October 2021

H_2Open – Open data in the water industry: a case for change

Ofwat

About this document

This paper outlines our initial views on how open data can enable water companies to create value for water customers, communities, and the environment.

We intend this paper to supercharge a public discussion about the benefits of open data and how it could be used to help address some of the challenges the water sector faces from climate change, the environment, changing customer demands and protecting the most vulnerable. In this paper we also:

- reflect what we have heard are the challenges the sector will need to overcome;
- shine a light on some of the open data and data sharing activities already underway in the industry; and
- set high-level expectations for areas where we would like to see more progress, sector wide, by autumn 2022.

We have engaged extensively with water companies and other stakeholders to understand where they see the opportunities and challenges of open data in the water industry. We have conducted open conversations with regulatory directors, specialist data teams and operational teams to best understand how water companies are approaching open data currently[1]. We have welcomed the insights and contribution of case studies by companies.

We have also engaged with industry and consumer bodies, including Water UK, British Water and the Consumer Council for Water (CCW), as well as other industry regulators to inform our understanding of the different elements that make up a successful industry-wide approach to open data.

The Open Data Institute work with businesses and governments to build an open, trustworthy data ecosystem. We have tapped into their support and expert advice. Their insights have informed our understanding of the processes, decisions, risk appetite and culture that underpin open data.

We know that there are a wide range of stakeholders with an interest in open data and achieving change will need the passion, energy, support and skills of many people. If you would like to be involved, details can be found at the end of this document.

1. For a full list of stakeholders engaged with, please see the appendix.

Executive summary

In our strategy, we said open data means making data freely available to everyone to access, use and share. What could open data look like in the water industry? What outcomes could it help to deliver?

Time to act, together: Ofwat's strategy sets out the strategic goals to transform water companies' performance, to drive companies to meet long term challenges and to provide greater public value for the benefits of water customers, communities and the environment. We set out that it was time for the water sector to embrace new thinking and innovation and that open data could play a central role. We shared that we intended to speed up the pace of change and build on what is in place to explore an open data approach within the sector.

In this paper we set out the case for open data in the water industry and how open data can enable water companies to create value for water customers, communities and the environment.

We examine the benefits of open data and how it could be used to help address some of the challenges the water sector faces from climate change, the environment, changing customer demands and protecting the most vulnerable.

In addition, we set out what we have learnt about the current position of the industry and reflect what we have heard from companies and stakeholders about how to unlock the best outcomes from open data. In doing so, we also highlight some of the open data activities already being undertaken by the industry.

We primarily focus on the 17 incumbent regional monopoly companies, however we consider that many of the benefits of this approach to open data could – and should – apply to data held by all operators including business retailers and new appointees and variations. We would therefore like to see those businesses also explore the opportunities and benefits that a more open approach to market data could present to business customers and the environment.

Huge untapped opportunity – case for change

The water sector is host to a wealth of data and information about not only water and wastewater services, but also how those services interact with communities, the environment and how we live. Across the water sector we have seen some good practice, but very few companies have introduced open access to their data sets. There is currently a huge untapped opportunities through open data for the industry including:

- stimulate more **innovation** and collaboration;
- encourage new business models and service offerings that increase **efficiencies** and enhance the **customer experience**;
- provide insights on company performance and assets, improving decision making;
- enable companies to work collaboratively to use open data to tackle shared challenges;
- improve **transparency** for customers, and the owners of companies, and build trust in the water companies.

It is also an important part of how we see public value in the industry is delivered to consumers, wider stakeholders and society.

Keeping pace

The growing interconnectedness of people, organisations and technology is changing customer and societal expectations around how services are provided. The water sector needs to maximise the benefits this opportunity presents and keep pace with the wider digital economy.

Open data can help to do this.

Through engagement with stakeholders, we have identified key enablers for unlocking the best outcomes from open data:

- a strong data culture and the development of capability and skills;
- improved collaboration on open data across the industry; and
- established data infrastructure.

We discuss these enablers in more detail in Section 2 of the document and highlight what we have heard about the challenges in delivering open data. These are not unique to the challenges that other sectors are encountering and further challenges may arise as open data develops. They are real and will need to be worked through. We are confident that the industry is well placed to address these.

