Building on the manifesto for sharing engineering data

Open Data Institute

Contents

About

This report was produced as part of a project funded by Lloyd's Register Foundation. It builds on research from the Open Data Institute's Sharing engineering data for the public good project, also funded by Lloyd's Register Foundation. The views in this report are those of the authors.

This report was researched and produced by the Open Data Institute, and published in March 2021. Its lead authors are Deborah Yates and James Maddison, with contributions from Jeni Tennison and Mahlet Zimeta. If you would like to share feedback or get in touch, please contact James Maddison.

Executive summary

This report is the result of a project by the Open Data Institute (ODI), funded by Lloyd's Register Foundation, which began in December 2019. The project follows our previous collaboration, which explored how sharing engineering data can create public good, by building on two key principles outlined in our manifesto for sharing engineering data:

- **Use challenges to drive innovation that solves problems** by supporting four project teams to explore how better access, use and sharing of data can contribute to public safety.
- **Build data literacy and skills** by assessing the current state of play of data skills in the sector.

This project demonstrates that there is a clear appetite for improving access, use and sharing of engineering data, but that there has been a general lack of coordination across the sector to drive it forwards. The four funded projects demonstrate that the principles outlined in the manifesto can provide a pathway to achieving impact with data, and that working towards one principle can unlock value in other areas of the manifesto.

The report recognises that there is a long way to go before the engineering sector is able to fully realise the value of wider data sharing. In particular, data literacy and skills across the engineering sector need to be improved. To enable this, continuing professional development (CPD) providers need to offer courses that have data literacy and skills built in.

We urge engineering organisations to consider how they can support the manifesto going forwards, whether through advocating for open or shared data, enabling innovation by supporting new projects or building data skills within your organisation. If you would like to discuss how you could act on the recommendations outlined in the manifesto, please get in touch.

Background

In December 2018, the ODI began a collaboration [with Lloyd's Register Foundation](#), an independent global charity that supports research, innovation and education to make the world a safer place. explores how sharing engineering data could benefit the public, with a particular focus on public safety and safety in the workplace. It culminated in the creation of a joint [insight report](#), which includes our [manifesto for sharing engineering data](#). The principles and actions within the manifesto are:

1. Data is infrastructure
2. Data must be stewarded
3. Opening and sharing data unlocks value
4. Explore new data sharing models
5. Use challenges to drive innovation that solves problems
6. Regulation must adapt to new technologies and uses of data
7. Build data literacy and skills
8. Ensure data is used legally and ethically
9. Share knowledge and insight

Applying these principles is a vital step in enabling organisations to create maximum value from data, for the public good. This set of principles and actions has so far been endorsed by 18 leading organisations from the sector

To demonstrate how these principles can contribute to the public good, the ODI and Lloyd's Register Foundation extended their collaboration into 2019. Our aim was to explore how we could enable the sector to deliver on the manifesto by:

- supporting innovative projects with an emphasis on safety
- increasing access to data
- supporting the sharing of data
- producing tools and guidance that enable others to publish or share data themselves.

We have also worked with key stakeholders to examine the extent of data literacy and skills in the engineering sector, with a view towards working with training and accreditation providers to build data literacy and skills into their continuing professional development (CPD) offerings.

Approach

We focused on addressing two core principles of the manifesto: 'Use challenges to drive innovation that solves problems' and 'Building data literacy and skills'.

Use challenges to drive innovation that solves problems

Open innovation can enable new solutions to pressing social, economic or environmental challenges, like improving safety in the workplace or on our road networks.

To drive innovation through a challenge-led approach, the manifesto recommends that:

- Governments and civil society should work together to identify social, economic and environmental challenges that might be addressed through innovative approaches.
- Funders should invest in programmes that enable collaboration across the private sector, startups and researchers to solve specific challenges through the better use of data.
- Funders should help innovators to scale and develop their solutions by supporting the creation of the data infrastructure needed to underpin new products and services.

For this project, we focused on how funding could support the development of new solutions to problems and enable greater collaboration across the engineering sector. We invited external organisations to present their ideas for new projects and ways of working that would advance the Lloyd's Register Foundation strategy and deliver on the vision of our shared manifesto. The invitation to tender was open to any public or private sector organisation, in the UK or internationally, in order to maximise the global impact of the funding and support.

