the processes for collecting, analysing and publishing the statistics against the Code of Practice. The assessment does not confirm that the data is necessarily correct - errors are to be expected. The review instead focuses on the processes by which those statistics are generated and made available. An assessment is carried out through a process of evidence collection and review that involves not just the public body but also users of the statistics.

The results of an assessment are publicly available.[58] If a statistical dataset is deemed compliant with the code then it is designated as 'National Statistics'. A list of these datasets is available from the OSR website.[59] Public bodies that produce National Statistics have a statutory duty to comply with the code.

Key insights:

This example highlights that regulators can support the assurance of both data and data practises. The focal point for this assurance work is the Code of Practice. It provides a standard which all producers of statistics can use to improve their data practises, regardless of whether they have a statutory duty to do so.

The Code of Practice also underpins the regulatory work of the OSR, providing a consistent benchmark for reviewing how statistics are produced.

A transparent process, which produces public outputs and an official register of compliant datasets, helps to build trust in both the data and the process by which it is assessed.

A clearly described process, with guidance about performing self-assessments, helps organisations understand what will be involved, allowing them to prepare for the process and ensure that their data practices produce the necessary supporting evidence.

Data assurance activities used:

- Development and maintenance of a data standard, in this case a common code of practice, that defines how data is collected, accessed, used and shared.
- Audits and compliance checks using a transparent and well-defined process.

[56] OSR, 'Assessment', https://osr.statisticsauthority.gov.uk/our-regulatory-work/assessment/

[57] OSR, 'Guidance about Assessment', https://osr.statisticsauthority.gov.uk/guidance/guidance-about-assessment/

[58] OSR, 'Publications: assessment report', https://osr.statisticsauthority.gov.uk/publications-list/?type=assessment-report

[59] OSR, 'National Statistics', https://osr.statisticsauthority.gov.uk/national-statistics/

- Publication of a register of compliant datasets.
- Clear labelling of compliant datasets, using the national statistic marker.
- Legal and regulatory requirements to maintain the above standard, register and to enforce compliance.

AI assurance

In the UK, as well as internationally, there is substantial engagement between regulators and other bodies in work on the assurance of products and systems that use Artificial Intelligence (AI).

The development and deployment of AI-enabled products and services requires the use of large amounts of data. Decisions made about how data is used in machine learning can introduce a range of harms.[60]

Data assurance is therefore an inherent part of the assurance of AI. This assurance involves, for example, reviewing the quality of the training datasets used to develop machine-learning models.

The ICO is in the process of creating an auditing framework for AI, which includes specific guidance for eight AI-specific risk areas including:

1. Fairness and transparency in profiling.
2. Accuracy.
3. Fully automated decision-making models.
4. Security and cyber.
5. Trade-offs.
6. Data minimisation and purpose limitation.
7. Exercise of rights.
8. Impact on broader public rights.[61]

Similarly, the Medicines and Healthcare products Regulatory Agency (MHRA) is working to improve trust in AI by updating the regulations that apply to software and AI when used as a medical device. Through these updated regulations, the MHRA requires that an assurance process be completed for a product to be licensed to market. The new regulations give

[60] 2021, Harini Suresh, John V. Guttag, 'A Framework for Understanding Sources of Harm throughout the Machine Learning Life Cycle', https://arxiv.org/abs/1901.10002

[61] 2019, Information Commissioner's Office (ICO), 'An overview of the Auditing Framework for Artificial Intelligence and its core components', https://ico.org.uk/about-the-ico/news-and-events/ai-blog-an-overview-of-the-auditing-framework-for-artificial-intelligence-and-its-core-components/

guidance for AI assurance processes, and also help to give insights into how to interpret regulatory requirements to conform to them.[62]

These two examples demonstrate regulators using their powers to guide AI assurance practices and to give guidance about new regulations.

Sharing data with the public

Assurance of data and data practices should not be thought of as only concerning the sharing of data between organisations. The public, private and third sectors all routinely share data and information with the public, as the ODI's Data Spectrum illustrates.[63] The way this data is collected and shared with the public can also be the subject of data assurance schemes.

One obvious example is the standardisation of product labelling which is overseen by the Food Standards Agency (FSA). See Case study 3 for more detail.

[62] 2021, CDEI, 'The roadmap to an effective AI assurance ecosystem', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1039146/The_roadmap_to_an_effective_AI_assurance_ecosystem.pdf

[63] ODI, 'The Data Spectrum', https://theodi.org/about-the-odi/the-data-spectrum/

Case study 3: Product labelling for health and safety

What the intervention is:

In the UK there are legal requirements that apply to the packaging and labelling of food, requiring businesses to make specific information available to consumers.[64] These rules and regulations are managed by the Food Standards Agency (FSA).

How it works:

The laws applying to food labelling have requirements to include specific information for consumers, such as the name of the food, listing ingredients in order of weight, declaring the percentage of each ingredient included,[65] as well as additional requirements relating to the labelling of common allergens.[66] Other rules and regulations describe necessary storage conditions, when 'best before' or 'use by' dates are provided, as well as information on the location of origin of food.

Standardising what data is available to a consumer increases safety, builds trust and improves decision making. Multiple studies have analysed the impact that food labelling has on the choices consumers make at the supermarket.[67,68,69]

To help businesses comply with these rules, the FSA provides free online training. The FSA also produces a range of guidance on food labelling, and conducts outreach with other professional bodies and groups.[70]

[64] 2021, Food Standards Agency (FSA), 'Packaging and labelling', https://www.food.gov.uk/business-guidance/packaging-and-labelling

[65] 2015, Department for Environment, Food and Rural Affairs (Defra) and FSA, 'Food labelling: giving food information to consumers', https://www.gov.uk/guidance/food-labelling-giving-food-information-to-consumers#give-a-quantitative-ingredients-declaration-quid

[66] 2021, FSA, 'Packaging and labelling', https://www.food.gov.uk/business-guidance/packaging-and-labelling

[67] UK Research and Innovation (UKRI), 'FLICC: Front of pack food labelling: impact on consumer choice', https://gtr.ukri.org/project/86FC4B0A-7D11-4C78-BFB4-1F8D9EB9AD5B#/tabOverview

[68] 2015, Peter Scarborough et al., 'Reds are more important than greens: how UK supermarket shoppers use the different information on a traffic light nutrition label in a choice experiment', http://dx.doi.org/10.1186/s12966-015-0319-9

[69]FSA, 'Consumers and allergen labelling literature review', https://www.food.gov.uk/research/food-allergy-and-intolerance-research/consumers-and-allergen-labelling-literature-review

[70]2021, British Standards Institution (BSI), 'Marketing and labelling safe, nutritious food and drink in 2021 webinar', https://www.bsigroup.com/en-GB/industries-and-sectors/food-and-drink/food-events-and-webinars/advertising-and-labelling-safe-nutritious-food-and-drink-in-2021-webinar/

The FSA also carries out enforcement to check that food meets necessary standards,[71] and works with the National Food Crime Unit (NFCU) works to prevent, detect, and investigate food crimes in the UK.[72]

By placing requirements on what data must be published, the FSA has also created requirements on the data the food businesses and their wider supply chain must collect, manage and share.