We can also learn from companies and public bodies in sectors such as banking, food, and transport, opening and sharing data. We have provided some case studies in this document about the emergence of new products and services as a result. More recently in energy, we have seen the introduction of regulatory requirements for network operators to open data to support decarbonisation and stimulate behavioural changes in the industry.

Expectations for delivery of open data in the water sector

We are confident that the industry is well placed to deliver open data and has the capability, expertise, and incentives to do so. However, the water industry is at the start of this journey and companies will need to make this more of a priority and work together to unlock the best outcomes from open data. We welcome the progress that has been made already and set out some of the good practice and learning that can be built on.

The next 12 months will be important for establishing some of the key foundations for open data, including progressing the industry wide collaboration initiatives we have seen emerging. In addition, we have stated in PR24 and beyond that we expect companies to be making better use of data, including by embracing open data. This insight should inform their PR24 business plans.

We have set out our expectations in this document around the key enablers set out above. We will review the progress the industry is making in autumn 2022.

We are excited and optimistic that the industry can deliver open data. We look forward to seeing the progress the industry makes in ensuring that water data is an asset that everyone can access, use and share.

Contents

www.ofwat.gov.uk

For more information about Ofwat and our work please contact:
mailbox@ofwat.gov.uk Phone: 0121 644 7500

See the water sector dashboard at discoverwater.co.uk
Find out about the water market for businesses at open-water.org.uk

Improving life through water | Gwella bywyd drwy ddŵr

1. Case for change: open data in the water industry

The water sector is host to a wealth of data and information. To overcome future challenges and achieve the long term aims of the sector, extracting insight from this data is critical. The benefits of open data also have the potential to reach beyond the water sector and supply chain, across multiple sectors, and to research and development.

In this section we explain what we mean by open data and set out our understanding of the current position of the water industry. We also explore how open data can be used to make a difference in the industry. By creating an open and trustworthy data ecosystem in the water sector open data can:

Innovation

Stimulate **innovation** creating opportunities to draw in and leverage new skills, technologies and stakeholders.

Support the development of **new business models and services.**

Enhance the **customer experience** through the development of new products and services such as apps.

Efficiency

Provide **information and insights** for companies on performance of assets by attracting new perspectives and encouraging fresh thinking.

Improve data **quality**, collection, and maintenance of data.

Enable companies to work collaboratively with other companies in the industry and beyond water to maximise the benefits of data to address shared challenges. **Collaboration** can also facilitate the linking and merging of data to create larger datasets, enabling deeper analysis and insight through research or machine learning.

Improve **decision making**. Data from multiple sources, along with experience and understanding, can allow companies to make better business decisions, benchmark themselves against others and work together. Open data can also provide better evidence and insights for policy makers and governments.

Transparency

Improve **transparency** for customers and the owners of companies and build trust in the water companies.

We cannot know how data might be used by others now or in the future. The value of some data may become immediately apparent while other data sets that might not be used much to start with may (or may not) become of more interest later as new collaborations, technologies and science develop. Making data freely available to everyone to access, use and share would allow this experimentation to happen.

“Open data means making data freely available to everyone to access, use and share”

– Ofwat strategy

What is open data?

Just like sewers, water pipes and treatment works, data (and the people, processes and technology which support it) is an asset and part of the critical water industry infrastructure. Data is essential for developing insight, making informed decisions and improving services.

Water companies serve 53.5 million people in England and Wales[2], with responsibilities for delivering safe and clean water and wastewater services for customers. The majority of household customers are served by the 17 incumbent regional monopoly companies.

In addition to these incumbent companies, there are eight new appointees and variations (NAVs)[3]. There are also 20 suppliers[4] within the business retail market, which serves 1.2 million businesses and operates under competitive conditions, allowing businesses to change retailer.

Together, these companies hold large amounts of data assets about the services they provide and their impact including (and not limited to):

- operational data;
- asset data;
- customer data (including personal data); and
- financial data.