The new stimulus fund offered a total of £150,000 of funding for up to six projects. The successful organisations would help to deliver on the recommendations in the manifesto by:

- exploring innovative ways to improve safety, in line with the mission of the Lloyd's Register Foundation.
- increasing access to data by supporting the publication of open data or sharing of data, across the public and/or private sector.
- encouraging collaboration across organisations, for instance, by building new data infrastructure, or exploring new ways to govern and steward data such as data trusts.
- demonstrating the value of open approaches by encouraging the creation of openly licensing outputs, tools and guidance.

Applicants were assessed based on the following criteria:

- **Alignment with the Foundation strategy.** Proposals should clearly demonstrate how they would support at least one aspect of the Foundation strategy.
- **Alignment with our manifesto.** Proposals are required to clearly demonstrate how they plan to deliver on specific principles within the manifesto.
- **Clarity of the proposal.** Project ideas need to be clearly scoped, have well-defined activities and outcomes and be deliverable within a six-month timeframe.
- **Ability to deliver.** The organisation making the proposal must demonstrate that they have the capacity and resources to deliver on the project within the funded period.
- **Adoption of open approaches.** Projects need to have a commitment to using one or more open approaches, such as publication of shared and open data, use of open standards and open innovation.
- **Value for money.** Projects need to deliver clear value, useful outputs and insights.

With this criteria in mind, five projects were chosen to receive funding and support, to further their ambitious ideas for improving safety. However, due to Covid-19-related constraints,, we were unable to support one of the projects – The Office for National Statistics's Data Science Campus collaboration with the Royal Air Force.

The four successful projects were:

- Unlocking data on brownfield sites, led by Atkins
- Creating a standard to help people find construction knowledge, led by Barbal
- Creating a digital version of a city, led by Slingshot Simulations
- Sharing data to improve the safety of lives at sea, led by the GeoData Institute

Unlocking data on brownfield sites

Area of the manifesto explored: Opening and sharing data unlocks value

There are over 1.1 million households in England on the social housing waiting list, even though there is enough brownfield land – land that is not currently in use, but could potentially be contaminated – to build an estimated 1 million homes.

Benefiting from the stimulus fund, the engineering and design consultancy company Atkins undertook a discovery project to identify how open datasets can be used to develop insights into contamination of brownfield sites. It then compared this approach with similar approaches that could be taken, using existing commercial and unused datasets.

Key findings from the project:

- Critical sources of contamination are currently only found in non-open datasets.
- Existing datasets could be made more open to help improve decision making and reduce costs associated with developing brownfield sites, particularly for smaller developers, and help address the national housing crisis.
- There is unused data held by companies like Atkins and its customers which, if shared in ways that built trust, could unlock further value.
- There is an opportunity for industry to collaborate and realise the power of location data about brownfield land contaminants. These conversations will be important in supporting Government housing commitments and delivering the ambitions in the UK Geospatial Data Strategy.

For more information about the project, please read our case study, the Atkins project report or watch the lunchtime lecture presentation.

Creating a standard to help people find construction knowledge

Area of the manifesto explored: Data is infrastructure

A third of construction practitioners (approximately 1 million people) do not have easy access to the knowledge they need to do their job. The number one barrier to access is that practitioners are not aware of what knowledge is available. There is therefore a need for quick and straightforward access to knowledge on demand.

Barbal, a Bristol based enterprise specialising in creating open standards for data, worked with the Construction Knowledge Task Group to collaboratively publish construction knowledge so that it becomes discoverable. In time, this will greatly improve practitioners' access to the knowledge they need to safely perform their role and implement a safer built environment during construction and in-use.

Key findings from the project:

- Taking an open and collaborative approach accelerates the development of technical solutions in the construction industry.
- Staying up to date with the latest regulations and best practice ensures industry professionals are competent to deliver safer buildings.
- Barbal's approach to improving discovery of construction knowledge could be the basis of a global coalition for change for construction knowledge professionals.

For more information about the project, please read our case study or watch the lunchtime lecture presentation.

Creating a digital version of a city

Area of the manifesto explored: Opening and sharing data unlocks value

Air pollution is one of the leading contributors to increased mortality around the world, with established links to numerous heart and respiratory diseases. Roughly 7 million deaths a year are caused by air pollution and more than 80% of people who are affected by poor air quality live in urban areas, such as cities. Decision makers in cities must do more to mitigate poor air quality and protect citizens by making better decisions around city planning.

