Tackling the climate crisis with data: what the built-environment sector can do

November 2021

Open Data Institute

Contents

About

Founded in 2012, the Open Data Institute (ODI) is an independent and not-for-profit organisation based in London, UK. The ODI works with companies and governments to build an open, trustworthy data ecosystem, where people can make better decisions using data and manage any harmful impacts.

For over 75 years, Arup has been recognised for its vision, talent and tenacity. Today, the firm is a collective of over 16,000 designers, advisors and experts working across 140 countries. Founded to be a humane organisation, one dedicated to excellent work across many disciplines, it collaborates with clients and partners to shape a better world.

1. Foreword

Foreword by Volker Buscher, Chief Data Officer, Arup Group

The Paris Agreement[1] is based on global climate models that work. Governments and businesses have agreed milestones for 2030 and 2050 that require us to act now. These climate goals require data that is specific to key industries, users and asset investors. For example, what is the carbon risk for an asset manager? How do property owners design and manage these assets? What is the role of the consumer or employee in the drive for net-zero properties? These questions apply across the built environment, including buildings, transportation, utilities, cities and agriculture.

Solving the climate emergency requires creativity in how we collect, use and share 'data at scale'. Data at scale has been transforming retail, banking and media for the past two decades. The built environment is now going through a similar transformation by combining high volume, variety and velocity of data with machine integrated data processing in architecture, engineering and asset management. A process that will unlock economic, environment and social growth.

For Arup this means investing in our data infrastructure and science, innovation, knowledge, culture and practices. But we are not an island and data sharing across the industry is vital. For example one of our early prototypes for data-enabled innovation addressed agent-based transport models. We spend 60% or our time on data ingestion and cleansing, due to inefficient bilateral data sharing, instead of focusing our entire efforts on climate-positive transport planning.

And it is not just the cost and friction regarding the use of data at scale in design and management of the built or natural world. It is also the economic opportunity for growth and innovation. Open banking[2] has grown into a $43bn[3] data-at-scale industry. How big is the opportunity for tech investors in the built and natural world? Can the emerging PropTech (property technology), ConstructionTech (construction technology), CleanTech (technology that aims to improve environmental sustainability) etc match what open banking has delivered?

This has led to our call for collaboration with the ODI as part of our research programme Data Supernova[4]. Our first research initiative focused on 'Exploring new approaches for sharing data in the built environment'[5]. This time our aim is to understand better the role data at scale can play in dealing with embodied carbon, operational carbon, human behaviour, resilience and transparency.

We would like to understand better why data matters in climate change. We would like to apply the ODI's theory of change[6] as a catalyst for the built world to develop the data foundations required to unlock the economic and environmental value generated by data at scale.

---

[1] United Nations Climate Change (2015), The Paris Agreement

[2] Dgen (2018), Development of UK open banking

[3] Allied Market Research (2020), Open Banking Market Size, Share | Opportunity, Trend and Growth 2026

[4] The Alan Turing Institute (2020), Volker Buscher and the Data Supernova

[5] ODI (2021) Exploring new approaches for sharing data in the built environment

[6] ODI (2018), Theory of change

2. Executive summary

In 2021, the Intergovernmental Panel on Climate Change[7] confirmed that human activities are having a profound impact on the planet, and will inform the global negotiations at Cop26[8], to accelerate the journey to net zero and keep the increase in temperature above pre-industrial levels to less than 1.5°C. ‘Pre-industrial levels’ refers to the approximation of the surface air temperature of this era from the IPCC 'Global warming of 1.5°C' report[9]. As with other significant challenges we face, effective use of data is becoming central to delivering emission cuts, and doing so in ways that deliver wider environmental, social and economic value. Data is increasingly seen as a powerful tool in solving societal challenges, which becomes even more powerful when standardised and shared in ways that meet a wide range of user needs.

At the moment there are growing concerns in the built environment in relation to the lack of accessible, reliable and trusted data and the missed opportunities for using it to develop innovative products and services. As such, we recognise the potential to improve the underlying data infrastructure across the sector, drawing on learning from other domains, and exploring some of the approaches needed to rapidly move towards realising these.

With this in mind, we examined several particular challenges, including bringing together data on the existing built environment to support strategic action and investment, and unlocking greater building and house-level data to support individual choices around energy saving and retrofit strategies. We identified the need to improve standards and working practices and the challenges relating to reliable and trustworthy carbon calculation, while pointing to the potential for modern data stewardship practices to address some of these.

And finally, we looked at the importance of bringing in data from other settings and sectors, working across professional and organisational boundaries. Some of the challenges are technical, some are political, and others are around culture and capacity, and they cannot be solved by organisations acting alone.

The need for urgent action on climate change is no longer a matter of debate. The question now for the built-environment sector is how to accelerate action to reduce emissions, build for resilience and protect biodiversity. We need to use all the tools at our disposal, including the significant power of data. Meeting this challenge demands coordination, cooperation, community building and culture

---

[7] Intergovernmental Panel on Climate Change (2021), [Sixth Assessment Report: AR6 Climate Change 2021:The Physical Science Basis]

[8] United Nations (2021), [COP26: It's (almost) here]

[9] Intergovernmental Panel on Climate Change (2019). [Global Warming of 1.5 °C]

change across the sector, and for cross-sector collaboration to break down data silos.

This draft paper also launches a call for collaboration – we are asking organisations and individuals to take action:

- **Give us your feedback on this working paper** to help focus future data collaborations and activities. Fill out this form by 19 November 2021 to send your feedback.

3. Introduction

The data is clear. We face a climate and biodiversity emergency.[10] Data that has been gathered, curated and analysed by thousands of scientists across the world has established the greenhouse gas cuts that we need over the next decade to keep global temperature rises below 1.5 degrees. Effective use of data is now central to delivering those cuts, and doing so in ways that deliver wider environmental, social and economic value.

The built environment contributes around 40% of the UK's total carbon footprint[11] (and a similar figure worldwide[12]), through a mix of embodied carbon in new construction, and operational carbon used to heat and power existing buildings. Add in the emissions from road and rail transportation, and the majority of emissions are shaped by choices made in the design and operation of our building and infrastructure. The UK Institute of Civil Engineers estimates that, whilst some progress has been made, we *'need to reduce the UK's annual carbon emissions from infrastructure more than 30% faster than we are doing now"* and although some of the needed reductions rely on substantial technological and policy shifts, decarbonisation can also be accelerated by *"such mundane measures as better data management*.'[13]

Data is a powerful tool. And it is made even more powerful when it is standardised and shared in ways that meet a wide range of user needs. The ODI's experience over the last nine years has shown that, when intentional design, investment and energy is put into creating effective data ecosystems, supported by the right mix of openness, incentives and data institutions, data can deliver a step-change in how problems are solved.