Manufacturers need to identify whether specific allergens may be included in a specific food and use common labelling (taxonomies) to describe them. Manufacturers must track the specific dates when an item was made to issue use by and sell by dates. They must also record and monitor the percentage of ingredients included in their products on an ongoing basis, to ensure the accuracy of their data provided on packaging.

Compliance with the broader regulation has an impact on the data ecosystem that supports production and distribution of food in the UK.

Key insights:
This example illustrates that existing regulatory interventions already include elements of data assurance.

Any regulation around statutory reporting or provision of data involves some element of data standardisation even if the requirements themselves do not directly relate to machine-readable data, or its automated consumption and sharing.

Data assurance activities used:

- Development of common standards, including taxonomies and information reporting guidelines.
- Training to support adoption of data standards.
- Compliance testing to assess data quality and accuracy, for example checking that ingredient lists are accurate.
- Legal requirements to drive and enforce compliance to data standards.
- Training programmes to guide adoption of and adherence to legal data sharing requirements.

[71] 2020, FSA, 'The food regulatory system', https://www.food.gov.uk/about-us/the-food-regulatory-system
[72] 2021, FSA, 'Food crime', https://www.food.gov.uk/safety-hygiene/food-crime

Regulatory sandboxes

Some regulators are creating trusted environments for sharing data such as regulatory sandboxes.

A regulatory sandbox is 'a safe testing space where participants can test their new business model, innovative products, services and delivery mechanisms without immediately incurring all the normal regulatory consequences of engaging in the activity in question.'[73]

Regulatory sandboxes give organisations in regulated sectors confidence in how to interpret regulatory requirements, and how those regulatory requirements apply in new use cases. They also allow regulators to test the interpretation, implementation and enforcement of regulatory requirements around new use cases, and give regulators the confidence to execute their regulatory mandate without unnecessarily inhibiting innovation.

The UK has pioneered regulatory sandboxes to help spur innovation and increase market competition. Regulatory sandboxes play a role in data assurance by creating the conditions to innovate with data in a trustworthy manner.[74] They help to fuel business and job creation and secure future scientific breakthroughs by making data more accessible and usable while at the same time protecting data rights and ensuring data is used responsibly.

An example of a regulatory sandbox is the ICO's Regulatory Sandbox, which supports organisations creating products or services that use personal data safely and in an innovative manner.[75] The Financial Conduct Authority's (FCA) Digital Sandbox looks to help firms in the UK financial services sector develop innovative products in a safe environment,[76] while the Civil Aviation Authority (CAA) Regulatory Sandbox provides an environment for companies in the aviation industry to safely test 'innovative solutions'.[77]

Researchers have analysed the effects of regulatory sandboxes, particularly in the finance sector. Some researchers have found that regulatory sandboxes had positive impacts on the growth of fintech venture investment,[78] while others have argued that it is still too early to know with certainty the effect of regulatory sandboxes on financial innovation.[79] The ODI's roundtable

[73] 2020, Jayoung James Goo, Joo-Yeun Heo, 'The impact of the regulatory sandbox on the fintech industry, with a discussion on the relation between regulatory sandboxes and open innovation', https://doi.org/10.3390/joitmc6020043

[74] 2021, ODI, 'How can regulators tackle challenges through innovative uses of data?' https://theodi.org/article/how-can-regulators-tackle-challenges-through-innovative-uses-of-data/

[75] ICO, 'Regulatory Sandbox', https://ico.org.uk/for-organisations/regulatory-sandbox/

[76] Financial Conduct Authority (FCA), 'Digital Sandbox', https://www.fca.org.uk/firms/innovation/digital-sandbox

[77] Civil Aviation Authority (CAA), 'Regulatory challenges for innovation in aviation', https://www.caa.co.uk/our-work/innovation/regulatory-challenges-for-innovation-in-aviation/

[78] 2020, Jayoung James Goo, Joo-Yeun Heo, 'The impact of the regulatory sandbox on the fintech industry, with a discussion on the relation between regulatory sandboxes and open innovation', https://doi.org/10.3390/joitmc6020043

[79] 2020, Christopher C. Chen, 'Regulatory Sandboxes in the UK and Singapore: a preliminary survey', https://dx.doi.org/10.2139/ssrn.3448901

report *Innovation and the data economy: opportunities for UK regulators* conducted with Better Regulation Executive (BRE) touched on the success stories from UK regulatory sandboxes, showing that they can help regulators to innovate even when working with sensitive data.[80] *The Kalifa Review* expressed that the FCA's regulatory sandbox has led the way in creating a system that protects customers and 'creates an enabling environment that encourages growth and competition'.[81] The review proposed the creation of a permanent regulatory sandbox in the finance sector.

Creating trusted data ecosystems

Trustworthy data ecosystems are environments in which individuals and organisations across the public, private and third sector trust that data is flowing in ways that will maximise benefits whilst minimising harms.

In the finance industry, regulators have worked with industry bodies to develop data standards. In 2021, the industry bodies Pensions Administration Standards Association (PASA), the Pensions and Lifetime Saving Association and the Association of British Insurers worked with the Pensions Regulator and the FCA to develop standards for data matching in pension dashboards.[82] The industry bodies and regulators worked together to determine the best set of personal data to use to match individuals with their pensions and to ensure that data is aligned with pension regulation. PASA also offers updated guidance on data management to members.[83] This data assurance scheme helps facilitate the trustworthy sharing and use of data in the pensions data ecosystem. It helps to make data more accessible and usable while ensuring it is used responsibly.

Regulators can also play a role in facilitating the creation of trusted data ecosystems through collaborative projects. These projects help build trust between organisations.

Ofgem and Ofwat are working together with the industry bodies Water UK, the Energy Networks Association, and the UK Regulators Network in a data access initiative to share non-financial data about vulnerable water and energy customers to help utility companies better serve them. A pilot programme has already been completed successfully, and non-financial data sharing between United Utilities and Electricity North West has been implemented permanently.

As part of their strategic plan, UKRN has expressed that they,

[80] 2021, ODI, 'How can regulators tackle challenges through innovative uses of data?' https://theodi.org/article/how-can-regulators-tackle-challenges-through-innovative-uses-of-data/
[81] 2021, HM Treasury, 'The Kalifa Review of UK FinTech', https://www.gov.uk/government/publications/the-kalifa-review-of-uk-fintech
[82] 2021, Maria Espadina, 'Industry bodies to set data standards for dashboards', Financial Times Adviser, https://www.ftadviser.com/pensions/2021/07/29/industry-bodies-to-set-data-standards-for-dashboards/
[83]'Data (DataWG)', PASA, https://www.pasa-uk.com/guidance/data/

> will facilitate information and best practice sharing on regulatory approaches or requirements on how firms record and use information about consumers with additional needs and to encourage collaboration on the responsible use and sharing of data.[84]

The UKRN are supporting a data access initiative and then producing data sharing best practices, a 'soft' form of data assurance.

