

United Utilities Open Data Strategy

2023

[Introduction](#)

[Our ambition and aims](#)

[What is open data?](#)

[How we will provide and use open data](#)

[Our open data guiding principles and values](#)

[Open data maturity model](#)

[Open data framework themes](#)

[Case study | Providing open data to support our environment](#)

[Case study | Using open data to simplify our customer journey](#)

[Case study | Collaborating across the sector with open data](#)

[The benefits of open data](#)

[Ongoing targets](#)



The purpose of the open data strategy is to provide a structured framework and plan for the publication and use of open data within United Utilities.

In this document we describe the main components, scope and deliverables of our open data activity and outline the benefits and timescales for implementation. Our strategy acknowledges the potential societal, environmental, efficiency and innovation value which can result from publishing and using open data. It provides a framework for us to ethically and securely enable an open data value chain aligned to [our purpose, strategic priorities and core values](#).



© 2023. This work is licensed under [CC BY-SA 4.0](#)

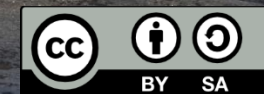


Our ambition for United Utilities is to be a data-driven, open data company.

That means not just using open data provided by others, but also publishing open data with a purpose, led by what our stakeholders and customers need.

We know that doing this will generate value for United Utilities, the water sector, our customers and others by:

- Improving efficiency and service to our customers.
- Improving openness, transparency and trust in our operations from our customers and stakeholders.
- Improving how we work with partners, including other utilities.
- Demonstrating the value of open data as an agent to drive innovation and economic value.
- Delivering improvements and benefits for the environment and society as a whole.



The Data Spectrum

Access to data exists on a spectrum, from **closed** to **shared**, to **open**, as shown by the Open Data Institute (ODI) [Data Spectrum](#).

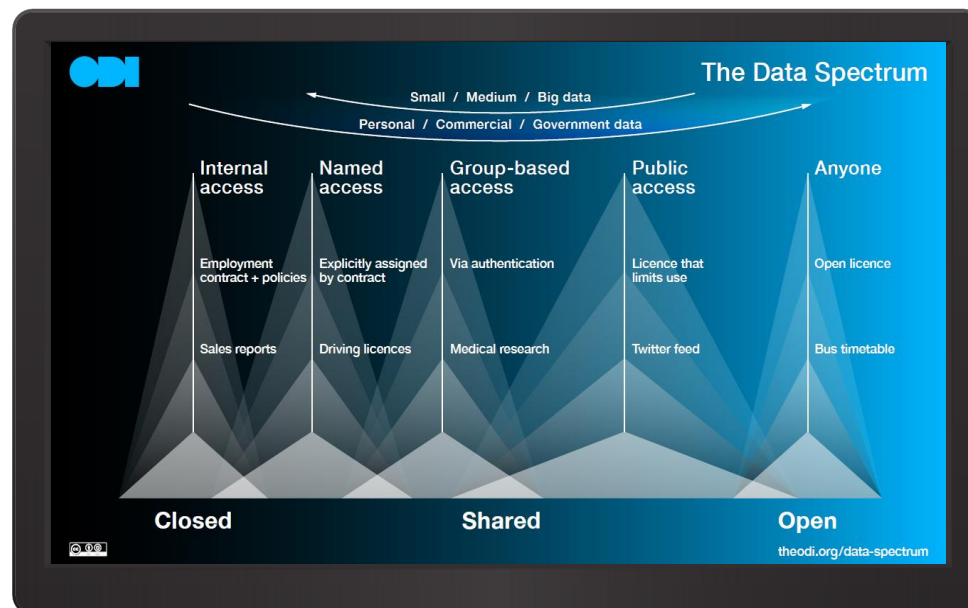
Open data refers to data that anyone can access, use or share.¹ Such data should be published with an open licence,² for example, the UK's [Open Government Licence](#), or those available from [Creative Commons](#). Catalogues of data can also be created and opened, and used to request access to the underlying datasets themselves.

Most data in the water sector falls under the umbrella of **shared data**. This is generally understood to be data that is accessible beyond where it was collected or created, but is not published with an open licence. Shared data has the widest range on the Data Spectrum, from data provided by one entity to another under a named or group-based access arrangement, such as through a data sharing agreement or commercial data portal, to data published on the web under licences that limit use, such as [non-commercial](#) or [no-derivatives](#) licences.