While this document is primarily focused on the 17 incumbent regional monopoly companies, we know that there is also a significant amount of market data in the non-household retail market. This includes data in the central market operating system (CMOS), which is maintained by the market operator (MOSL) and which is accessible to licenced Trading Parties (incumbent companies and business retailers) but not more widely. We consider that many of the benefits and this approach to open data could – and should – apply to such market data (or at least subsets of it). We would therefore like to see market participants work collaboratively to explore the opportunities and benefits that a more open approach to market data could present to business customers and the environment.

An open data approach presumes that water data should be available and accessible publicly, and where data is closed or only shared between certain parties, requires justification as to why it should not or cannot be opened.

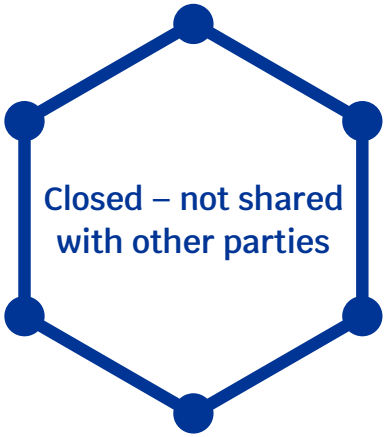
Opening data in this way is an important part of how we see public value in the industry being delivered to consumers, wider stakeholders and society. We set out in our Public value principles that we believe that open data can facilitate engagement and build trust with customers, the supply chain and research communities (principle three).

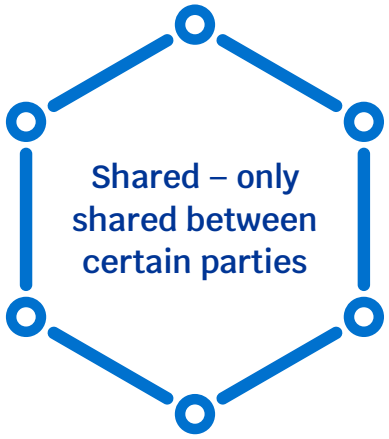
2. Environment Agency. (2021). Water and sewerage companies in England: environmental performance report for 2020.
3. Albion Water Ltd, Albion Eco Ltd, County Water Ltd, Icosa Water Services Ltd, Independent Water Networks Ltd, Leep Networks (Water) Ltd, Severn Trent Services (Water and Sewerage) Ltd, Veolia Water Projects Ltd.
4. Current retailers in the market.

Circumstances where data from the water sector is being opened, shared and kept private

Personal, private customer data, for example household usage

Data protected by GDPR[5], for example names and addresses of customers





Priority Services Register data (data sharing of vulnerable customer address and contact data)

Asset information, for example water mains in England span over 347,000 km, and sharing the location of these is being piloted through the National Underground Asset Register (NUAR)

Open – accessible and available to all

Environment Agency flood risk mapping

Leakage data, for example Yorkshire Water opened a leakage data set through Data Mill North

5. General Data Protection Regulation

Case study

Opening data sets

Data Mill North is an ongoing collaboration between Leeds City Council and other stakeholders, spanning both data providers and data users. **Yorkshire Water** has published customer meter datasets, drinking water quality data and leakage data through Data Mill North, generating engagement from data scientists and third parties.

Ideas generated using this leakage data spanned predictive analytics, new technologies for leakage detection and designing leakage identification techniques that recognise night-time water usage for those working night shifts. Other data publishers include government departments, city and borough councils, community organisations and start-ups.

Case study appendix

Like in other industries, companies that generate data do not always come up with all the best uses of the data. **To ensure that the most value is gained from industry data, it should be made as open as possible.**

Getting the most out of open data means adopting a mindset that is open by default, recognising and managing risks to maximise opportunities. However, we recognise that there are circumstances when data should remain private or can only be shared between certain parties. For example:

- how personal customer data is used; and
- where the use of certain data may raise issues for the security of our critical national infrastructure.

Case study

National Underground Asset Register

The National Underground Assets Register (NUAR) is a project run by the Geospatial Commission. It recognises the value of location data in improving the ways in which national infrastructure is planned, built and managed through better mapping underground assets.

The NUAR is a combined map of existing underground asset data, indicating where electricity and phone cables, and gas and water pipes are buried. Bringing this information together in one accessible place aims to prevent both accidents and disruption to the economy (currently estimated at £2.4 billion a year). At present, NUAR has only engaged with regional water companies in the North East and the London area (Northumbrian Water, Thames Water and SES Water) in its pilot phase.