Regulators are also working with industry bodies to incorporate data assurance into wider regulatory interventions aimed at increasing competition within sectors. An example of this is Open Banking, an intervention led by the Competition and Markets Authority (CMA) which also involved the FCA. The CMA directed the industry to create a new body to act as steward for the new open banking standards. See Case study 4 for more detail.

[84] 2021, UK Regulators Network, 'UKRN annual report and multi-year workplan', https://www.ukrn.org.uk/wp-content/uploads/2021/03/UKRN-workplan-and-annual-review-2021-for-publication-.pdf

Case study 4: Open banking

What the intervention is:

In 2016, the CMA published the findings of a market investigation into retail banking.[85] In their investigation, they found the UK financial services market lacked competition, leading to lower-quality services for customers. To help increase competition, the FCA and the CMA worked together to create a data assurance scheme to share financial data safely and securely. This is the CMA's *Open Banking remedy* regulation.[86] The CMA required the development and adoption of shared standards for financial data, datasets and for gaining access to data. The work to implement the standards was then delegated to a new industry body, the Open Banking Implementation Entity (OBIE). This entity has the responsibility to coordinate the development and adoption of open banking in the UK.

How it works:

Open banking requires the UK's nine largest banks and building societies to make certain data available under consent, while smaller banks can share data voluntarily.[87] The Open Banking Standards govern how that data is provided technically, from the contents to the dataset.[88] Open banking standards also cover the content of Third Party Providers (TPPs), and the entire open banking ecosystem. The Open Banking Implementation Entity (OBIE) maintains the open banking standards.[89]

TPP's must be approved and authorised by the FCA. The FCA regulates payment service providers and the OBIE provides tools for service providers to test whether they meet the open banking standard.[90] The OBIE conducts certifications of payment service providers to determine that they meet functional, security profile, customer experience, and operational standards. Open banking incorporates the data assurance activities of the development and adoption of standards for data. These standards describe how data is collected, accessed, used and shared in the finance sector.

[85] 2016,Competition and Markets Authority (CMA), 'Retail banking market investigation', https://www.gov.uk/cma-cases/review-of-banking-for-small-and-medium-sized-businesses-smes-in-the-uk

[86] Open Banking Implementation Entity (OBIE), 'CMA publishes approved Roadmap for the final stages of Open Banking implementation', https://www.openbanking.org.uk/news/cma-publishes-approved-roadmap-for-the-final-stages-of-open-banking-implementation/

[87] OBIE, 'Questions about Open Banking', https://www.openbanking.org.uk/faqs/

[88] OBIE, 'Welcome to the open banking standard', https://standards.openbanking.org.uk/

[89] OBIE, 'About the OBIE', https://www.openbanking.org.uk/about-us/

[90] OBIE, 'Design and testing', https://standards.openbanking.org.uk/operational-guidelines/dedicated-interface-requirements/design-and-testing/v3-1-3/

Key insights:
Open banking illustrates how regulatory interventions can incorporate data assurance. In particular, because the CMA's open banking initiative concerned increasing access to data, this regulatory intervention needed to incorporate data assurance activities so that all stakeholders would trust in the solution. As part of this data assurance scheme, the CMA determined that no existing body in the finance sector could take responsibility for developing and maintaining the open banking standards. The CMA, therefore, recommended the creation of a new industry body and data institution[91] to help improve market competition. This may be an intervention for the assurance of data and data flows replicated in other sectors.

Since the implementation of open banking, UK consumers have become more comfortable with sharing their financial data with third parties. The Open Banking Implementation Entity announced in September 2020 that the number of people using products enabled by open banking had doubled since January 2020 to more than 2 million people.[92] As of November 2021, the UK government estimated that more than half of UK small businesses and more than 4 million consumers use technology powered by open banking.[93]

Data assurance activities used:

- Development and adoption of data holding, use and sharing standards.
- Development of an industry body to guide the introduction of data holding, use and sharing standards.
- Accreditation of actors involved (banks, TPPs and customers) in the holding, use and sharing of data.
- Regulation of service providers in the holding, use and sharing of data.
- Certification of service providers in the holding, use and sharing of data by the OBIE.

The success of Open Banking has prompted regulators in other sectors to consider similar data access initiatives that build trust.

[91] ODI, 'Data institutions', https://theodi.org/project/rd-data-institutions/#:~:text=Data%20institutions%20are%20organisations%20whose,public%2C%20educational%20or%20charitable%20aims.&text=combining%20or%20linking%20data%20from,those%20that%20have%20contributed%20data.

[92] OBIE, 'Real demand for open banking as user numbers grow to more than two million', https://www.openbanking.org.uk/news/real-demand-for-open-banking-as-user-numbers-grow-to-more-than-two-million/

[93] 2021, CMA, 'Corporate report: update on open banking', https://www.gov.uk/government/publications/update-governance-of-open-banking/update-on-open-banking

According to the Department for Business, Energy and Industrial Strategy (BEIS) the next UK sector to have a 'live data sharing ecosystem' will be the UK energy sector as part of the Open Energy Initiative.[94] As part of Open Energy, Ofgem is developing open data best practices.[95] Non-profit Icebreaker One is working on standards for data licensing and metadata to help support the move to net zero.[96] The Icebreaker One Trust Framework includes a set of principles designed to help data flow more easily among organisations.[97]

Accreditation for data assurance

Regulators, industry bodies, and professional bodies are all involved in different types of accreditation and certification practices to improve data assurance.

The ICO has developed a GDPR certification scheme which functions through a collaboration between multiple bodies.[98] The ICO publishes the accreditation requirements for certification bodies to meet, approves and publishes certification criteria, and maintains a public register of approved GDPR certification schemes. The United Kingdom Accreditation Service (UKAS) carries out accreditations of the bodies responsible for conducting GDPR certifications. Other marketplaces built around assurance practices function in a similar way. See Case study 1 for more information.

Developing best practice

In other cases, industry bodies and professional bodies set data assurance best practice themselves and develop their own frameworks, guidance, and certification schemes to drive adherence.

The professional bodies the Royal Statistical Society, British Computer Society, The Chartered Institute for IT, The Operational Research Society, The Institute of Mathematics and Its Applications, along with the National Physical Laboratory standards institute, the Alan Turing Institute, and with the support of the Royal Academy of Engineering and the Royal Society formed an alliance in 2020 to develop 'new professional standards for data science'.[99] The

[94] 2021, Department for Business, Energy and Industrial Strategy,' Smart Data Research report: third party accreditation', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/993342/smart-data-research-accreditation.pdf

[95] 2021, Office of Gas and Electricity Markets (Ofgem), 'Data best practice guidance', https://www.ofgem.gov.uk/sites/default/files/2021-11/Data_Best_Practice_Guidance_v1.pdf

[96] Icebreaker One, 'Open Energy', https://energy.icebreakerone.org/

[97] Icebreaker One, 'What is Icebreaker One doing?' https://icebreakerone.org/

[98] ICO, 'Certification', https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/accountability-and-governance/certification/