It is important that some data within the water sector remains **closed**, meaning it is not shared outside of the organisation that collected it. This is to protect privacy, commercial interests or national security as and where relevant. However, even if the raw data itself remains closed, aggregated or anonymised versions of it can potentially still be shared or made open.

¹ [Glossary \(2022\) | The ODI](#)

² [Publisher's Guide to Open Data Licensing \(2013\) | The ODI](#)



We recognise that the data we hold is a valuable asset that can be put to wider use. Our open data strategy sets out how we intend to use open data to deliver value, whether that's economic, social or environmental.

We will do this by releasing a series of high-value datasets and facilitating collaboration on what we share through our open data action plan. We will also develop an open data culture with our colleagues and with stakeholders, including commercial, research and public organisations and the public.

Our strategic objectives are to leverage data sharing and open data opportunities via innovation challenges with external stakeholders and partners to unlock new insights and drive innovation. We want to use shared and open data to drive situational awareness of the water and wastewater ecosystem and increase data driven decision making.

Through implementation of our strategy, we want to create a well-governed and efficient open data framework that enables innovators, both within and outside our company, to help address key water sector challenges.

With these activities in mind, we are designing our open data action plan to:



Increase efficiency and insight into how we deliver our services, through improving accessibility, use and combination of open datasets.



Increase business innovation, creativity and promotion of economic growth and value.



Improve the transparency of our decisions and operations, leading to increased trust and reputation with customers and stakeholders.



Improve standardisation of shared data, increasing the opportunity to improve data quality through preparation for publishing and subsequent feedback.



Make consideration of open data opportunities part of our way of working.



Improve our ability to report and encourage open innovation to power products, services and insights designed to increase efficiency and positively impact our environment and our contribution to net zero.



In 2022 we commissioned the Open Data Institute (ODI) to support the development of our open data capability.

At this time we were already active in using open data as evidenced in the Ofwat publication [H₂Open](#), and had published a number of open and shared data items in response to individual drivers.

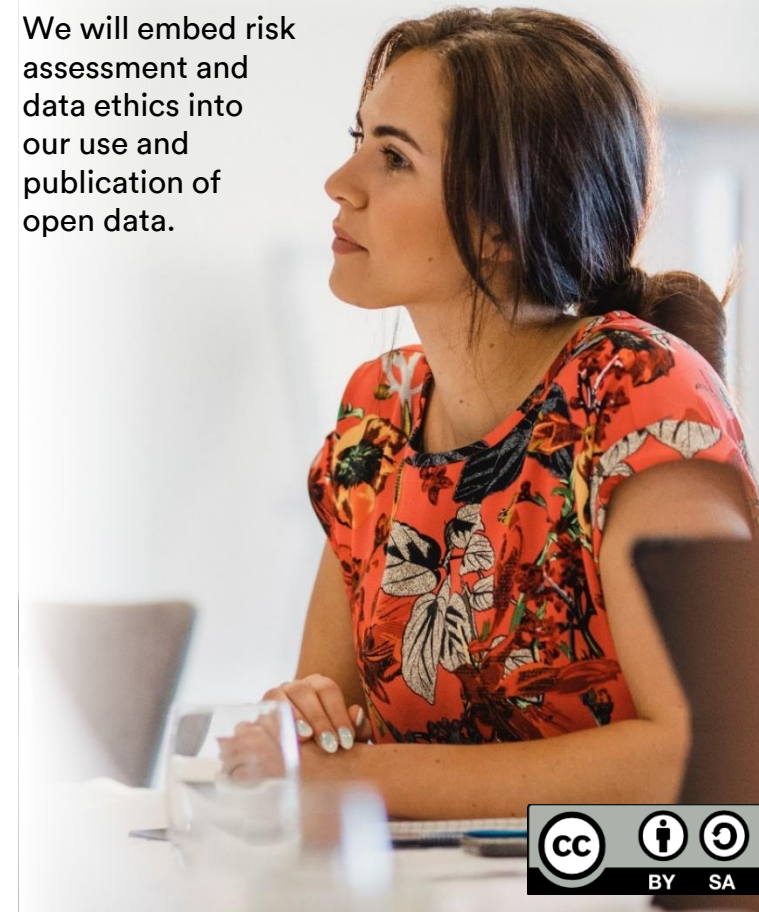
However, the work with the ODI was instigated to accelerate progress on opening targeted datasets and establishing the culture and practice of open data within the organisation. We focused on key priorities to increase our open data maturity by:

- Engaging our senior leadership team to set strategic direction and vision.
- Engaging stakeholders and existing United Utilities data consumers to understand their open data needs.
- Engaging colleagues to identify opportunities to support open data publication.
- Creating our open data strategy and delivery plan.