[10] Built Environment Declares (2021), Built Environment Declares Climate and Biodiversity Emergency

[11] UKGBC (2019), Climate change

[12] World GBC (2021), The building and construction sector can reach net zero carbon emissions by 2050

[13] Institute of Civil Engineers (2021) State of the Nation 2021 | Six ways for civil engineers to act on climate change

Across the built-environment sector there are common concerns about a lack of accessible, reliable and trusted data. The data ecosystem for decarbonising the built environment has developed over many years, reliant on scattered spreadsheets, ad-hoc research projects, proprietary databases and conflicting calculations, leading to a landscape in which targets are hard to both reliably set, and measure. For emerging big data, the Internet of Things, sensor network and real-time data sources, questions over standardisation, privacy protections and governance can frustrate data sharing and re-use.

Meeting climate goals requires transparency and accountability mechanisms. How do governments, businesses and citizens know if society is meeting or exceeding climate change goals? The level of change and complexity across the built and natural world will require transparent reporting based on data where the provenance and accuracy can be assured. This will help meet goals, protect asset value and enable companies to compete based on the quality of their work.

Addressing the data foundations of the built environment has significant potential. From helping drive intelligent retrofit strategies that reduce the operational carbon footprint of our buildings, to ensuring new infrastructure projects minimise their embodied carbon, and helping local communities to invest in measures that reduce their exposure to climate risk – better flows of data can support innovation, collaboration and transparent measurement of progress.

In the sections that follow we explore a range of opportunities for data to be used in the built environment to tackle the climate emergency, and, drawing on learning from other sectors, explore some of the approaches needed to rapidly move towards realising these.

4. Opportunities for better data sharing across the built environment

When it comes to reducing the carbon emissions from the built environment, there are different elements at play. First, the vast majority of operational emissions from the built environment are as a result of human behaviour. Energy use for heating, cooling and powering our homes and offices, use of transportation, and the intensity with which buildings are used all affect the carbon emissions linked to each square metre of built space. So better data and the products and services that can be built upon it can help planners, companies, consumers and citizens to make better decisions. Yet, those decisions, and their impacts, are also substantially shaped by the design and quality of new and existing building stock.

Looking at existing building stock, to target investment in retrofit strategies and low-carbon upgrades, we need a better understanding of that stock and infrastructure, including its condition, ownership and options for improvement. In the UK, where the vast majority of the housing that will exist in 2030 is already built, and very little is currently optimised for low-carbon living, data has a central role to play in shaping national, local and householder strategies to cut climate impacts.

Transforming our built environment will certainly need a skilled workforce and well functioning markets to make sure labour is available to deliver improvements. Better data about the 'green' workforce, and a workforce better equipped to use data, will both be vital to meeting the challenge of climate change. In both new construction, and in relation to retrofitting strategies, it is vital that gains in operational efficiency are not undermined by embodied carbon costs of construction materials, technologies and processes. Approaches to lifecycle assessment (a method for assessing environmental impacts associated with all the stages of the lifecycle of a commercial product, process, or service) and accounting for the full carbon costs of each choice made in built-environment schemes demand detailed and trustworthy data.

Lastly, investment in a low-carbon transition will require both new finance and high-quality company and government reporting, increasing transparency and accountability mechanisms to make sure that global efforts are not derailed. Trustworthy data will also be vital to support market mechanisms for decarbonisation, including carbon taxation. The right data can help investors better understand, prioritise and respond to climate change risks, and avoid holding stranded assets.

Right now, data supports each of these elements of low carbon transformation, but we are a long way from harnessing its full potential. No single organisation holds or controls all the relevant data, or alone has the resources and skills needed to realise the full benefit data can bring. In the sections below we explore four themes to understand the potential of data, highlighting current data sharing initiatives and exploring the challenges that organisations are facing to fully harness data. We identify some initial opportunities for better sharing and use of data to address the climate emergency through the built environment.

Operational carbon

With good design, data can help shift the day-to-day decisions made by consumers, commercial firms, policy makers and developers towards lower-carbon outcomes. From in-home smart meter displays and energy performance certificates (EPCs), to platforms providing at-scale analysis of the carbon footprint of large property portfolios,[14] there are already many established data-driven tools that encourage action to reduce energy consumption and operational emissions. However, these tools often rely on relatively limited data about the existing built environment, and their ability to provide tailored recommendations is held back by missing data, or barriers to data sharing.

For example, whilst over 20m UK properties now have smart meters and many are able to see their real-time energy consumption through a device in the home, there are obstacles to be overcome in providing safe and trusted access to that data for third-party analysis that might be able to suggest individual strategies for lowering energy use, or that might support more systemic changes. The ODI has been working with the Data Communications Company (DCC), who provide the licensed monopoly platform for smart-meter communications, to explore ways to increase access to this data, while respecting the significant legitimate privacy concerns around its use[15].

[14] For example, Parity Projects

[15] Data Communications Company (2021), Data for Good

When it comes to understanding the properties of existing building stock, we face similar challenges of bringing together data from different stakeholders, and navigating issues of privacy, commercial interest and the appropriate level of detail to share. Energy-performance-certificate data, for example, provides basic information about millions of residential and non-residential properties in the UK, and it is available through a government provided application programming interface (API).[16]

However, its use raises some practical challenges. Firstly, the data licence explains that *'address level data concerning the energy performance of buildings constitute personal data for the purposes of the General Data Protection Regulation (GDPR)'*[17], placing certain restrictions on the kinds of applications to which the data may be put.

Secondly, the data only covers properties that have been sold since EPC regulation was introduced in 2007, and as certificates are valid for up to 10 years, data can be significantly out of date. Thirdly, EPCs are based on only a basic assessment, and are generated using assumptions that may not reflect the full condition of a building. However, many other stakeholders may also hold information about buildings: from insurers and mortgage companies, to local authorities and construction regulation schemes, maintaining registers of boiler or window installations. Add to this remote-sensed data and satellite imagery that might provide additional context about buildings, and the natural environment surrounding them, and the opportunity for an understanding of buildings based on multiple data signals, rather than one-off surveys and certificates, becomes clear. The Royal Institution of Chartered Surveyors has highlighted the potential of a growing PropTech (property technology) sector to support revitalised EPCs[18].

With the right institutional structures in place, it may be possible to develop data access initiatives that can bring together different building-level datasets, and unlock new finance for reducing operational carbon. Crowdsourcing can also fill key data gaps, and engage citizens in thinking about individual and collective action. For example, as in the case of Colouring London, a citizen science platform hosted by the Alan Turing Institute designed to gather open data on London's building stock. By colouring in the building map, users are able to provide structured data in categories including building age, construction methods, and sustainability performance, and participate in making sense of the opportunities and challenges for transforming the built environment[19].