The 'Breathing City' project – an initiative led by Slingshot Simulations – demonstrates how open data can be used to create a digital representation of a city, which can provide valuable insights about the impact of pollution, and help decision makers to improve the wellbeing and safety of city populations.

Key findings from the project:

- Slingshot Simulations has created a digital twin of Leeds city centre, using multiple open data sources, which can be used to understand the challenges faced around air pollution in the city.
- The digital twin enables city decision makers to simulate different scenarios in which changes to city infrastructure could have a significant impact on the effects of air pollution.
- The digital twin has the potential to be rolled out across different cities, with a reusable simulation being produced as part of the project. Newcastle-under-Lyme council has approached the team to explore replicating this in Staffordshire.
- Slingshot's work has created a foundation for helping city planners in a world impacted by the coronavirus, which will be further explored through its follow-on project A Living Model for People and Place, which is funded by Innovate UK and supported by the University of Leeds VirtuoCity Centre for City Simulation, Leeds City Council and Newcastle-under-Lyme Borough Council.

For more information about the project, please read our case study or watch the lunchtime lecture presentation.

Sharing data to improve the safety of lives at sea

Area of the manifesto explored: Explore new data sharing models

As an island nation, the UK sees numerous coastal incidents and emergencies every year. In 2019, Royal National Lifeboat Institute (RNLI) lifeguards responded to 17,356 incidents near the shore and lifeboat crews were launched 8,941 times in response to incidents at sea.

The Open Marine Data project, led by the GeoData Institute based at the University of Southampton, supports emergency responders to improve access to the data that they collect, with a view to providing an open marine dataset for use by the whole sector. This will enable innovators to create new tools and approaches that support emergency responders to make better decisions.

Key findings from the project:

- There is an opportunity to improve the way that data is shared between key emergency response stakeholders, such as the Maritime Coastguard Agency (MCA) and the Royal National Lifeboat Institution (RNLI). Data is currently shared manually. To improve data sharing, MCA is looking to adopt the Multi Agency Incident Transfer (MAIT) schema, a standardised UK government format for sharing incident data, which enables device-to-device data transfer without the need for human involvement.
- Currently the MCA and RNLI collect different records of a particular incident, which means that often, neither organisation has a complete picture. Because this data is not collected in a standardised format, it can be difficult for the organisations to use the data that is being shared.
- The GeoData team has created a blueprint for these organisations to adopt a standardised approach to collecting data, which will help to resolve most of the existing issues that prevent data from easily being shared, such as ambiguity around the location of an incident and a lack of common understanding around attributes.
- The blueprint lays the foundation for development of an open marine-incident data standard, which can be adopted by other emergency responders working in the marine sector. The standard will enable responders to publish open data about incidents for others to use and learn from.

For more information about the project, please read our case study.

Building data literacy and skills

Organisations need the capacity to make use of data and an understanding of how to share it safely. To achieve this, data literacy and skills need to increase across the engineering professions and organisations, and a range of stakeholders need to collaborate.

To help build data literacy and skills in the engineering sector, the manifesto recommends:

- Professional bodies and societies should ensure that their professional development and certification courses are building data skills
- Universities and research organisations must ensure that they are teaching the necessary data skills required for the future of the profession
- Private sector companies should design their internal skills and learning programmes to build an understanding of the value of data at all levels of their organisation.

Our work in this area focused on validating these three points, by providing evidence for the need and demand for data skills to the relevant stakeholders listed above.

We first undertook a review of the available training, support and guidance for engineers regarding access, use and sharing of data. We conducted this research using a combined approach of desk research and expert interviews. Insights from both elements informed the structure and focus of a workshop, which convened a multi-stakeholder group from across the sector. Based on the findings from the workshop, a roundtable was convened with engineering CPD providers to explore opportunities for adding data skills to existing CPD offerings.

Desk research and interviews

Our research on data literacy and skills in the engineering sector focused on two key questions:

- What codes of practice or guidance are available to organisations in the engineering space around accessing, using and sharing data?
- What training courses or certification are available to organisations in the engineering space around accessing, using and sharing data?

The desk research uncovered a number of gaps in guidance, training and skills, which were used as the basis of questions for seven expert interviews, with a focus on current best practice. We engaged with a range of stakeholders working in the sector, including a regulator, funder, research organisation and construction firm, as well as with organisations that play an advisory role to the sector.