Our strategy will deliver the following open data principles and values:

- The quantity of datasets will grow over time, selected by relevance and value-add, in line with our open data ambition and aims.
- We will publish open data with a purpose led by the needs of our customers, stakeholders and wider society.
- Our open data will be freely available and of high quality (complete, accurate, valid, of known provenance, maintained and updated at appropriate intervals).
- Our open data will be published using a suitable open licence.
- Our open data will meet the requirements of an appropriate technical framework (standards, formats, metadata and licence) to ensure compliance and interoperability.
- Our open data will adhere to the FAIR principles (findable, accessible, interoperable and reusable).

- Our open data will be aligned to our information security and management policies, design principles and will be guided by our open data policy.
- We will embed risk assessment and data ethics into our use and publication of open data.



In order to support our action plan, we have used the findings from Ofwat’s open data assessment and conclusions and the ODI’s [Open Data Maturity Model](#) to assess where we stand and the actions we can take.

The model provided a progress framework based around five themes:

- 1 Data management processes**
Key business processes that underpin data management and publication including quality control, publication workflows, and adoption of technical standards.
- 2 Knowledge and skills**
The steps required to create a culture of open data within an organisation by identifying the knowledge sharing, training and learning required to embed an understanding of the benefits of open data.

- 3 Customer support and engagement**
The need for an organisation to engage with both their data sources and their data re-users to provide sufficient support and feedback to make open data successful.
- 4 Investment and financial performance**
The processes and mechanisms needed by an organisation to understand the cost and benefit generated by both the release and use of open datasets.
- 5 Strategic oversight**
The need for an organisation to have a clear strategy around data sharing and re-use, and an identified leadership with responsibility and capacity to deliver that strategy.

We have organised our action plan to ensure it aligns with these five themes.



1

Data management processes

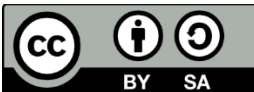
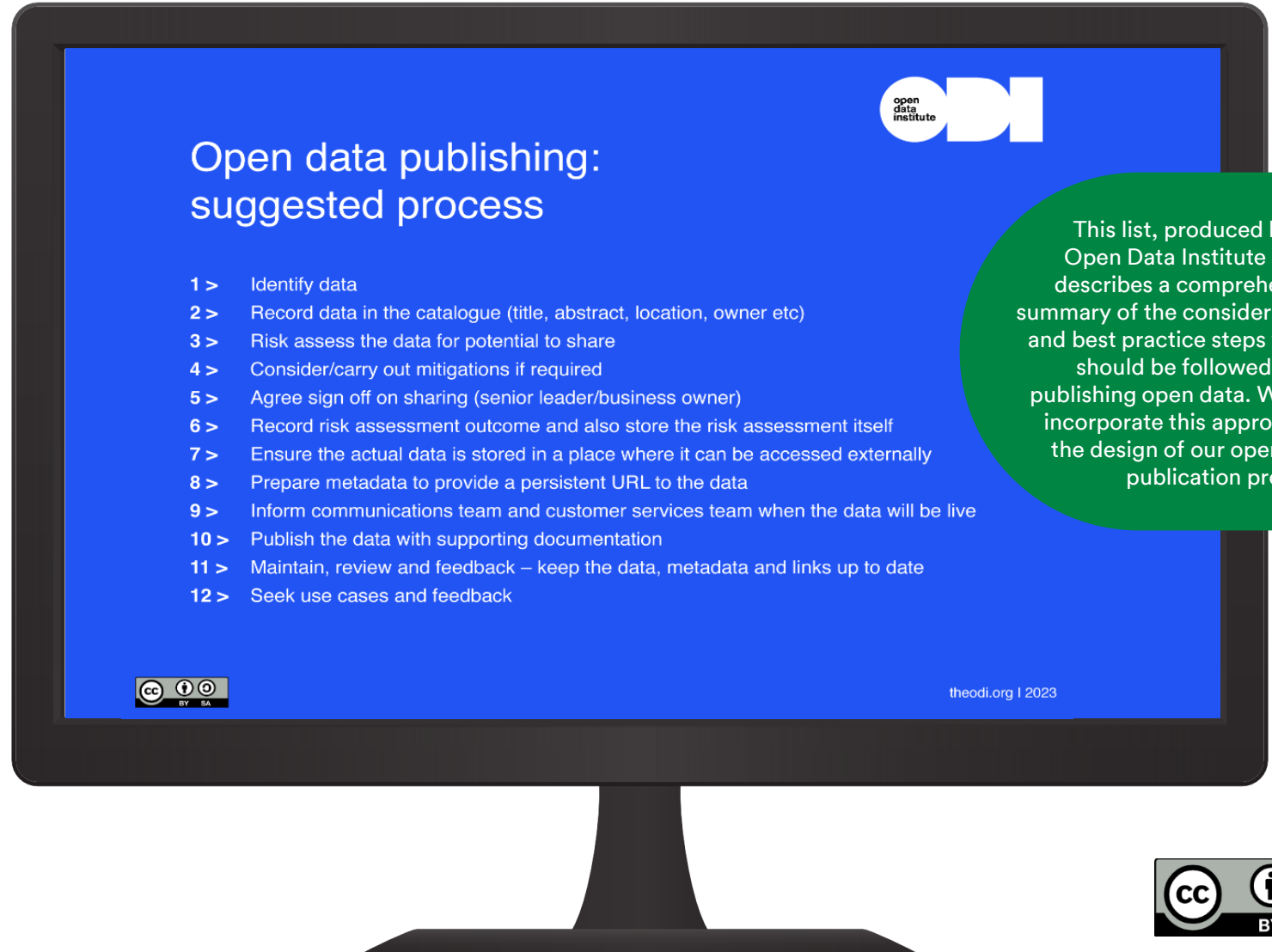
The processes that underpin how we govern and manage the data we publish and use to ensure we comply with our information security framework, quality standards and give appropriate consideration to usability and data ethics.