---

[16] Open Data Communities Energy Performance of Buildings Data England and Wales

[17] Energy Performance of Buildings Documentation

[18] RICS (2020), Retrofitting to Decarbonise the UK Existing Housing Stock: RICS net-zero position paper

[19] Colouring London

Case examples: using building data to plan carbon reductions

In 2017, Arup worked with the City of Boston, Massachusetts, US, to develop a classification of 86,000 buildings into 75 distinct ‘building energy models’. Calibrating these models with anonymised data from Boston’s utility providers, the project was able to explore which combination of policies was most likely to help the city achieve a target of net-zero carbon emissions by 2050[20]. The work fed into wider recommendations that called for a mix of new regulatory requirements, upfront funding, and workforce training.

In the UK, housing associations and local authorities are responsible for millions of properties. To work out how to sequence improvement plans, social landlords could potentially draw on data on building age, construction and condition, current energy demand, future overheating risks, and scope for improvement, as well as data on the local labour markets and supply chains that could carry out property updates. Just as for individual homeowners, data to understand the financial cost of improvements, and the likely costs of doing nothing, is particularly important, as it shapes forward planning and cost-benefit calculations. Right now this data is often sparse, hard to access, and lacks standardisation – leading to a reliance on data intermediaries to make best-effort calculations.

Much of the data that is needed to plan for more efficient use of building space comes from outside the built-environment sector: calling for cross-sector collaboration. For example, to meet operational carbon reduction goals, companies need to make more intensive use of existing space, reducing overall space heating and power requirements, *and* reducing work commuting and travel. This requires action to understand, and shift, current working and travel behaviours. The Covid-19 pandemic has seen datasets from transport providers, mobile networks and payment processors used to provide near real-time insight into how lockdowns transformed the use of urban space, and how it might shape future infrastructure demands[21]. Qualitative surveys and dialogue with employees, partners and customers can help firms interpret what aggregate trends mean for each individual business, and can support planning that makes more efficient use of existing resources, reducing demands for new construction.

[20] Arup (no date), [Carbon Free Boston]

[21] National Infrastructure Commission (2021), [Behaviour change and infrastructure beyond Covid-19]

Ultimately, the energy demand of homes and offices can only be brought down far enough through appropriate upgrades, widespread retrofitting of insulation and low-energy technologies. Data is at the heart of retrofit planning, but right now, consumers face a confusing picture when it comes to weighing up the costs and benefits of different interventions. Without accurate data on upfront costs and potential benefits, consumer trust can be undermined, and householders may be less willing to take on retrofit projects. Carbon Coop in Manchester has been exploring the use of data to drive community-powered retrofit plans. Its findings highlight the importance of thinking at the community level, as well as at the household level: combining statistical data with local knowledge to identify neighbourhoods with the greatest potential for multiple retrofit projects, and then working at the grassroots to develop cost-effective retrofit plans[22]. Carbon Coop has also worked with Open Data Manchester and the ODI to explore opportunities to create a data cooperative – where people or organisations agree to pool and share data – to support this work[23].

For central governments and local authorities, there are also substantial returns to be had through taking a data-driven approach to policy making. From national incentive schemes, to local authority regulations, better data has the potential to help set shared targets, track progress towards goals and inform interventions. One area where this is particularly important is around workforce development. Carbon Coop found that *'whilst there isn't currently a ready-made 'Retrofit Contractor' workforce, there are quality RMI (Refurbishment, Maintenance and Improvement) contractors who can apply themselves to a given specification with the right support"'*[24], yet there is little structured data available about this primarily word-of-mouth workforce. A recent Nesta project using its Open Jobs Observatory to apply machine-learning techniques to explore green jobs through data shared on key job-seeker websites shows promise in finding ways to gain more labour market intelligence at scale to inform government policy[25], while Carbon Coop is building local software and data infrastructure to better connect generalist contractors with retrofit opportunities.

---

[22] Carbon Coop (2019), [People Powered Retrofit: A community led model for owner occupier retrofit Project Report]

[23] Open Data Manchester (2021), [Developing a Data Cooperative Model for Small Energy Cooperatives]

[24] Carbon Coop (2019), [People Powered Retrofit: A community led model for owner occupier retrofit - Project Report]

[25] Nesta (2021), [Finding jobs in green industries results]

Embodied carbon

With the right standardisation and flows of data, emissions forecasts for construction projects could be as commonplace as financial forecasts. Several platforms and tools already exist to take annotated digital building information models (BIMs) created during the design and construction process, and to use the data from these to drive live cycle analysis: giving estimates of the embodied carbon for each design or material choice. Forecasts of an asset's operational carbon emissions, and the risks of issues such as overheating under future climate-change models, can also be generated by various platforms. Digital-twin projects go further, creating a dynamically linked digital representation of built assets that are intended to exist for the whole lifecycle of the structure. In some cases, these are updated with real-time sensor data, and are able to support predictive modelling that can be used to optimise buildings for lower carbon. Wider adoption of such tools has significant potential. One estimate suggests that if structural engineers alone were to use such tools to reduce the embodied carbon of buildings by just 10%, there could be around 3m tonnes of carbon savings each year.

However, a lack of common practice and standardisation – as digital models move between commissioners, architects, engineers, contractors, installers and builders – creates a significant barrier to making clear carbon calculations the default. There is often an information gap between digital twins and the physical performance of a building. Currently, changes to a design (and the subsequent impact on operational performance) are not communicated back to designers, and data is lost.

Current tools to support carbon estimation are often the result of industry-academia collaboration, but often appear to lack pathways to adoption beyond the few firms that created them. This creates the risk of many competing 'version 1.0 tools', rather than collaborative investment in common and open version 2.0 and 3.0 platforms that are both sustainable, and able to deliver a more joined up user experience across the process from commissioning to construction.

Case examples: Carbon calculation

In the last few years, a number of new tools have been launched to streamline carbon calculation. For example: the Institution of Structural Engineers 'Structural Carbon Tool' is an open-source Microsoft Excel-based estimator[26] that gives broad estimates of the carbon impact of different structural designs. OneClickLCA provides a commercial platform to perform building and infrastructure LCA, and assess projects against more than 40 different certification schemes[27]. Mott Macdonald's Moata Carbon Portal[28] offers integration with BIM systems to provide its clients with real-time calculation of the carbon cost of building models. Arup's PECC tool[29] provides similar functionality, layering carbon calculations onto virtual models

The Icebreaker One project is exploring ways to streamline carbon calculation by rebooting an open dataset of carbon calculation models[30]. The dataset, previously maintained for several years by startup AMEE, seeks to provide a unified interface and API onto the different methods available to calculate embodied, operational or life cycle carbon of particular materials or construction methods. Just as a financial forecast may provide a range of possible outcomes based on different assumptions and scenarios, being able to easily access different methods can help projects to more clearly communicate to decision makers the range of possible carbon impacts of each option.