[99] 2021, Royal Statistical Society (RSS), 'Alliance formed to create new professional standards for data science', https://rss.org.uk/news-publication/news-publications/2021/general-news/alliance-formed-to-create-new-professional-standar/

alliance is working to develop 'industry-wide standards' that will 'look to address current issues, such as data breaches, the misuse of data in modelling and bias in artificial intelligence.' They intend to 'give people confidence that their data is being used ethically, stored safely and analysed robustly'.[100]

This example shows professional bodies working together in a data assurance scheme to develop industry-wide best practice. The UK's *National Data Strategy* says that the UK Government hopes to build on the work of these professional bodies to 'publish a working definition of data skills for the wider economy, set out a clear distinction between data skills, digital skills and AI skills, and consider the benefits of providing information on pathways into data-related careers'.[101]This initiative ties in with larger trends, part of the changes following the UK's exit from the EU to create more consistency among UK data professionals. The ODI's *Data Skills Framework* demonstrates that data skills are broader than just data science; instead, organisations need both technical data skills as well as skills that enable data innovation.[102]

Certification for data assurance

Multiple bodies engage in data assurance schemes involving certification.

The industry body Electronic Data Management (EDM) Council sets best practice for data management with their Data Management Capability Assessment Model (DCAM) framework and for data management in the cloud with their Cloud Data Management Capabilities (CDMC) framework.[103] Companies can receive CDMC certification based on an assessment conducted by an independent partner. The EDM Council also offers CDMC training courses for individuals. Individuals can take a training course which is followed up with an exam to receive a DCAM professional development certification. Organisations can use the EDM Council's self-assessment tool to evaluate their own internal data management capacity, or receive a formal assessment from a certified authorised partner.[104] The EDM Council has stepped in to create these frameworks to help develop best practice and educate members in how to conform to it.

The BSI offers multiple training courses and certifications in privacy and data protection, and several which specifically pertain to General Data Protection Regulation (GDPR). Some of the courses and certifications offered include the EU General Data Protection Regulation (GDPR)

[100] 2021, Royal Statistical Society (RSS), 'Alliance formed to create new professional standards for data science',
https://rss.org.uk/news-publication/news-publications/2021/general-news/alliance-formed-to-create-new-professional-standar/
[101] 2020, DCMS, 'National Data Strategy',
https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy#data-2-5
[102] 2020, ODI, 'Data Skills Framework',
https://theodi.org/article/data-skills-framework/#1561999285194-8defa99a-39a52584-8cfe
[103]Electronic Data Management (EDM) Council, 'Training and ELearning',
https://edmcouncil.org/page/TrainingOverview
[104] EDM Council, 'DCAM assessments and support',
https://edmcouncil.org/page/dcamassessmentsuppor

Foundation training course, which gives an overview of GDPR legislation in one day, the GDPR auditor/self assessment, designed to help stakeholders within organisations ensure that their organisation is GDPR compliant, the Certified Information Privacy Professional Europe training course and certification, the Certified Information Privacy Manager training course and certification, and the Certified Data Protection Officer training course and certification, among others.[105]

The professional body Data Management Association (DAMA) similarly offers the Certified Data Management Professional (CDMP) certification.[106] See Case study 5 for more detail.

Case study 5: Data management professional certification

What the intervention is:
DAMA is an international professional body for data management professionals organised into local, country-specific chapters.[107] DAMA offers the Certified Data Management Professional (CDMP) certification.[108] Data professionals in any sector can become certified and employers internationally accept and seek out CDMP certification.

How it works:
In order to become certified in CDMP, a data management professional must pass the Data Management Fundamentals exam offered by DAMA. The exam tests an individual's knowledge of the most recent version of the DAMA Data Management Body of Knowledge.[109]

The exam covers specific areas of knowledge, along with data management process, ethics, and Big Data. There are four different levels of certification (associate, practitioner, master, and fellow) that represent different levels of career development for a data professional. In order to receive the practitioner or master levels of certification, data professionals must pass two specialist exams along with the Data Management Fundamentals exam.

Those who become certified must pay a yearly certification fee and complete a 3 year cycle where they continue educational and professional development activities.[110] The CDMP certification tests practitioners in the assurance of data practices and also in data being fit for purpose.

[105] BSI, 'Privacy and data protection training', https://www.bsigroup.com/en-GB/our-services/training-courses/Data-Protection/
[106] Data Management Association (DAMA), 'Certified Data Management Professionals', https://cdmp.info/
[107] DAMA, 'Home', https://www.dama.org/cpages/home
[108] DAMA, 'Certified Data Management Professionals', https://cdmp.info/
[109] DAMA, 'DMBoK v2x Initiative', https://www.dama.org/cpages/forthcoming-dmbok-2x-initiative
[110] DAMA Dach, 'Certified Data Management Professional (CDMP) Certification Program', https://damadach.org/certified-data-management-professional-cdmp-certification-program/

Key insights:
DAMA’s CDMP certification demonstrates a professional body undertaking a data assurance activity and their work has an impact in multiple sectors. The certification benefits data professionals, who can use it to demonstrate their expertise to potential employers and support their career development. It benefits organisations that employ data professionals as it helps them to ensure that the employees they hire have the appropriate knowledge to handle organisational data in a trustworthy manner. The certification also benefits the ecosystem of data practitioners because it helps to create a common reference and set of best practices for the profession. Other professional bodies can follow DAMA’s lead in adopting certification schemes for data assurance.

Data assurance activities used:

- Development and maintenance of DMBoK, data management code of practice.
- Certification of individuals in their knowledge of the data management code of practice.

4.3 Opportunities and challenges for data assurance

Our research and interviews have highlighted a range of challenges and opportunities that might hinder or support the introduction of data assurance across different sectors.

Rapidly changing regulatory context

Following its exit from the EU, the UK is going through a period of wide-ranging regulatory reforms. These reforms are rapidly changing the regulatory and policy context for the access, use and sharing of data.

Multiple recent UK policy documents emphasise changing regulatory remits to facilitate innovation. The *Regulation for the Fourth Industrial Revolution* white paper makes this explicit, stating 'we need a more agile approach to regulation, that supports innovation while protecting citizens and the environment'.[111] *Build Back Better: our plan for growth* similarly discusses changes to the UK's regulatory system to support innovation by 'easing the regulatory compliance red tape burden on business'.[112] It includes plans for a new digital strategy, innovation strategy, and Better Regulation Committee, along with a Taskforce on Innovation, Growth and Regulatory Reform (TIGRR).

TIGRR looks at the opportunities arising from the UK's exit from the EU and considers changes to regulation going forward.[113] One of the most significant proposed reforms concerning data is to replace GDPR with a new framework for data protection, which would 'give people greater control over their data while allowing data to flow more freely and drive growth across healthcare, public services and the digital economy'.[114] Other proposals include positioning the UK as a regulation standard setter. These reforms will not necessarily be taken forward precisely as expressed in the report, however.