Key insights from the research:

- There is little generic guidance or training around data in the engineering sector. Most organisations follow the rules outlined in the General Data Protection Regulation (GDPR), and may have internal guidelines around accessing, using and sharing data.
- Some organisations have their own training programmes around data, but they are not always mandatory. There are a small number of organisations that offer external guidance, such as Royal Academy of Engineering's 'Towards Trusted Data Sharing' work.
- Interviewees agreed that data guidance needs to be embedded into everyday engineering practices, and that people working in the sector need to be able to understand the value of data for their role and organisation. If organisations are to invest in upskilling staff in data knowledge and skills, it needs to be aimed at addressing specific business or organisational problems or opportunities.
- Data is rarely well understood across the organisation, which means that culture change around data and digital is driven by a few individuals, and seldom by those at the top of the organisation.
- There is sometimes a disconnect in understanding between departments or roles, which can cause miscommunication and a misinterpretation of each other's needs.

Workshop

To build on and validate the findings from our early research, we facilitated a workshop to explore the extent of support, training and guidance for data literacy and skills in the engineering sector.

The workshop included nine participants from nine organisations:

Organisation	Organisation role	Attendee Role
Health & Safety Executive	Regulator	Senior Scientist
Colouring London/Alan Turing Institute	Researcher	Programme Director
EPSRC	Funder	Senior Portfolio Manager
GeoData Institute	Creator	Enterprise Fellow
Atkins	Steward/Creator	Data Scientist
Mission Drive	Data skills consultant	Data Skills Consultant
Lloyd's Register Foundation	Funder	Programme Manager
RAEng – Engineering X	Researcher	Senior Manager
Tideway	Steward	Engineering Information Manager

Together with these stakeholders, we discussed the role that data plays in the everyday work of engineers; examined the balance of data skills needed across engineering organisations; and discussed how existing support compares to the needs of the sector, in order to inform interventions that would drive greater use and sharing of data.

While it was felt that in some areas data skills and knowledge were fairly strong, the general consensus from participants was that the level of data skills and knowledge across the sector needs to be improved and requires more investment. The group also agreed that the level of support available was not currently sufficient.

For more information about the findings of the workshop, please read our workshop report.

Based on the findings from this workshop, we made recommendations on next steps to be delivered over the short term (within six months) and long term (beyond six months). These recommendations include:

1. Producing a white paper on data skills in engineering.
2. Convening a roundtable meeting with engineering CPD providers to explore opportunities for adding data skills to existing CPD offerings.
3. Working with universities, research organisations and professional bodies to increase capacity in data skills and knowledge.
4. Developing an 'introduction to data in engineering' learning programme.

We decided to focus on the roundtable with CPD providers for the remainder of the project. By convening organisations that offer training and support to the sector, we sought to understand how they incorporate data literacy and skills into their offerings.

Roundtable meeting

At the beginning of February 2021, we brought together representatives from three engineering institutes that provide CPD opportunities for the engineering sector, to explore opportunities for improving the provision of data skills and literacy support.

We framed discussions around three key questions:

What data skills and knowledge are needed to improve access, use and sharing of data in the engineering sector?
Participants felt that data skills were becoming increasingly important for people working in the engineering sector. They recognised the importance of technical data skills, such as using platforms or analysing data, but they were less familiar with strategic data skills, such as managing organisational change or delivering policies with data. The group acknowledged the importance of building strategic data skills in parallel with technical skills to ensure that the curation and use of data is valued at all levels of an organisation.

How well are data literacy needs already supported?
Most CPD providers offer training and accreditation that applies specifically to a job role or domain, but it was clear that these courses rarely help people to consider how skills to access, use and share data applies to their context. Many CPD providers are already in a good position to provide relevant sector or role-specific guidance, but would need to invest time and resources in their own training staff, so that they could incorporate data skills and literacy into their courses.

Participants suggested that they would need to see more demand from engineering organisations for data literacy and skills support in order to invest in changes to their offerings.

What steps do we need to take to improve the support available around data skills and knowledge in the sector?