Data & Analytics Facility for National Infrastructure (DAFNI)[31] is an £8m investment from the UK Collaboratorium for Research on Infrastructure and Cities (UKCRIC) intended to provide 'world leading infrastructure systems research capabilities'. Through provision of data storage, compute, visualisation and analysis capacity it aims to remove technical barriers that prevent researchers exchanging and working with complex digital models of infrastructure and buildings. Pilot work to date has explored how DAFNI can act as a digital twin platform[32], and how it can provide a technology platform for the OpenClim Climate Impacts Modelling Framework[33].

[26] The Institution of Structural Engineers (2021), The Structural Carbon Tool

[27] One Click LCA: World's fastest Building Life Cycle Assessment software

[28] Mott Macdonald (2021), Moata Carbon Portal

[29] RIBA Journal (2020), Carbon calculation tools to cut embodied CO2

[30] Icebreaker One (nd), Discover

[31] DAFNI: Data & Analytics Facility for National Infrastructure

[32] Data & Analytics Facility for National Infrastructure DAFNI Champions: DAFNI as a Digital Twin Platform

[33] Data & Analytics Facility for National Infrastructure (2020), OpenCLIM

Data can guide decisions over the projects that should receive funding, subsidy or prioritisation as we work towards ambitious carbon reduction targets. The UK's Infrastructure and Projects Authority (IPA) forecast that more than £650bn will be spent over the next decade on new infrastructure, from roads and rail schemes, to new schools, flood management schemes and broadband networks[34]. However, with the construction phase accounting for up to 40% of the lifetime emissions of an infrastructure project, reducing the carbon cost of this investment over the next 10 years has a key role to play in meeting 2030 targets.

Both local and national public works pipelines are frequently based on static cost-benefit analysis calculations that have not been updated to reflect both changing patterns of demand, and the increasing urgency of carbon reduction. A move from document-driven processes, to planning and decision-making based on data-driven and standardised digital models, accompanied by appropriate training for decision makers, has the potential to transform the choices made over what to build, how, and when.

Earlier this year, the UK Net Zero Infrastructure Industry Coalition published a report drawing on data from the IPA's National Infrastructure and Construction Pipeline dataset to try and calculate the locked-in embodied carbon that will result from proposed infrastructure projects[35]. While this work offers a powerful demonstration of how data could be used to set the direction of public investment, the analysis was frustrated by *'significant issues with data availability, quality, and transparency across sectors that needs remedying from both top-down and bottom-up'*. In particular, the project team found it hard to locate data stewards in construction organisations who could supply access to the key datasets that might offer insight into the likely carbon footprint of projects.

Effective data stewardship is also a key issue for the reference datasets used in greenhouse gas accounting and lifecycle analysis (LCA). At present, the landscape of reference datasets and models for carbon calculation can be complex and confusing: leading to variation between different embodied carbon estimates. There is also an interplay between regulation or legislation, and carbon estimation datasets, with different national or sectoral requirements leading to use of different datasets. Greater policy harmonisation could help simplify the data landscape.

Maintaining these reference datasets takes considerable effort: requiring ongoing research into the carbon emissions of materials, keeping track of changing global supply chains, and converting research outputs into reusable calculations. This results in a mixed marketplace of government or academic provided datasets, and commercial offerings, but few that are maintained as robust or collaborative[36] open data infrastructures with built-in version control, APIs and supporting

[34] UK government (2021), Analysis of the National Infrastructure and Construction Pipeline 2021

[35] Net Zero Infrastructure Industry Coalition (2021), Is our carbon wallet empty? The embodied carbon of the National Infrastructure Pipeline

[36] ODI (2020), Launching the collaborative data maintenance guidebook

software toolkits. For example, the widely used Inventory of Carbon and Energy (ICE) database[37] is provided as an Microsoft Excel download, under a permissive, but non-open licence. Modernising the management of greenhouse gas accounting and LCA datasets through creation of the right data institutions to steward data holds out the possibility of substantially increasing the usability and use of this vital data. This could foster a shift towards more open business models for carbon accounting services across the market, and challenge data-hoarding practices that can hold back progress towards shared goals.

Transparency

Trustworthy greenhouse gas accounting is central to improved corporate sustainability reporting, and to the effective functioning of financial markets to direct built environment activities towards carbon reduction. Companies face an increasing number of mandatory and voluntary reporting standards that ask for accounting of direct emissions of greenhouse gases ('scope 1'), indirect emissions from the generation of purchased energy ('scope 2') and emissions generated within their value chain, both upstream and downstream ('scope 3')[38], alongside wider reporting on biodiversity, social and governance issues. An increasing number of investors are drawing on company-reported sustainability data to shape their portfolios towards low-carbon futures, and because of concerns about the impact of both climate change, and potential regulatory changes, in creating 'stranded assets'. A stranded asset is something that once had value or produced income but no longer does, often due to some kind of external change. Greenhouse gas accounting data will also be a vital component of any future carbon tax regimes.

Well designed reporting standards, that are informed by consideration of the data infrastructure through which reporting will be generated, can incentivise the creation of robust data systems inside firms, that not only serve external reporting, but that also support in-house analysis and decision making. However, a proliferation of reporting standards, and a limited focus on data-infrastructure design, can create barriers to effective reporting and benchmarking: This can undermine the quality of reported data and even incentivise 'greenwashing'. At present, considerable effort goes into simply extracting data from reports. If that effort could be instead harnessed to verify, audit and quality assure data, the impact of the corporate sustainability data ecosystem could be increased.

[37] Circular Ecology (2021), Embodied Carbon Footprint Database

[38] ODI (2021), Accelerating progress on tackling the climate crisis through data collaboration

Case examples: supporting reporting transparency

Bloomberg[39], Refinitiv[40] and other financial-market information providers offer products based on standardising the data from company environmental, social and governance reports. Limited data standardisation at source creates the need for intermediaries to analyse published reports, and raises the costs of access to this data.

Arup's Beacon platform draws on companies' financial information to calculate Scope 1, 2 and 3 emissions across a whole supply chain, identifying areas to reduce both operational and embodied carbon emissions, and supporting greenhouse gas reporting[41].

The Open Footprint Forum[42], hosted by the Open Group and co-chaired by leaders from ERM and Shell, is working to develop *'a common model for footprint-related data covering all types of emissions, consumptions (e.g., water, land, energy), and base calculations to normalize and aggregate data'* implemented through an open-source based reference software stack.