The *Data: a new direction* consultation is also looking into reforms in the use of data in the

[111] 2019, BEIS, 'Regulation for the Fourth Industrial Revolution', https://www.gov.uk/government/publications/regulation-for-the-fourth-industrial-revolution/regulation-for-the-fourth-industrial-revolution
[112] 2021, HM Treasury, 'Build Back Better: our plan for growth', https://www.gov.uk/government/publications/build-back-better-our-plan-for-growth
[113]2021, Prime Minister's Office, 10 Downing Street, 'Taskforce on Innovation, Growth and Regulatory Reform independent report', https://www.gov.uk/government/publications/taskforce-on-innovation-growth-and-regulatory-reform-independent-report
[114] 2021, UK Government, 'Taskforce on Innovation, Growth and Regulatory Reform independent report', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/994125/FINAL_TIGRR_REPORT__1_.pdf

UK,[115] including changes to the ICO's remit,[116] introducing changes or alternatives to GDPR, reducing barriers for responsible innovation, and removing barriers to cross-border data flows.[117] TIGRR has influenced the proposals in *Data: a new direction*. Though these policy documents include similar themes of changing regulatory remits to encourage innovation, the 'new direction' for data in the UK still remains somewhat uncertain.

The *Benefits of Brexit* report further highlights regulatory opportunities pertaining to data, with plans to reform 'our data laws and' set a 'new direction for data regulation'.[118] The *National Data Strategy* also emphasises greater data availability to encourage innovation and growth by alterations to regulatory roles and data protection laws, among other potential changes.[119] The UK's *Innovation Strategy* expresses an aim to create 'the world's most agile regulatory system' by reforming some regulatory standards deemed 'overly restrictive,'[120] while the *National AI Strategy* highlights that the government hopes to create a 'progressive' regulatory environment that will support innovation and 'keep pace with the fast changing demands of AI'.[121]

As a result of these reforms, the powers and remit of UK regulators may change and key legislation and policies are under review. This changing regulatory context presents an opportunity. The UK may move towards a looser regulatory framework with fewer protections. *Global Britain in a Competitive Age: the integrated review of security, defence, development and foreign policy* expresses an ambition for the UK to be a global data and digital services hub by 'working with our international partners to overcome barriers to data flows and promote international data standards that enable growth and innovation'.[122] With a potential loosening of the UK's regulatory framework, it is likely that there will be a greater role for data assurance in

[115]2021, DCMS, 'Data: a new direction', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1022315/Data_Reform_Consultation_Document__Accessible_.pdf
[116] 2021, DCMS, 'Data: a new direction', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1022315/Data_Reform_Consultation_Document__Accessible_.pdf
[117] 2021, ODI, 'Data: a new direction: Open Data Institute response', https://docs.google.com/document/d/1DUN51AR57gDUS3Ck2hg2zmuYIGBq0s9ikx837ws7wh8/edit#
[118] 2022, HM Government, 'The Benefits of Brexit: How the UK is taking advantage of leaving the EU', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1052148/benefits-of-brexit-document.pdf
[119] 2020, DCMS 'National Data Strategy', https://www.gov.uk/government/publications/uk-national-data-strategy/national-data-strategy#data-2-5
[120] 2021, BEIS, 'UK Innovation Strategy: leading the future by creating it', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1009577/uk-innovation-strategy.pdf
[121] 2021, UK Government, 'National AI Strategy', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1020402/National_AI_Strategy_-_PDF_version.pdf
[122] 2021, Cabinet Office, 'Global Britain in a Competitive Age: the Integrated Review of Security, Defence, Development and Foreign Policy', https://www.gov.uk/government/publications/global-britain-in-a-competitive-age-the-integrated-review-of-security-defence-development-and-foreign-policy/global-britain-in-a-competitive-age-the-integrated-review-of-security-defence-development-and-foreign-policy

building trust.[123]

This could, however, also create a challenging context for introducing new data assurance schemes, particularly 'harder' data assurance activities that may rely on existing powers and remits that could be subject to change. Furthermore, any change in the regulatory framework for data risks diminishing trust in data practices. Assurance, therefore, will be a valuable tool in maintaining confidence in data practices and should be an essential part of implementing these reforms. Depending on the direction taken by reforms to data protection law, data assurance might need to play a greater role in maintaining public trust in the data ecosystem.

Industry bodies and professional bodies are also facing contextual shifts spurred by increasing globalisation, the shift to online work accelerated by the pandemic, and the shift to job automation, among other changes. There is an ongoing process of data-driven disruption changing the way that many industries operate. The report *Review of the UK business to business data assurance market* articulates that the biggest opportunities and needs for data assurance are in data-driven products and services, new data ecosystems, markets and new use cases.[124] There is an opportunity, therefore, for industry bodies and professional bodies to support their members in this changing context by adopting data assurance schemes to help create more trust in data ecosystems by helping make data more accessible and usable while protecting data rights.

Standards-based interventions

Standards can help to increase access to data, making it easier to use and share in a trustworthy manner. There are many different types of standards,[125] which cover not only technical aspects of data sharing but also standards for data governance, data practices, as well as audit and certification. Creating and driving the adoption of standards is an important part of all forms of assurance and can help ensure that data is used responsibly.

Regulators have an opportunity to use their powers to require the adoption of standards. Regulators can make meeting certain standards a condition of licensing products. An example of this is the MHRA which requires AI and software used in medical devices to meet specific standards.

[123] 2021, ODI, 'Data: a new direction, Open Data Institute response', https://docs.google.com/document/d/1DUN51AR57gDUS3Ck2hg2zmuYIGBq0s9ikx837ws7wh8/edit#heading=h.f22474fe7t8c

[124] 2021, Frontier Economics and the ODI, 'Review of the UK business to business data assurance market', https://theodi.org/wp-content/uploads/2021/07/PRS-ODI-Data-assurance-FINAL_23072021.pdf

[125] ODI, 'Types of open standards for data', https://standards.theodi.org/introduction/types-of-open-standards-for-data/

Regulators can also drive the adoption of standards by introducing them as a requirement for a licence to operate, or as a necessity to receive government contracts. Standards, however, can be difficult to develop and can take time to be fully adopted.[126]

Building data literacy

Our interviews highlighted that some sectors may need extra help in building data literacy. The ODI defines data literacy as the ‘ability to think critically about data in different contexts and examine the impact of different approaches when collecting, using and sharing data and information’[127] and in this way, it goes beyond data science. There is an opportunity for regulators, industry bodies and professional bodies to collaborate with other organisations with data assurance expertise to help these three types of organisation build data literacy or collaborate to pool resources.

Some smaller regulators lack the resources to be able to further explore data assurance schemes, a representative from the UK Regulators Network expressed. Other interviewees echoed similar concerns about industry bodies and professional bodies.

Some regulators, industry bodies and professional bodies, therefore, may not have the resources to be able to design certification schemes, develop standards, or revise professional training to include aspects of data assurance. Interviews highlighted that some organisations, in particular professional bodies, have hesitations about sharing their own data or uncertainty about the benefits. The ODI’s Data Literacy programme is an example of the type of initiative organisations with expertise in data assurance can implement to address this challenge.[128] It supports organisations through upskilling, training programmes and engagement with company directors to improve data literacy across sectors.[129]

Varying maturity of sectors

Our research found that different sectors are at different stages of data and digital maturity.