By the end of 2023, we plan to:

- ✓ Embed a standardised approach to assessing risk for open datasets, using the ODI risk assessment tool which includes consideration of security and data ethics within our existing robust data governance process.

By the end of 2024, we plan to:

- ✓ Publish a further three open datasets, following the publishing of our first open dataset from our new open data initiative in 2022.



2

Knowledge and skills

The education we need to provide to develop and embed understanding and grow our ability to create and innovate with open data.

By the end of 2023, we plan to:

- ✓ Design an Open Data Masterclass course in collaboration with the ODI to help embed an open data culture within United Utilities. This course will maintain engagement and support from the senior leadership on our open data maturity journey. It will focus on risks and myths, benefits and opportunities, case studies of successful open data projects, and how to engage teams to discover and use open data.
- ✓ Grow and further develop an internal citizen data science community through a series of data science modules targeted at a variety of internal subject matter experts. This training covers the fundamentals of data science, to enable them to identify the types of problems data science can help resolve.

These colleagues will use and leverage advanced analytics and open data to address business challenges and promote data-driven decision-making.

By the end of 2024, we plan to:

- ✓ Establish an open data co-ordination role to champion, manage and further facilitate the creation and utilisation of open data throughout the organisation.



3

Customer support and engagement

A feedback loop of communication and support to ensure the quality and value of our open data.

By the end of 2023, we plan to:

- ✓ Align our open data release process with our communications strategy so that new datasets and changes to open data developments are communicated to our wider data community of partners, users and other stakeholders.

By the end of 2024, we plan to:

- ✓ Gain feedback on our data aims and priorities and shape our future open data releases through engagement with stakeholders and data users.



4

Investment and financial performance

How we intend to assess and measure the benefit and usefulness of the open data we publish and utilise to ensure cost effectiveness.

By the end of 2024, we plan to:

- ✓ Derive and apply a consistent mechanism to track the cost and benefits of open data to ensure we are adding value and publishing with a purpose.
- ✓ Launch an open data innovation challenge in order to better understand the value of the data we publish openly, and deliver greater impact from our open data initiative. This will follow on the success and learnings from our wastewater flooding hackathon which used open datasets from Public Health England to understand how certain activities may increase the risks of wastewater flooding, creating tangible benefits for customers.