Circular Ecology list a range of greenhouse gas emission calculation tools designed specifically developed for the buildings sector[43], though notes that not all are easily or openly accessible.

[39] Bloomberg (2021), ESG Data

[40] Refinitv (2021), ESG data

[41] Arup (no date), Beacon: a new methodology

[42] Open Group (no date), Open Footprint Forum

[43] Circular Ecology (2021), Carbon footprint calculators for construction

Transparent reporting should not just be about compliance, but should recognise the potential of shared data for benchmarking and sector-wide learning. For example, shared and structured reporting on project-level data, as opposed to company level, can support identification of data-backed targets for new building projects, informing both contractors and commissioners. Data from certification schemes, and voluntary data contributed to projects such as the Royal Institute of Chartered Surveyors nascent Building Carbon Database[44] can help practitioners to compare the performance of a planned project portfolio with that achieved by peers, supporting conversations about ways to further optimise designs.

Data is already enabling a move to project and asset level reporting. Initiatives such as the Spatial Finance Initiatives GeoAsset project[45] are generating open datasets with information about built assets, their locations, and ownership in order to support both financial markets, policymakers and non-governmental organisations drive a shift towards greater sustainability. The first GeoAsset dataset covers cement, iron and steel production plants, drawing on multiple data sources, remote sensing and machine learning in order to exceed the approximately 70% coverage that proprietary asset databases tend to reach. Global asset-level reporting is able to leverage the considerable progress made in recent years on data infrastructure to provide unique identifiers for companies, and to link this to corporate structure information (see below). Bringing greater transparency to the ownership, location and condition of existing assets is also critical to preparing for climate shocks, and planning for climate adaptation.

Resilience

Even with action to reduce emissions, the built environment will have to adapt to a changing climate. Data is vital to understand and increase the resilience of buildings and infrastructure. For example, the Climate Just mapping tool draws upon multiple public datasets to provide localised assessments of populations vulnerable to flooding, household overheating, or fuel poverty, as climate change impacts hit[46]. Work by University College London used the tool to explore overheating risk in care homes, mapping care home locations and exploring the interaction of urban heat islands, population age and building characteristics[47]. Data-driven research such as this can inform short and long-term campaigning and planning to avoid climate harms disproportionately affecting the most vulnerable.

---

[44] RICS (2019), RICS Building Carbon Database

[45] Spatial Finance Initiative (no date), GeoAsset Project - Greening Finance and Investment

[46] Climate Just (no date), Mapping tool page

[47] Climate London (2018), Mapping heat vulnerability in London

Case examples: open data for resilience

Since 2011 the Open Data for Resilience Initiative, run by the Global Facility for Disaster Reduction and Recovery, has been encouraging greater data sharing, collection and use for resilience and disaster response, with an emphasis on geospatial data. The project field guide explores how to bring together multiple stakeholders from governments, universities, civil society and private sector organisations to collaborate in building an open data ecosystem. As the guide describes *'OpenDRI catalyzes a change to mindset: it builds a community of practitioners who apply open data to their daily problems, and in so doing, creates a sustainable ecosystem around a living and growing corpus of data that describe a dynamic society'*[48].

Insurers have a key role to play in pushing for resilience of the built environment: with the potential to mobilise billions of dollars of preemptive finance for risk reduction. In 2019, work began on a set of open data standards for describing insured assets, such as buildings and infrastructure[49], and modelled loss reports used in setting premium. This has the potential to both support adoption of open risk models, and greater portability of data between insurer systems and other parties.

One pathway towards more resilient buildings involves the introduction of smart systems, able to respond to greater pressure on energy grids through adapting demand and supply of power. The adoption of open standards for data communication within smart buildings, whether closed, shared or open, is vital to avoid vendor lock-in, maximise the value of the data generated, and ensure smart features remain interoperable as technology develops. US-based Project Haystack, which seeks to *'standardize semantic data models and web services with the goal of making it easier to unlock value from the vast quantity of data being generated by the smart devices that permeate our homes, buildings, factories, and cities'* responds to this challenge by providing a set of agreed taxonomies that can be used to provide compatible datasets and API endpoints for building systems such as lighting, heating and air conditioning[50].

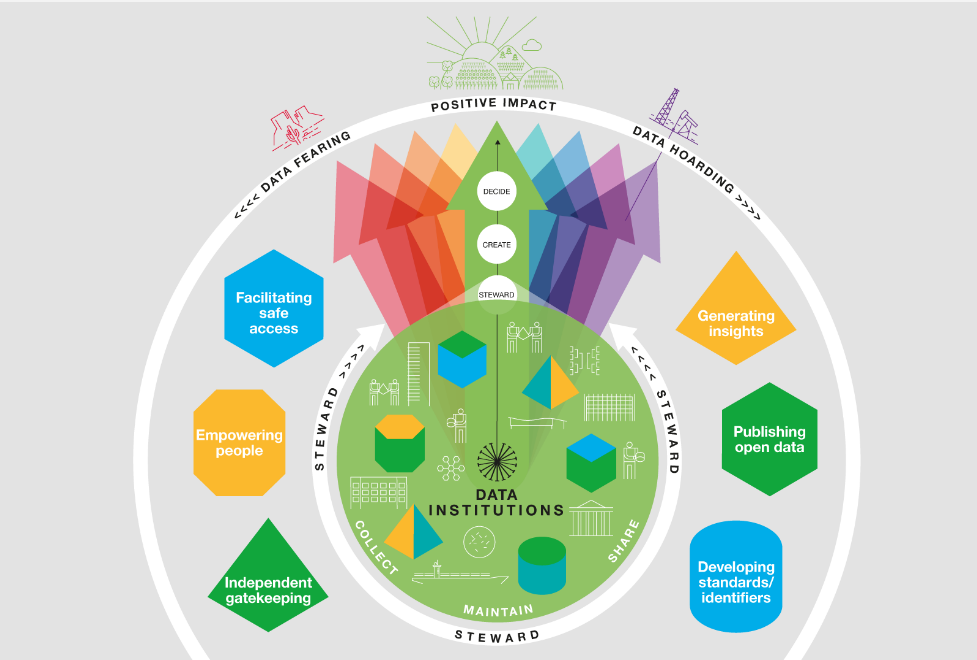
As the smart-building case shows, making effective use of data to address climate change in the built environment isn't all about creating vast central datasets, or using the latest smart technology. Small, standardised and decentralised data can help put individuals and communities in control of charting a low-carbon future, enabling the combination of local skills and knowledge with data-derived insight.

[48] Open DRI (2014), Open Data for Resilience Initiative: Field Guide

[49] Oasis Loss Modelling Framework (no date), Open Data Standards

[50] Project Haystack

5. Approaches for creating open and trustworthy data ecosystems



The Open Data Institute theory of change

The ODI wants those who steward data and those who create information from this data to act in ways that lead to the best social and economic outcomes for everyone.