The *Insight report on sharing engineering data*, produced by the ODI and Lloyd's Register Foundation discussed the differing levels of data sharing maturity across different sectors.[130]

The report found that the finance sector has undertaken some of the most substantial data access initiatives with the implementation of Open Banking, and there has also been work

[126] ODI, 'Open standards for data', https://standards.theodi.org/

[127] 2021, ODI, ‘Data literacy: what is it and how do we address it at the ODI?’, https://theodi.org/article/data-literacy-what-is-it-and-how-do-we-address-it-at-odi/

[128] 2021, ODI, 'Introducing the ODI’s data literacy programme', https://theodi.org/article/introducing-the-data-literacy-programme/

[129] 2021, ODI, 'Introducing the ODI’s data literacy programme', https://theodi.org/article/introducing-the-data-literacy-programme/

[130] 2019, Leigh Dodds, et al., 'Insight report on sharing engineering data', https://www.lrfoundation.org.uk/en/publications/insight-report-on-sharing-data/

undertaken in the agriculture sector to adopt data access initiatives.[131] Interviewees also highlighted similar work undertaken in the utilities sectors, especially with the Open Energy initiative.[132]

Regulators can also follow the model of the ICO, FCA, CAA and others in creating regulatory sandboxes. Regulators in specific sectors can also collaborate with cross-sector regulators like the ICO to tailor existing schemes such as the AI auditing framework or the GDPR certification scheme for their own use rather than designing similar data assurance schemes from scratch.

There is also an opportunity for different sectors to learn from each other. Sectors can, and are, adopting common models and approaches to help them develop their data ecosystems and increase trust. The Open Energy[133] initiative is a good example of this. Drawing on the success of Open Banking, the Open Energy initiative concerns 'opening up data in the UK energy sector to help create a fairer energy market'.[134] Ofgem is developing open data best practices,[135] while non-profit Icebreaker One is working on standards for data licensing and metadata to help support the move to net zero.[136] The Icebreaker One Trust Framework includes a set of principles designed to help data flow more easily among organisations.[137]

The varying maturity of sectors, however, may make it difficult to implement data assurance schemes and learn from lessons elsewhere. Sectors may have to first address underlying issues, such as a lack of data literacy, general fears about data sharing, and concerns about the capacity to share data. The first round of the BRE's Regulators' Pioneer Fund competition focused on addressing challenges related to data and AI and is a good example of an initiative to help address the varying maturity of sectors.[138] The competition was intended to help develop the maturity and confidence of regulators in those areas and help to address the four 'grand challenges'' highlighted in the 2017 *Industrial Strategy*.[139] Some of the initiatives funded included regulatory sandboxes to facilitate trustworthy innovation.

[131] 2019, Leigh Dodds, et al., 'Insight report on sharing engineering data', https://www.lrfoundation.org.uk/en/publications/insight-report-on-sharing-data/

[132] 2021, BEIS, 'Smart Data Research report: third party accreditation', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/993342/smart-data-research-accreditation.pdf

[133] Icebreaker One, 'Open Energy', https://energy.icebreakerone.org/report-meda-3/

[134] ODI, 'R&D: Open standards for the UK energy sector', https://theodi.org/project/open-standards-for-the-uk-energy-sector/#1558340061441-cb8100c5-47f3

[135] 2021, Office of Gas and Electricity Markets (Ofgem), 'Data best practice guidance', https://www.ofgem.gov.uk/sites/default/files/2021-11/Data_Best_Practice_Guidance_v1.pdf

[136] Icebreaker One, 'Open Energy', https://energy.icebreakerone.org/

[137] Icebreaker One, 'What is Icebreaker One doing?' https://icebreakerone.org/

[138] 2021, BEIS, 'Evaluation of the Regulators' Pioneer Fund (round 1) - Main Findings Report', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/966974/evaluation-of-rpf-main-report.pdf

[139] 2021, BEIS, 'Evaluation of the Regulators' Pioneer Fund (round 1) - Main Findings Report', https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/966974/evaluation-of-rpf-main-report.pdf

Existing levers and powers can be repurposed

All three types of organisation have both 'harder' and 'softer' levers that they can use to drive changes in behaviour. Regulators can use legislation to require the adoption of data assurance activities. Industry bodies and professional bodies could require members to conform to certain criteria for membership or to meet specified professional standards.

Though data assurance may require some new standards or technical approaches, the fundamentals are not new.

New regulators or bodies are not necessarily required to implement data assurance schemes in a sector. Instead, all currently existing bodies have a role to play in driving the creation of trustworthy data ecosystems.

Commercial opportunities

In their report, *Review of the UK business to business data assurance market*, the ODI and Frontier Economics identified potential gaps in the data assurance market.[140] While data assurance firms are over-represented in the computer programming and management consulting sectors, they are under-represented in other sectors including office administrative and other business support activities, information service activities, professional, scientific, and technical activities, and publishing activities. There is an opportunity, therefore, for these three types of organisations to fill in the gaps in the UK data assurance market.

For professional bodies, these gaps might offer a commercial opportunity. By offering training and certification programmes that touch on aspects of data assurance, professional bodies can differentiate themselves from others in their sector and gain an advantage.

Regulators have an opportunity to stimulate and create markets around assurance, as CDEI expressed in *The roadmap to an effective AI assurance ecosystem*[141] and as illustrated in the Cyber Essentials scheme outlined in Case study 1. This will help create new commercial opportunities whilst increasing trust, helping to fuel business and job creation. Creating a marketplace for auditors allows the data assurance scheme to scale to meet demands of applicants. Incentives such as procurement requirements can drive the adoption of data assurance schemes independent of legal or regulatory requirements.

[140] 2021, Frontier Economics and the ODI, 'Review of the UK business to business data assurance market',
https://theodi.org/wp-content/uploads/2021/07/PRS-ODI-Data-assurance-FINAL_23072021.pdf
[141] 2021, CDEI, 'The roadmap to an effective AI assurance ecosystem',
https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/1039146/The_roadmap_to_an_effective_AI_assurance_ecosystem.pdf

Support in designing and implementing data assurance schemes

Our research highlighted that regulators, industry bodies and professional bodies could benefit from guidance and support in designing the data assurance schemes to address the challenges they hope to solve. They could also benefit from support in creating confidence in data and ensuring it is fit for purpose. Developing this guidance would help organisations estimate the costs and effort involved, and understand the return on investment for different data assurance activities. Organisations with expertise in data assurance, therefore, have an opportunity to develop this guidance and offer support to these three types of organisations.

5. How can regulators, industry bodies and professional bodies create conditions for the trustworthy sharing, use and reuse of data?

Based on our research, we have developed suggestions for activities that regulators, industry bodies and professional bodies can undertake to help create conditions for the trustworthy sharing, use and reuse of data. The suggestions fall under four broad themes:

- Leading by example.
- Integrating data assurance into existing activities and services.
- Sharing and collaboration to build trust.
- Reviews for effective data assurance schemes.