The first step for many CPD providers will be to examine existing data skills and data literacy gaps within their own offerings. Participants suggested that while they were aware that their courses and training did not cover data skills in detail, they were not sure where to begin. Leaders within CPD organisations can support this change by investing in the data literacy of learning and development specialists, who develop the courses and deliver the training to customers in the engineering sector.

By building data skills and literacy internally, CPD providers will be able to embed modules covering access, use and sharing of data within existing training and courses. This will help the engineering sector understand how better data practices can support their work.

For more information about the findings from this workshop, please read our blog post on improving data skills in the engineering sector.

Conclusion

The most important insight from our work exploring data sharing in the engineering sector is that there is a clear appetite from organisations to engage more with data. There seems to have been a lack of coordination to realise the ambition in the past, but during this project, we have seen an enthusiasm for data manifest in the following ways:

- The initial wave of endorsement for the manifesto, from organisations such as Health and Safety Executive, Royal Academy of Engineering and Arup, gave us an indication that organisations in the engineering sector are keen to explore the potential of using and sharing data.
Our engagement as part of this follow-on project, with a wide variety of stakeholders, through interviews, a workshop and roundtable meeting, has solidified that organisations are interested in being able to improve how they use data to solve problems.
- External engagement with project outputs demonstrates that people want to learn how data can enable innovation, address challenges and save lives. Our case studies, blogposts and lunchtime lecture videos associated with the stimulus fund projects have received a combined figure of over 2,000 views.
- The expression of interest in joining our stimulus fund – with applications coming from 10 organisations – also suggests that engineering organisations are ready to explore new approaches to data. This suggests that many of the barriers to sharing data that have been cited to us in the past (see our joint insight report), such as commercial sensitivity and concerns around public perception, become less of an issue when an organisation is given the time and resources to explore the feasibility of a new idea.

Not only has the ODI's work with Lloyd's Register Foundation demonstrated that there is an appetite for better access, use and sharing of data in engineering, but the outcomes of each stimulus fund project have shown that the principles outlined in the manifesto can provide a pathway to achieving impact with data, and that working towards one principle can unlock value in other areas of the manifesto. For instance:

- The project led by Barbal to create a shared way of publishing construction knowledge is a great example of how treating data as infrastructure, and creating rules around how we access and use it, can lead to better access to knowledge and insight.
- The projects led by Slingshot Simulations and Atkins both demonstrate how open and shared data from third party organisations can be used to address societal challenges like air pollution and a lack of affordable housing, in innovative new ways.
- The GeoData Institute has been able to show that using new models for sharing can help organisations to access and use data more effectively, saving time and resources for all involved.

We have also learned that there is still a long way to go before the engineering sector is able to fully realise the value of wider data sharing. Our engagement activities with people in engineering roles suggests that the level of data literacy and skills across the sector needs improvement.

There is an opportunity for organisations that provide CPD to equip people in engineering roles to understand the relationship between data and their work. We encourage universities, research organisations and professional bodies that deliver CPD support for the sector to consider how they can build data literacy by building data skills into their existing offerings. There are tools available to help with this assessment, such as the ODI's Data Skills Framework, which can help organisations to inform their data literacy programmes, identify skill gaps and clarify the best places to invest in learning and capacity development.

Demonstrating demand for data literacy and skills will be an important motivator for CPD providers to invest more in data-enabled courses and training. Engineering organisations need to show CPD providers that there is a demand for learning relevant data skills. Without more evidence of demand for data skills, training providers might not be able to justify investing in adapting their offerings, so engineering organisations must play a stronger role in advocating for more data skills support.

Endorse our manifesto

Our work this year has demonstrated that following the principles outlined in the manifesto for sharing engineering data can help address issues of safety, and that improving how we access, use and share engineering data could be instrumental in helping to solve other social and economic challenges in the sector.

The good news is that your organisation could be contributing to the public good being created from data too.

Start by asking yourself one of the following questions:

- Do you feel equipped to work with data in your role?
- Could you be solving problems with the data that your organisation collects?
- Have you considered how sharing data can create value for your organisation?
- Do you offer courses and training for engineers that build data skills and literacy?
- Do you have a story to share about how engineering data can be used for the public good?

We urge engineering organisations to consider how they can support the manifesto, whether it be through advocacy for opening or sharing data, enabling innovation through supporting projects with new ideas or building data skills within your organisation. If you would like to discuss how you could endorse and act on the recommendations outlined in the manifesto, please get in touch.