5

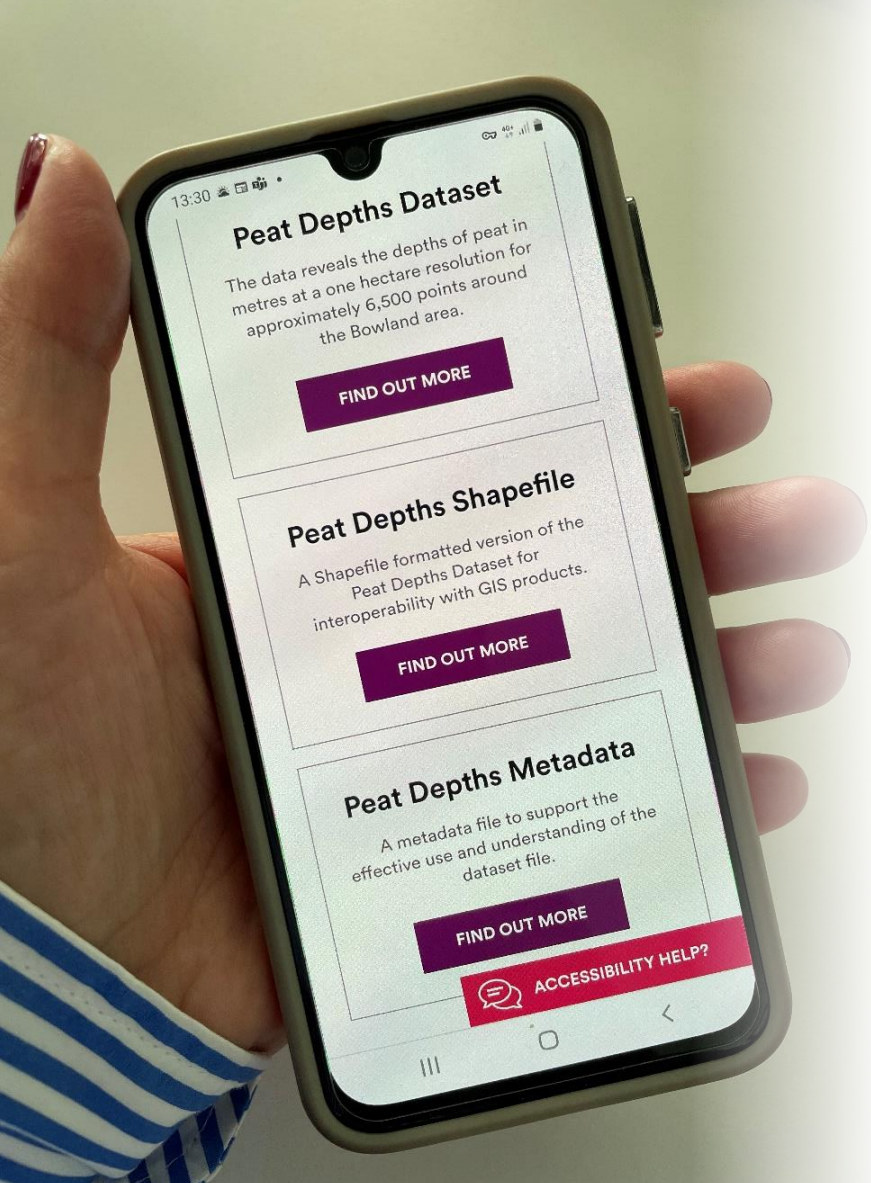
Strategic oversight

A clearly communicated governance framework for open data sharing and re-use, to support and inform our colleagues, customers and stakeholders.

By the end of 2024, we plan to:

- ✓ Create an open data policy and guidance which will inform how we intend to openly publish data and define our commitment to data sharing and the consumption of open data.





With the ability to hold more carbon per acre than a tropical rainforest, provide quality water, and home species from bugs to birds of prey – peatlands are nature’s miracle worker. In the UK, almost 12% of our land is peat, with a large portion of that lying in the North West.

Our catchment team, through collaboration with Natural England, were aware of an ongoing country-wide project to collect information on all aspects of peat to support ecosystem improvement. As a result, they undertook additional survey work as part of Bowland, Water Industry Natural Environment Programme (WINEP).

The data from this survey, which covered every part of our Bowland moorland, provided valuable, high quality peat depth information which Natural England were keen to access and utilise. This need, and the importance of the information in maintaining our peatland environment, was a key driver in our decision to publish it under open data licence in December 2022.