Leading by example

- **All three types of organisation** can ensure they are leading by example, modelling industry-leading data practices.

  In their report on *Mapping data in the UK government*, the ODI expressed that the UK government can 'unlock the wider value of data across the economy' in part by providing 'examples, resources and skills for others to benefit from, as well as behaviour for others to model'.[142] The ODI's roundtable report *Innovation and the data economy: opportunities for UK regulators* conducted with Better Regulation Executive (BRE) expressed that regulators should demonstrate their own trustworthiness with the data they collect, use, and share.[143] Regulators, industry bodies and professional bodies can take a similar role as models for data assurance. As part of their function, regulators often collect and hold data for their sector. Industry bodies also play a role in stewarding some of the data for their sector. Professional bodies collect and hold membership data, and sometimes other relevant data for their sector as well. With the data that these three bodies regularly collect, they can model best practices for data assurance, helping to ensure data is used responsibly.

- **Regulators** and **industry bodies** can themselves act as data institutions or data intermediaries, taking on a role as trustworthy stewards for data in their sector or across sectors.

  Both regulators and industry bodies often hold data for their sector or across sectors. The HSE acts as a steward for a dataset of places storing potentially hazardous

[142] 2021, ODI, 'Mapping Data in the UK Government', https://theodi.org/wp-content/uploads/2021/10/OPEN_REPORT_Mapping-data-in-UK-government_ODI_2021-10.pdf

[143] 2021, ODI and BRE, 'Innovation and the data economy: opportunities for UK regulators', https://docs.google.com/document/d/12sjzcxHp_o6QkzHr-A6-QEnAAenkzT9oHM-4c5gRFbc/edit#

substances.[144] They do this as part of their role as a regulator and the dataset is made available to the public through a web interface. The industry body UK Finance also acts as a data intermediary by requiring members to agree to share certain data with them, which UK Finance will then, under specific conditions, share more publicly with researchers, universities and others, and use to inform UK Finance's policy work.[145]

Sometimes, however, as in the case of open banking, creating a new organisation to steward data in a sector may be a better option. The decision about whether to use an existing organisation to steward data or to create a new one will depend on the resources, capability and remit of existing organisations and whether they have sufficient trust in the sector to act as an intermediary. While data institutions can play an important role in creating trust,[146] problems have also arisen in some of these institutions.[147] New institutions may not have as robust checks and balances and therefore may be open to abuse. Existing institutions can have similar problems, but can be more easily judged based on their track record. Choosing to create a data institution should be a step taken with deliberation.

Integrating data assurance into existing activities and services

- **Regulators** can consider data assurance as an essential part of the activities they currently undertake.

  The MHRA regulation of AI used in medical devices and the FSA's use of data assurance in food labelling demonstrate that data assurance should be built into multiple regulatory interventions.

- **Regulators** can adopt proven models for trusted sharing of data, like regulatory sandboxes.

  Multiple UK regulators are already undertaking data access initiatives, from the HSE's Discovering Safety Programme[148] to the ICO's Regulatory Sandbox.[149] Regulators in other sectors should also adopt similar proven models to support trust and innovation.

- **Industry bodies** and **professional bodies** can ensure that training and professional development programmes introduce trustworthy data practices.

[144] HSE, 'Control of Major Accident Hazards (COMAH)', https://www.hse.gov.uk/comah/
[145] 2020, UK Finance, 'Privacy policy: how we use and protect your personal data', https://www.ukfinance.org.uk/privacy-policy
[146]2021, ODI, 'What are data institutions and why are they important?' https://theodi.org/article/what-are-data-institutions-and-why-are-they-important/
[147]2021, Katherine Griffiths, 'Open banking chief Imran Gulamhuseinwala resigns over 'bullying culture'', https://www.thetimes.co.uk/article/open-banking-chief-imran-gulamhuseinwala-resigns-over-bullying-culture-rtd5jqrdw
[148] Health and Safety Executive (HSE), 'Discovering Safety', https://www.discoveringsafety.com/
[149] ICO, 'Regulatory Sandbox', https://ico.org.uk/for-organisations/regulatory-sandbox/

Industry bodies and professional bodies already offer a range of training, professional development and certification programmes to help their members learn new skills and further their careers. Training about trustworthy data practices can be added to these programmes relatively easily. By adding modules, these bodies will help to build literacy around trustworthy data practices in their sectors. This is a low-risk intervention that could also potentially help industry bodies and professional bodies gain an advantage. By including these offerings, these bodies can help to differentiate themselves from the others in their sector and to attract members.

- **Industry bodies** and **professional bodies** can make compliance with trustworthy data practices a condition of membership.

  With membership conditions, industry bodies and professional bodies have a lever they can use to encourage the adoption and compliance of trustworthy data practices. Both bodies can then carry out checks to ensure that members are compliant with these conditions.

  If industry bodies and professional bodies do not wish to require their members to comply with specific practices, they could instead recommend their members undertake certifications in trustworthy data practices.

Sharing and collaboration to build trust

- When designing certifications and other standards-based approaches to improve data assurance, **all three types of organisation** can collaborate with multiple stakeholders in development.

  The ODI's research on open standards for data found that standards are most successful when the development includes people with different backgrounds.[150] Successful standards build markets around the standard, and the creators of the standard must take into account the perspectives of the many stakeholders that will use or be affected by the standard. All three types of organisation, therefore, should work with multiple stakeholders when creating standards for data assurance.

- **All three types of organisation** can provide guidance and share best practice to help build trust in data access initiatives.

  Our research highlighted that there is still work to be done in demonstrating the value of data access initiatives in many sectors. As leaders in their sectors, all three types of organisation can work to understand where there are concerns and where data assurance activities may prove useful.

[150] ODI, 'How open standards are developed', https://standards.theodi.org/introduction/how-open-standards-are-developed/

- **All three types of organisation** can collaborate with the researchers assessing emerging issues in machine learning datasets to help review key datasets and decide where further assurance is needed.

  Many academic researchers are currently engaging with data assurance by investigating the quality of datasets used in machine learning algorithms. All three types of organisation should engage with researchers in this field who have conducted research to review the quality of standardised training datasets,[151] and worked to develop frameworks for algorithmic auditing that include aspects of data collection and use.[152] Regulators, professional bodies and industry bodies can identify researchers to engage with through investing in horizon-scanning capacities or by being open to ad-hoc engagement with companies and research organisations operating in their sector, helping to secure new scientific breakthroughs.

- **Regulators** can collaborate with others to develop impactful data assurance schemes.

  Regulatory activities that involve data assurance often involve multiple stakeholders that need to collaborate. The ODI's report on the development of open standards found that standards are most successful when the development process includes people with multiple skills and backgrounds.[153] This is because a variety of people and organisations will be impacted by these regulatory activities. To create the most successful intervention, their perspectives should be taken into account.

  Regulators can use different types of formal knowledge sharing mechanisms to collaborate with other bodies. Some formal knowledge sharing mechanisms including running a program, prototype, or user research investigation. The Energy Data Taskforce, which brought together the UK government, Ofgem and Innovate UK to create a data and digital strategy to help decarbonise the energy sector,[154] is another example of a formal knowledge sharing mechanism that may be effective in other situations as well. The taskforce created a report with a set of recommendations to build upon to develop a data access initiative in the energy sector.