Future plans include expansion of the NUAR across England, and the Geospatial Commission are currently working to build a minimum viable product for the North East, Wales and London before roll out to the remaining regions. Once operational, NUAR is expected to deliver £350 million a year in benefits through avoiding accidental asset strikes.

Case study appendix

Current position of the industry

The use of open data will grow and mature, and we know this will take time. The water industry has started this journey, with some companies more advanced than others. We believe – and have heard through our engagement – that there is much more that can be delivered through open data.

This section reflects our understanding of the current position of the water industry.

At present, very few water companies provide open access to their data sets, although we have seen a growing number of individual open data initiatives in the form of hackathons, pilots and sharing data. We have also seen companies participating in specific successful collaborations such as the National Underground Asset Register (NUAR) which is encouraging.

What is a hackathon?

A hackathon is an event designed to bring together digital experts, coders, graphic designers, software engineers and subject-matter experts to solve a series of challenges. They may be framed around a particular data set, theme or challenge.

Companies do publish some specific information; for example, Ofwat requires companies to make publicly available their Annual Performance Reports (APRs) and five-yearly business plans. While this is not currently in the form of open data[6], they provide valuable information about how companies are performing and how they plan to deliver their services in the future. Companies also provide data and information to Defra and the Welsh Government, the Environment Agency, Natural Resources Wales, Drinking Water Inspectorate and CCWater. At present however, **no company has opened large numbers of data sets beyond what is required by Government or a regulator for regulatory purposes.**

6. This is often hosted by companies in different formats and with differing degrees of accessibility to non-specialist audiences, with further consideration required to make these data sets into accessible open data capable of generating efficiencies, innovation and improving transparency and trust.

Most companies told us they used open data from other sources to improve the quality, completeness, and utility of their own data. For example, many companies draw on Met Office data sets to strengthen internal forecasting of weather and climate patterns. Most companies also recognised the potential benefits of combining open data across multiple industries.

> **In the future this will allow us to plan better for other abnormal events including weather patterns and to be able to monitor the impact much more closely of our work to reduce PCC through our other campaigns"**
>
> – Andrew Morris, CIO Affinity Water

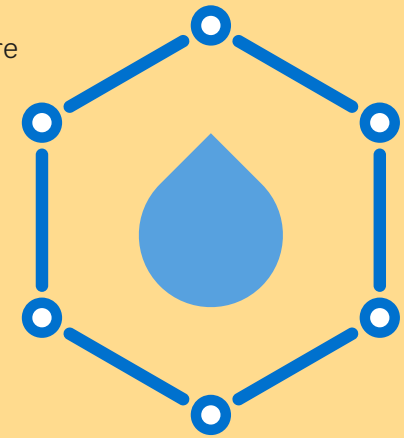
Case study

Using open or shared data

Through conducting a hackathon with external partners, **United Utilities** utilised open data sets from Public Health England on the location of food service establishments alongside a range of its own data. This data hack found a strong correlation between food service establishments and an increased risk of wastewater flooding, with potential to inform United Utilities' approach to targeted campaigns on how certain activities may increase the risks of wastewater flooding, creating tangible benefits for customers.

Yorkshire Water taps into shared network data across Sheffield to harness collective benefits. Established by British Telecomms (BT), the narrow-band Internet of Things is a pilot collaboration between BT, Yorkshire Water and Stantec to harness existing technology for asset maintenance and operational decision making across all three networks.

Anglian Water has utilised mobile phone data from BT to predict water consumption hotspots. Individual mobile phone usage is transformed into population equivalents across Anglian Water's district monitored areas (DMAs), creating a new dataset that can be used to generate further insights about water consumption, such as how per capita consumption might have been impacted by more people living at home during Covid-19.



Affinity Water is pioneering new AI driven data sharing to allow companies to better monitor the impacts of the weather and Covid-19 on future water demand. Using meteorological data within their AI systems, Affinity Water has been able to accurately predict per capita consumption (PCC) use across its network.

Case study appendix

Case studies indicate that companies are identifying significant benefits from individual activities and open data initiatives, including improvements in operational efficiency and environmental quality, innovative solutions to a range of challenges (climate change, affordability) and improvements in transparency and relationships with customers.