There are three activities that create impact from data: stewarding data under which we mean collecting, maintaining and sharing data; creating information from that data in the form of products and services, analyses and insights, or stories and visualisations; and deciding what to do informed by information from multiple sources along with experience and understanding .
Each of these activities can be carried out by individuals, companies, communities or governments.

The ODI's theory of change identifies the threats that can stop positive impact from data. One threat is a data-hoarding approach, where misconceptions about the value of data leads to monopolies, reduced innovation and invasion of privacy. Another is a data fearing approach, leading to missing, inaccurate and non-representative data, potentially because data subjects give false details, or withdraw consent for data use.

To combat these threats, we need to build openness and trustworthiness into each data ecosystem, which will help us to realise the greatest social and economic value from data for everyone.

More specifically, this means applying an open culture to the three pillars of how data is used in our economies: a data infrastructure that is as open as possible; data literacy and capability for all; and open innovation. Equally, it means putting in place practices that ensure the way data is collected and used is trustworthy and just, including building ethical considerations into data collection, management and use; ensuring there is equity around who accesses, uses and benefits from data; and engaging widely with affected people and organisations.

The Exponential Roadmap, a call to action setting out 36 solutions with the potential to halve emissions by 2030, argues that *'digital services have potential, tenfold their footprint, to reduce energy and materials across the economy and could directly enable a third of the emissions reductions needed by 2030.'*[51] Yet, to harness that potential demands substantial actions that can move us from a world still largely defined by data silos, disjointed data access, and missed opportunities for data use – to one with a vibrant and trusted data ecosystem for the built environment.

A vibrant data ecosystem involves a constellation of **data infrastructure, data literacy and open innovation. Data infrastructure** encompasses crucial data assets, standards, technologies, guidance and policies. Data infrastructure is supported by organisations that govern the data infrastructure, including **data institutions or data access initiatives** that bring together multiple stakeholders to solve a common social, environmental or economic challenge. These components are all necessary to enable diverse independent stakeholders to operate and interact within a culture of open and shared data.

In this section we explore learning from other sectors and domains, to highlight elements of the way forward. For each example, we have sought to draw out some of the key features that may provide inspiration to action on climate change in the built-environment sector.

[51] Falk, J. and Gaffney, O. (2020) Exponential Roadmap: Scaling 36 Solutions to Halve Emissions by 2030 (Version 1.5.1).

Data infrastructure

Data infrastructure consists of data assets supported by people, processes and technology[52]. The data assets at the heart of any given infrastructure vary: technical schemas for data exchange; metadata registries for locating relevant data; data lakes and processing environments designed for specialised datasets; and core reference data and identifier lists provided via APIs and downloads to improve interoperability.

Developing infrastructure in the open

The ODI has articulated seven design principles for data infrastructure: design for openness; build with the web; respect privacy; benefit everyone; think big but start small; design to adapt; and encourage open innovation[53].

Since 2013, the **Open Contracting Partnership** has been working to develop a global data infrastructure to enable transparent and effective public procurement, and supporting its adaptation and adoption in different sectoral and country contexts. At the heart of the open contracting data infrastructure is the **Open Contracting Data Standard (OCDS)**, a data specification that describes how to represent each stage of public procurement processes as structured data. Prior to the introduction of the standard, it was often hard to join up the data on what governments planned to buy, what they tendered for, the contract awards made, and final delivery of goods, services or works. Building on a design process that considered private sector, public sector and civil society use cases for procurement data, OCDS introduced the idea of a common identifier to tie together stages of procurement, and set out key fields of data that existing systems should map to. The first proof-of-concept prototype was built in just 10 days, followed by a year-long series of development sprints commissioned by the World Bank, each one delivering an iteration on a prototype for comment and feedback.

[52] ODI (no date), Data infrastructure

[53] ODI (2016), Principles for strengthening our data infrastructure

Figure N: What can be covered by data standards?
Source: http://standards.theodi.org/introduction/

As a result of this open and iterative development process, before OCDS was even finished, the data structures it provides were forming the basis for Ukraine's new ProZorro born-digital procurement data infrastructure. This is estimated to have saved upwards of $1.9bn against budgeted spend[54] through better purchasing and reduced fraud and corruption (a significant return on investment given the estimated €4.7m cost of establishing the ProZorro system[55]). Now around $363bn of annual public contracting is described in standardised data[56]. Open Contracting Partnership (OCP) research found that several countries that had used OCDS to shape their procurement data infrastructure were able to respond more readily to the pressures of emergency Covid-19 procurement, using transparency as a tool for intelligent purchasing, rather than seeing it as a bureaucratic hurdle[57].

From the start, OCDS was designed with flexibility in mind: seeking to make data interoperable enough to support re-use of analytical tools across countries and contexts, but also recognising the different priorities, risks and opportunities around public procurement that exist in each country. A key part of the OCP's role as steward of OCDS is to engage with the community of users, and spot

---

[54] Organisation for Economic Cooperation and Development, Observatory of Public Sector Innovation (2016), eProcurement system ProZorro

[55] Results for Development (2017), Costing the ProZorro e-Procurement Program

[56] Open Contracting Partnership (2020), Global Procurement Spend

[57] Open Contracting Partnership, Global procurement responses to COVID-19: how to do better in an emergency

opportunities for extending or updating the standard so it continues to support learning, collaboration and innovation.

> Technical data standards to enable the built-environment sector to respond to the climate crisis stand to have greater impact if developed through open, agile approaches – learning from models like OCDS that bring together policy and technology.

Creating appropriate infrastructure

Effective standards and infrastructures are built with a clear understanding of the daily experience of the community that will be using them.

The Flexible, Appropriate, Structured, Transparent (FAST) Standard[58] provides a set of rules on the structure and design of spreadsheet based models, particularly for the financial sector. Managed and maintained under an open licence by a UK non-profit organisation, the standard is focused on creating models that are *'as simple as possible, but no simpler'* by setting out rules for workbook organisation, sheet layout, and formula design.

Rather than encouraging creators and consumers of financial models to move to new tools and platforms, the FAST Standard sets out to 'tweak' existing Microsoft Excel use practices, with the goal of reducing complexity in how data is expressed (even if there is considerable complexity to the underlying model).

> Recognising the widespread use of spreadsheets for data exchange in built environment projects, standards like FAST may provide inspiration for ways to strengthen good data-management culture and practice, and to increase the usability of spreadsheet-based modelling tools.