  Through forums like the UKRN, **regulators** can also share successful approaches to introducing new data assurance schemes with each other. Despite their differences in sector and remit, many regulators face similar challenges.

[151] 2021, Inioluwa Deborah Raji, Genevieve Fried, 'About Face: A Survey of Facial Recognition Evaluation', https://arxiv.org/abs/2102.00813
[152] 2020, Inioluwa Deborah Raji et al, 'Closing the AI Accountability Gap: Defining an End-to-End Framework for Internal Algorithmic Auditing', https://arxiv.org/abs/2001.00973
[153] ODI, 'How open standards are developed', https://standards.theodi.org/introduction/how-open-standards-are-developed/
[154] UK Government, 'Energy Data Taskforce', https://www.gov.uk/government/groups/energy-data-taskforce

- **Industry bodies** can convene members to help address sector challenges through data assurance schemes.

  Industry bodies play a convening role in their sector. They can bring members together to seek solutions to common challenges. Some formal knowledge sharing mechanisms including running a program, prototype, or user research investigation. This will help organisations to come to solutions to solve pressing data assurance challenges across the sector.

Reviews for effective data assurance interventions

- **All three types of organisation** can conduct a systematic review of their sector or ecosystem and how best to improve trust before adopting a data assurance scheme.

  'Hard' forms of data assurance activities, such as creating new data standards or producing accreditation, are time consuming and potentially resource intensive to implement. These 'hard' data assurance activities may not always be the most effective or efficient way to improve trust in the holding, using and sharing of data. All three types of organisation should take time before implementing a data assurance scheme to understand their ecosystem and where trust is lacking. All new data assurance schemes should be designed only after this systematic review has taken place.

- Furthermore, for **all three types of organisation**, practical projects can help identify the areas needing most assurance.

  As mentioned in the previous section, regulators, professional bodies and industry bodies are all currently involved in various data assurance schemes. These schemes are highlighting areas where data assurance may be useful. Drawing on existing schemes, all three types of organisation can identify practical areas where it may be useful to implement data assurance activities.

5.1 Next steps

The successful development and implementation of data assurance schemes requires collaboration. Our research has demonstrated that regulators, industry bodies and professional bodies with fewer resources may need help to explore data assurance activities. Researchers and other organisations with data assurance expertise could offer targeted support to these three bodies. Organisations with expertise could convene these bodies and other stakeholders to help design sector and cross-sector data assurance activities like standards and certification schemes, helping to make data more accessible and usable while ensuring it is used responsibly.

These organisations could also use formal knowledge sharing mechanisms such as forums to convene bodies. The Digital Regulation Cooperation Forum, which brings together the CMA, ICO, the Office for Communications (Ofcom) and the FCA, was created to facilitate cooperation among these regulators and to help them address the challenges posed by regulating online platforms.[155] Similar forums could help to enable cross-sector collaboration and facilitate data assurance.

Organisations with data assurance expertise could also conduct further research to determine where complementary data assurance activities are occurring. These activities could support multiple bodies across sectors and this will help to discover where sectors can build on and re-use other work. This would help to develop cross-sector consistency in data assurance. Further research should also explore how the changing regulatory and policy landscape will affect the data assurance activities of the three bodies. Organisations with data assurance expertise should conduct further investigations to address the future impacts of this changing policy context.

[155] 2021, CMA, 'The Digital Regulation Cooperation Forum', https://www.gov.uk/government/collections/the-digital-regulation-cooperation-forum

6. Methodology

To conduct this research, we used a mixture of desk research and interviews. We began with initial desk research on previous ODI work on data assurance that touched on the roles of regulators, industry bodies and professional bodies. We consulted UK government publications and relevant work by other bodies including the Alan Turing Institute and Lloyd's Register Foundation. We also consulted publications by regulators, industry bodies and professional bodies about their own work, and turned to recent academic research on data assurance for conceptual grounding. From the desk research and initial conversations with the ODI we created a list of interviewees.

For our interviewees, we sought representation from regulators, industry bodies and professional bodies. We prioritised interviewing a mix of sector-specific bodies and cross-sector bodies at different stages of engagement with data assurance. We conducted fifteen online interviews with stakeholders and experts for their insights.

Broadly, our interview questions covered:

- How important are data assurance activities to tackling issues in your sector or profession?
- Which data assurance activities are important to the role, goals, and mission of your organisation?
- What data assurance activities is your organisation currently involved in?
- What is working well and where are there challenges?
- Does your organisation conduct any formal data assurance mechanisms?
- What are the barriers for your organisation to do more work on data assurance?

We then synthesised the interviews, looking for common themes among the points raised by interviewees, and conducted further desk research to expand on and corroborate the points raised in the interviews. From this combination of desk research and interviews, we created our recommendations.

Interviewees include:

Regulators and government bodies

- Competition and Markets Authority (CMA)
- Health and Safety Executive (HSE)
- Centre for Data Ethics and Innovation (CDEI)

We engaged with another regulator that prefers not to be listed by name.

Industry bodies

- UK Regulators Network (UKRN)
- Electronic Data Management (EDM) Council
- Financial Data and Technology Association (FDATA)
- UK Finance

Professional bodies

- Royal Institute of British Architects (RIBA)
- Data Management Association (DAMA)
- Energy Institute (EI)

Other organisations

- Global Open Finance Centre of Excellence (GOFCoE)

Data experts

- Professor Sylvie Delacroix, The Alan Turing Institute
- Craig Civil, BSI, Director of Data Science and AI
- Dr. Mahlet Zimeta, ODI
- Ed Evans, ODI
- Lisa Allen, ODI

Glossary

Data access initiative: initiatives or programmes with a strong focus on collecting, using and sharing data as part of their work, involving multiple stakeholders actively working together to solve a problem. They often have a clear challenge, in the form of a specific social, environmental or economic problem that is the focus for the collaboration.

Data assurance: the process, or set of processes, that increase confidence that data will meet a specific need, and that organisations collecting, accessing, using and sharing data are doing so in trustworthy ways

Data assurance activity: a specific activity to create trust in data, such as conducting an audit, validating a dataset, or carrying out training

Data assurance scheme: a specific project that works by applying one or more data assurance activities and may involve multiple actors

Data ecosystem: data infrastructure and the people, communities and organisations that benefit from the value created by it

Data literacy: the ability to think critically about data in different contexts and examine the impact of different approaches when collecting, using and sharing data and information

Data practices: the processes by which data is collected, shared, held and/or managed

Quality infrastructure: formal mechanisms of assurance such as standardisation, conformity assessment, measurement or accreditation

Standards: documented, reusable agreements that solve a specific set of problems or meet clearly defined needs, often used to help define or support a data assurance activity

Trustworthy data ecosystem: an environment in which individuals and organisations across the public, private and third sector trust that data is flowing in ways that will maximise benefits whilst minimising harms