[Bowland WINEP webpage](#)



It’s a great website and a great example of open data... I’m hoping others will look at it and be inspired.

Natural England



Case study | Using open data to simplify our customer journey



Recognising the need for affordable bills, we have implemented a range of industry-leading support schemes including lower tariffs, capped bills, and payment matching schemes. Tailoring payment plans to customer affordability is a key goal and the advent of improved data availability from open banking stimulated an innovative method of improving our customer journey.

Our idea was to utilise open banking technology to verify customer income in real time to improve the accuracy and efficiency of our customer affordability assessments. We implemented an improved customer journey in three key steps:

Open data now forms a key part of one of our most sensitive customer journeys, and initial results showed 45% of customers who were offered the option to use open banking accepted.

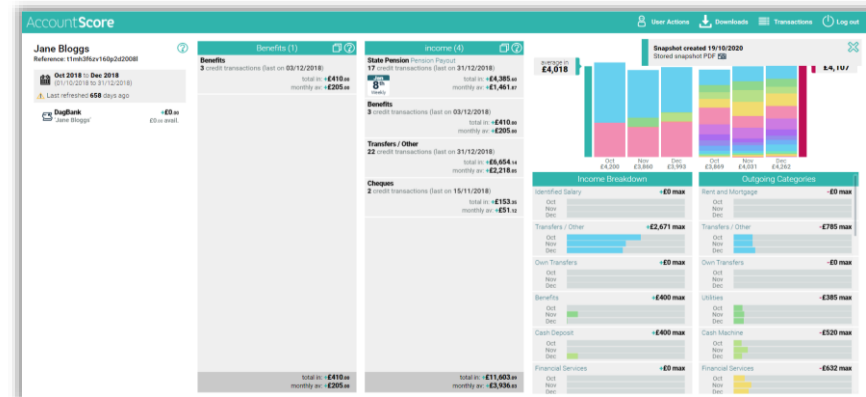
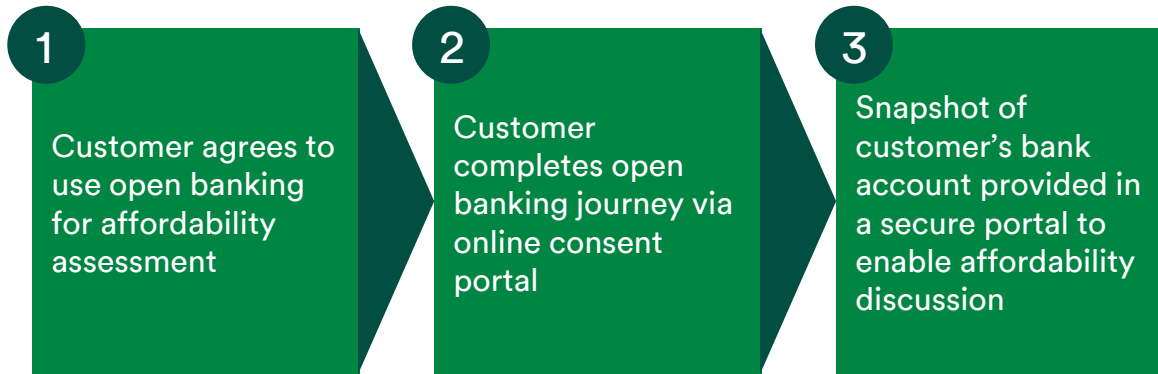
Customer feedback on their experience of open banking is very positive, with customers saying it was easy to use and 88% saying they would use it again, despite never having used it before.

The use of open banking has streamlined customer eligibility for reduced-rate social tariffs. Were it not for this solution, customers applying for help with payment of their water bill would have had to

manually collate their income and expenditure information, including evidence of benefit receipt, in preparation for their telephone affordability assessment.

What previously would have taken weeks, can now be done in minutes, with the added benefit of increased accuracy.

Open banking improves first-time completion rate, meaning customers are given a decision on tariff eligibility there and then and a sustainable payment plan can be agreed.



Much media focus has been given recently to the importance of our sewer networks and their impact on public health and the environment. Maintaining this network and improving its operation has always been a key water industry-wide objective and presented a shared challenge. We led a cross-sector innovation project to accelerate the use of artificial intelligence (AI) to improve efficiency and reduce the cost of maintaining this critical piece of our infrastructure.