Data institutions

Data institutions are organisations whose purpose involves stewarding data on behalf of others, often for public good. Well designed data institutions can provide trustworthy data governance, aligning the different stakeholders' incentives and solving collective action problems that would otherwise lead to an underinvestment in data, limited data sharing, or unsustainable data supply.

In recent years, significant work has gone into exploring the range of institutional models for data stewardship, from trusts with fiduciary responsibility to protect personal data; to co-operatives that seek to work to ensure benefits are shared between all those who contribute data; to commons, created through use of appropriate data licensing. Whether it's a small reference dataset, or a large global data infrastructure, often more than the technical infrastructure is required to provide the foundation for collaboration and innovation at scale.

[58] The FAST Standard Organisation

Connecting governance, regulation and technology

Data institution design involves work to align governance, sustainable funding and technical architectures.

The **Global Legal Entity Identifier (LEI)**[59] was created in the wake of the 2007/8 financial crisis, following recognition of the systemic risks created through patchy data on the market participants actually involved in trades, derivatives and other financial products. Prior to the LEI, financial data systems either had to accommodate hundreds of different local company identifiers, or to make use of proprietary 'black box' company identification products. The LEI is an identifier that firms and other market participants register through a Local Operating Unit (LOU), licensed to issue identifiers in a given jurisdiction or sector by the Global LEI Foundation's (GLIEF). The GLEIF itself is governed by a Regulatory Oversight Committee and operates under globally set principles and a charter. Over time, standards for structured data attached to each Legal Entity Identifier have been developed by GLEIF, allowing the central infrastructure to aggregate and provide API access to the details and corporate structure of over 1.8m financial sector firms and actors.

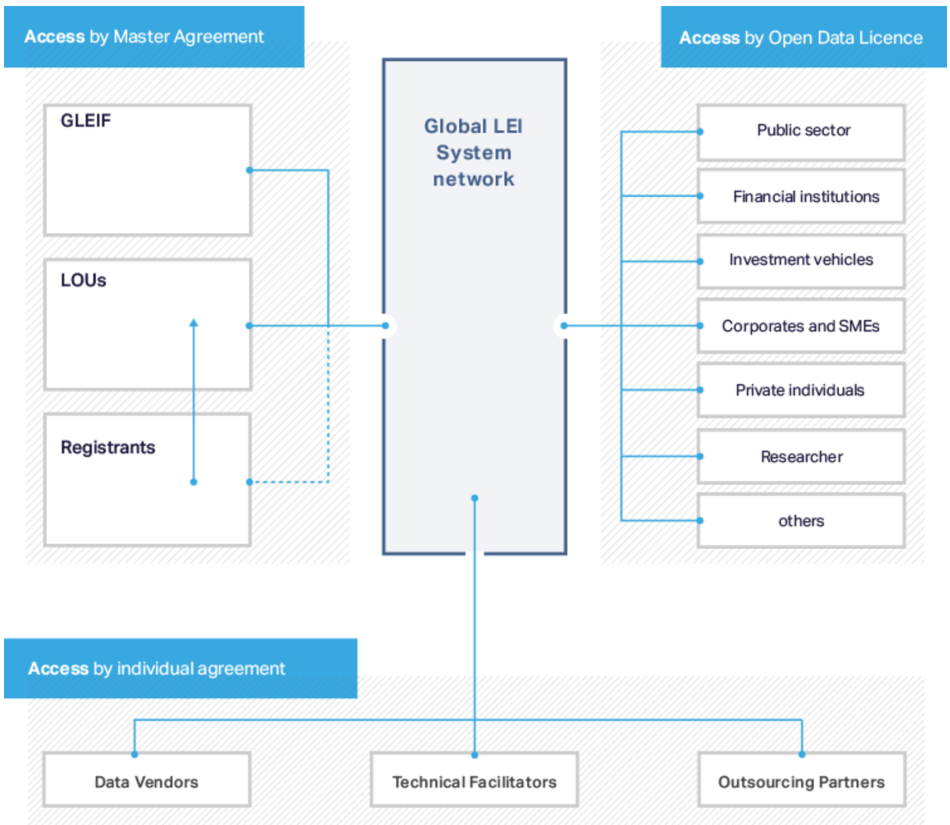


Figure N: the Global LEI System as a Legal Network (Source: GLEIF.org)

Early advocacy during the set-up of GLEIF ensured that the LEI system is committed to open data principles for data access, while it has also developed a model of registration fees, creating financial sustainability for the system. A 2018 review notes the substantial progress made by the system, and its potential, but cautions that it will take continued work from regulators and senior management

[59] Global Legal Entity Identifier Foundation

for the LEI to become *the* standard identifier in financial markets, displacing legacy data practices[60].

> Global challenges for built environment firms around common carbon measurement and reporting may call for substantial work on designing governance institutions, and securing regulatory change, to align different stakeholders around common data standards and data sharing processes.

From scraped datasets to trusted open infrastructures

There are cases where existing data projects have migrated into more formal data institution structures over time, in order to maximise their contribution to the public good.

Before the intergovernmental efforts to develop the LEI had even begun, many people were already grappling with the challenge of how to bring together data about companies, and to understand and map supply and value chains. **OpenCorporates** was first launched in 2010 as an open database of information scraped from company registers, hosted by a small independent company[61]. As the platform has grown, it has become a critical resource for journalists, NGOs and researchers who are often unable to afford alternative black-box commercial data products. In 2018, ownership of OpenCorporates was transferred into a Trust[62], allowing the platform to continue to generate commercial income, but locking in a public-benefit business model that allows others to build on the platform without fearing that access to the data for public good uses could be removed in future[63].

> There are numerous (and important) elements in the data ecosystem that support climate action around the built environment that lack stable institutional foundations. Identifying ways to improve the sustainability, support for, and guaranteed open licensing of key datasets could unlock significant new activity.

[60] Journal of Risk and Financial Management (2019), The Global Legal Entity Identifier System: How Can It Deliver?

[61] State of Open Data (2019), Corporate Ownership

[62] Open Corporates (2018), Announcing the OpenCorporates Trust

[63] Open Corporates (no date), Governance

Data access and use initiatives

Many thriving data infrastructures and institutions are embodied within, or provide the foundations for, wider **data access and use initiatives**. These initiatives seek to catalyse action around specific social, environmental or economic problems, and bring together multiple stakeholders in order to address the particular data collection, sharing and use parts of the problem space[64]. Initiatives take many forms, from ad-hoc coalitions, to long-term programmes with high-level policy backing.

Improving data discovery

Icebreaker One was launched in 2019 as a non-profit organisation with a vision to 'empower decision-makers to mandate, measure and act upon the data-flows that enable net-zero'. Through scoping research, rapid prototyping and convening stakeholders, the project quickly identified a need to address data sharing in the energy sector.