When utilising AI, the data used to ‘train’ the machine learning algorithm is critical to its ability to recognise patterns and make predictions. The better the data, the more accurate and reliable the predictions made by the software. Working with our delivery partners, Water Research Centre, and in collaboration with other water companies, we have been able to gather, classify and release a wide range of sewer network CCTV footage.

By making this combined dataset available to existing and future AI software companies, the project offers a fantastic opportunity to better train machine learning software to more accurately recognise and understand sewer defects.

It is envisaged this project could enable and encourage international, cross-sector collaboration throughout Europe and the US by providing a functional capability which will directly help to move forward AI and its application in inspection.

This collaboration will provide the first global open dataset of sewer defect images for use by AI software providers working across multiple sectors.

By using AI we can:

- Reduce the time it takes to survey a kilometre of pipe by up to 20%.
- Enable a better more accurate understanding of sewer wear and tear and improve prediction of failure.
- Better target our investment to improve accuracy of maintenance and operation.
- Reduce the potential impact of sewer failure on our customers and the environment.



The research we conducted, which involved interviews and workshops with our internal data users, external stakeholders and findings shared through the sector-wide [Stream](#) open data group, helped formulate and shape our open data ambition and aims in a number of ways:

- Providing insight into the economic and societal value of data, data sharing and open data, reinforcing our commitment to increasing openness and transparency around the data we hold.
- Enabling better understanding of the needs of customers, stakeholders and consumers which helped shape and prioritise how we should implement our open data programme prioritising focus on high value use cases and datasets.
- By seeing and understanding the work being conducted in other sectors and areas we gained practical insight into methods and approaches that worked well as well as identifying potential pitfalls to avoid.

Open data in the water sector

To tackle complex challenges, including climate change, population growth and creating sustainable supply chains, we need to build open and trustworthy data environment, or ‘ecosystem’. This ecosystem must enable collection, transfer and use of data to stimulate innovation and collaboration. For customer-facing sectors such as water utilities, these societal goals must be tackled alongside ensuring that customer demands are still met.

- When data is more open, more people can access, use and share it to address problems and make better decisions. For the water sector, there are currently significant untapped opportunities from increased open data and data publishing,³ including:
 - Stimulating more innovation and collaboration.
 - Encouraging new business models and service offerings that increase efficiencies and enhance the customer experience.
 - Providing insights on company performance and assets, improving decision-making.

- Enabling companies to work collaboratively to use open data to tackle shared challenges.
- Improving transparency for customers, and the owners of companies, and building trust in the water companies.

The water sector is facing a number of current and imminent challenges, from the impact of the recent pandemic on household abilities to pay bills, to the developing expectations of customers regarding the environment and climate change.⁴

Collaboration through data sharing can stimulate innovation within our own organisation, within the water sector and beyond. It can help to develop better insights, identify more efficient solutions to the challenges faced, and build confidence among consumers.

³ [H₂Open – Open data in the water industry: a case for change | Ofwat](#)

⁴ [Time to act, together: Ofwat’s strategy | Ofwat](#)



It is well-established that data is valuable. Companies have been powering their business models from data-driven insights for decades, but only in the last ten years have we started examining the value of open data in particular.

As recently as 2020, the European Data Portal estimated that the value of open data for the EU28+ was €184bn in 2019, and forecast it to reach between €199.51bn and €334.21bn by 2025.⁵

Besides the clear economic benefits, we are hearing the need for more open and shared data in our immediate network: from customers, business ecosystem partners, local authorities, environmental groups and more.

Our research in the water sector

Through independent research, including leveraging insights from the [Independent Challenge Group](#) for United Utilities and the wider [Stream](#) consortium, we now have a good picture of the open data needs of the water sector, including the use cases and datasets in highest demand.

Key insights from the Independent Challenge Group include:

- United Utilities data is not just data about a company, but is a valuable resource for understanding the North West region as a whole.
- Often data users access and use downloadable data, visualisations and reports about United Utilities on a weekly basis not just from our own website, but also from the [Consumer Council for Water](#) and [Ofwat](#).
- Users have a good experience, and appreciate data being freely available in an accessible format with associated documentation. Users trust us as a data provider, but mention some issues with data quality and findability.
- Key data interests are company performance and ROI, customer engagement and support, and water quality and taste. Data related to flooding is also a particular interest, with a strong concern about the implications this may have on properties.

- Some of the greatest concerns around the data are the comparability and standardisation with other companies and with specific definitions. For example, spills and blockages could relate to both the relatively minor case of a brick in a sewer, or the major case of a collapsed sewer. Context is needed for clarity.

⁵ [The Economic Impact of Open Data: Opportunities for value creation in Europe | European Data Portal](#)



We need to be a part of this change, driving the potential to realise this value. Data and digital transformation are driving changes across all sectors, and the water sector is no exception. Outside of the water sector, there are many examples of how the adoption of open, collaborative approaches is helping to tackle social, environmental and economic challenges. We can learn from these sectors in developing our own approaches.

Built environment and finance

The water sector does not stand alone, and is connected both physically and digitally to other parts of the built environment, such as energy. Early adopters of open data publishing and open data infrastructure were banks and the finance sector. Both sectors drive many lessons with potential relevance for the water industry.

Energy

The energy sector has been undergoing transformational change through recommendations from the [Energy Data Taskforce](#) to move the UK towards a 'Modern, Digitalised Energy System', one of which is for energy companies to openly publish asset and network data under the principle of 'presumed open'. This principle encourages energy companies to start from a position of 'open' and then, through a 'data triage' process, consider whether there are limiting factors which may require data to be less open. This sector-wide realignment has led to other important initiatives that we can learn from, for example:

- The [Modernising Energy Data Access](#) programme has launched the [Open Net Zero](#) platform to help improve access to net zero data, which has inspired our own development of [Stream](#).
- Ofgem has implemented mandatory [Data Best Practice Guidance](#) for licensees, which is expanding both within and beyond the energy sector.
- Organisations like [UK Power Networks](#) are developing [open data portals](#) with net zero use cases already designed.
- Open Climate Fix and National Grid ESO are [building an AI-powered tracking system](#) to match cloud movements with solar panel locations.

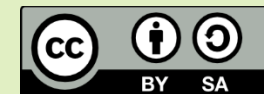
Finance

Open Banking is a data portability initiative designed to increase innovation in the UK's banking sector, by allowing individual and SME customers to securely share their data with third parties, allowing a broad range of businesses to compete in the provision of better financial services. The scheme covers current accounts and credit accounts as well as some savings accounts.

Following the Competition and Market Authority's Retail Banking Order, which mandated participation of the nine largest UK payment services providers, the Open Banking scheme has continued to gain momentum and benefit more users. The Open Banking ecosystem now has more than 3 million people and businesses using Open Banking-enabled apps and services, with over 300 firms active in the market, including:

- HMRC, which became the first tax authority in the world to integrate Open Banking into its systems, reducing the risk of payments being misallocated to the wrong account.
- Nationwide, which launched the 'Open Banking for Good' initiative to help the UK's 'financially squeezed', working with fintechs and charities, and investing £3m to create useful tools and services, such as interactive debt and money management apps.
- United Utilities, which is already using open banking to verify customer income in real-time, which improves the accuracy and efficiency of customer affordability assessments, reducing time and effort for determining customer eligibility for reduced-rate social tariffs.

McKinsey predicts that economies that embrace data sharing for finance could see [GDP gains of between 1% and 5% by 2030](#), with benefits flowing to consumers and financial institutions.



Publishing open data with a purpose

We strive to leverage the opportunities of open data to deliver value to our customers, partners and the environment. Initially we will start publishing through external platforms but will develop our own data portal, utilising a secure, well-governed release process, with open data seen as business-as-usual.

In progress

Collaborating across the sector

As a core member of the Stream initiative, we have the opportunity to not only publish open data, but to contribute directly to high-value use cases, and more easily collaborate with academia, local authorities, and other businesses.

In progress

Implementing an approach to data ethics

An understanding of data ethics is critical to ensuring a trustworthy data ecosystem. We will begin developing a standard approach to assessing the ethics of not just publishing, but also collecting and using data.

Planned

Building and improving our data literacy and culture

Starting with an open data masterclass, we will ensure there is a standard curriculum of open data resources and available training to United Utilities staff that will be working with data in any way. We want data, and especially open data, to be baked into the culture of the organisation.

In progress



United Utilities Group PLC

Haweswater House
Lingley Mere Business Park
Lingley Green Avenue
Great Sankey
Warrington
WA5 3LP

unitedutilities.com/corporate



Water for the North West