Thousands of different entities are involved in the energy generation and distribution market in the UK: from generators and distributors, to regulators, suppliers, service providers and consultancies. Many of these actors generate, hold or need access to data. Better data sharing could accelerate a low-carbon energy transition, but, as the Icebreaker One Open Energy initiative[65] identified, it's often difficult to discover when data is available, and once discovered, to work out the terms on which it might be accessed. As a result, the Open Energy initiative has worked with the energy industry and other stakeholders to develop a federated data search, backed up with clear metadata standards that support 'preemptive licensing'. This gives data holders a structured language within which to express that, for example, they are willing for a particular dataset to be directly accessed and freely used by academics or public sector users, whereas commercial users should get in touch to arrange charged access, supporting sustainable cost recovery for data maintenance.

The technical specification at the heart of Open Energy is not, however, just a tool for metadata exchange: it also supports conversations with energy sector stakeholders about how they can improve their data stewardship. To see data included in the catalogue, organisations need to think about the data they hold, have conversations about the terms under which it can be shared, and consider the data from others they may potentially draw upon. By creating a meeting place and marketplace, the project surfaces data demand, and highlights gaps where further data institution building may be needed to improve the flow of data. The Open Energy project has also been exploring the incentives that different stakeholders need in order to participate: ranging from recognition to regulatory mandates.

[64] ODI (2020), [What do we mean by data access initiatives?]

[65] http://energy.icebreakerone.org

There is a wealth of built environment data that could be harnessed to tackle climate change if it were more discoverable. Declaring the data assets that are available, and exploring the terms under which they could be more widely shared, can be a powerful early step in surfacing the value of data, and catalysing greater focus on data stewardship.

Fostering innovation and cooperation

Climate Subak gets its name from a ninth century cooperative irrigation system used in Indonesian rice fields that *'allowed for the sharing of critical resources (water) to ensure the survival and growth of the community.'*[66]. Recognising data as a 21st century critical resource that is vital to ensure survival and growth in the face of climate change, Subak has created a technology accelerator programme, aiming to stimulate data use that both engages with, and feeds back into, a rich climate data infrastructure.

As well as providing funding and support to data-driven climate-focused non-profit organisations to enable them develop and scale ideas that drive faster and better climate action, Subak asks each organisation they work with to commit to a **data pledge.** The pledge seeks to create a data cooperative ecosystem, in which all the data gathered by accelerator members (and other voluntary signatories) becomes available to other climate-focused projects. It includes commitments to take responsibility for data quality, standardisation and interoperability. It also places emphasis on data accessibility, recognising that attention is needed to curating complex climate data so that each datasets is constructed *'with a user experience, in which interested people irrespective of skill set would be able to understand its relevance."'*[67]

Built environment stakeholders should explore how to incentivise small- and large-scale innovation with data, while also building collaborative and commons based solutions. They should not put effort into competing initiatives that seek to hoard rather than share data.

[66] Subak (no date), About Us

[67] Subak (no date), The Data Cooperative

Data cultures: independent actions in a vibrant ecosystem

While infrastructure, institutions and initiatives provide the groundwork to realise value from data, it is action from many different individual entrepreneurs, firms, nonprofits and forward-thinking government agencies that leads to the growth of thriving data ecosystems. This requires people with the capacity and interest to engage in collaborative data projects, and community-building efforts that join the dots between different stakeholders.

The journey towards data culture

Over the last 20 years, the international development field has progressively embraced data as a powerful tool for planning, prioritising and monitoring activity, and has developed a widespread culture of open data and data use. It has invested in capacity building in large institutions and grassroots organisations. Driven initially by a transparency agenda, the international development landscape now hosts a wide range of actors producing data standards, providing data sharing platforms, advocating for improved data quality, and investing in support for open data activity.[68] One community-building node in this ecosystem is The Global Partnership for Sustainable Development Data which provides a network bringing together hundreds of organisations – from national statistical agencies, to two-person startups – to support peer learning and knowledge exchange[69]. National, regional and global fora also provide a space to showcase projects, to secure ongoing high-level support for data programmes, and to deepen networks between individual practitioners.

> For built environment firms and individuals used to solving their data problems alone, or with black-box off-the-shelf products, moving to a world in which they have the ability to draw on data as a shared or common resource for problem solving can involve considerable culture change, not to mention calling for new skills and working practices. Accelerating this culture change will take dedicated focus.
>
> It is also worth noting that there can be significant differences between the professional cultures of civil engineering and software engineering, or between the forms of dialogue and action of climate activists and construction specialists. Building deeper open and shared data cultures to respond to the climate emergency will require two-way dialogue between data specialists, and built-environment experts, understanding the knowledge, experience and insights that each brings.

[68] Zenodo (2019), Open Data Sectors and Communities: Development Assistance and Humanitarian Action

[69] Global Partnership for Sustainable Development Data

6. Summary and call to action

The need for urgent action on climate change is no longer a matter of debate. Nor are we short of viable changes that can be made that would substantially reduce short- and long-term energy demand. The question now for the built-environment sector is how to accelerate action to reduce emissions, build for resilience and protect biodiversity. We need to use all the tools at our disposal, including the significant power of data.

Meeting this challenge calls for coordination, cooperation, community-building and culture change across the built-environment sector, and for cross-sector collaboration to break down data silos. When stakeholders come together to identify shared challenges, ambitious goals, and clear plans, they can mobilise resources to deliver rapid change at scale.

In this report we've surveyed a number of particular challenges, including bringing together data on the existing built environment to support strategic action and investment, and unlocking greater building and house-level data to support individual choices around energy saving and retrofit planning. We have identified the need to improve standards and working practices so that digital representations of projects during the planning stage can be better used to drive low-carbon design. We have highlighted challenges relating to reliable and trustworthy carbon calculation, and pointed to the potential for modern data stewardship practices to address some of these. And we've looked at the importance of bringing in data from other settings and sectors, working across professional and organisational boundaries.

Some of the challenges are technical, some are political, and others are around culture and capacity.

None can be solved by organisations acting alone.

We are calling on data creators, data stewards, researchers, data users, and those whose actions could be shaped by better access, sharing and use of built environment-related data, to join us to chart a shared route map for 2022 and beyond.

This paper launches a call for collaboration – we are asking organisations and individuals to take action and:

- **Give us your feedback on this working paper** to help focus future data collaborations and activities by filling this form between now and 19th of November

- **Sign up to collaborate with us –** making 2022 a year of action on built-environment data to tackle the climate crisis.

- **Tell us about existing data-sharing and open data projects, priorities and needs** to help build a shared map of initiatives and opportunities.

- **Commit to identifying how you can play a role in improving data flows in support of the priority areas in the built-environment sector**, building a deeper culture of data sharing to address climate challenges.