We also heard from companies that they liaise across a network of Chief Information and Chief Technology Officers (CIOs and CTOs) and share insights with one another. Active collaboration with the wider water sector and supply chain can be seen in company participation in the Water Data Taskforce and engagement with groups such as BIM4Water.

Through our engagement, we have observed that companies are at different stages in terms of their use of data, data management and considerations of open data. We have seen that differences in culture around data, capability, and levels of ambition really impact upon the maturity of a company in terms of open data which we explore further in section 2 below.

We would like to see companies maximising their use of open data and keeping pace with the associated digital transformation happening across all sectors.

How open data can be used to make a difference

As outlined in our strategy, the water sector faces multiple interconnected challenges, spanning climate change, population growth and affordability. In addition, customers' demands are changing, with greater emphasis placed on the environment, customer experience and building a culture of transparency and trust.

Addressing these challenges – and associated opportunities – will require the sector to adopt a long term perspective and utilise new insights and technologies. In this section we have considered how open data can help with this journey in the following areas:

- adapting to climate change;
- protecting the environment; and
- meeting customer's expectations and diverse needs.

Adapting to climate change

Climate change threatens resilience, both of networks and water supply. We are already seeing drier summers, more frequent and intense rainfall, more variable river flows and biological changes in water bodies.

We know that effective climate policies and approaches are driven by access to high-quality data and the water industry holds important and relevant data. To drive a co-ordinated national and global approach to reaching net zero, the water industry will need to collect and make its data available, as well as using other data to understand its impact.

Open data could enhance this co-ordination, enable innovation, support investment decisions and track progress. It will also support effective reporting and forecasting and build the public's trust and awareness of what the industry is doing.

Case study

Energy

Within the energy sector, Icebreaker One's Open Energy system aims to improve and modernise access to energy data through creating open data standards and governance. Easier access to energy supply and demand data forms part of the energy sector's decarbonisation plan; for example, renewable energy is highly dependent on weather data and this variability can make it harder to model renewable power generation and customer demand – better data sharing can help to better address the challenges associated with wide-scale roll out of renewables.

> **Data is the single biggest enabler of a decarbonised, decentralised and digitised energy future. It's the tool that will bridge the gap between where we are now vs where we need to be to achieve net zero"**
>
> – Matt Hastings, Innovate UK, Modernising Energy Data Access

Case study appendix

Protecting the environment

Our strategy reminds us that the natural world is fundamental to the water sector, but it is facing profound challenges in the form of climate emergency, wildlife loss, and environmental degradation.

If we are to build a sustainable, resilient future, we must step up to address these challenges. Companies will need to consider the environment as an integral part of their business and the value of the natural environment is a key aspect of PR24, with an increasing shift towards more nature-based solutions. We expect that open data may form part of companies' long-term strategies for asset management through increasing understanding, operational efficiencies and catalysing innovation. This is reflected in our recent Asset Management Maturity Assessment – Insights and Recommendations publication.

Making data open and available, and using others' data could support this in a number of ways. We have heard through our industry engagement that open data could enable:

- a more in-depth understanding of operational environmental challenges such as real time monitoring of combined sewer overflows;
- better predictive analytics and forecasting for resource shortages with changing climates;
- better analysis and monitoring of network solutions such as smart metering data and performance of nature based solutions; and
- improved understanding of pollution triggers and thresholds.

Open data could also aid in the development of long-term projects and solutions that protect the environment, providing a cross-sectoral data set to draw upon to deliver predictive forecasting and better understand the state of the environment.

Case study

Environment Agency

The Environment Agency collects, uses and shares data relating to waste, fisheries and other environmental areas. The EA recently undertook a programme of data releases, opening data sets across a range of different environmental concerns.

These data sets are published online and can be accessed and used by third-party organisations. For example:

- the British Red Cross combine EA flood data with data from the Met Office to provide advanced emergency weather warnings via their app;
- software start-up Dsposal uses data from the EA's public register to provide waste disposal companies with a real-time picture of the compliance status of their supply chain; and
- the Pang Valley Flood Forum utilises EA datasets to display live dashboards and datasets for local communities.

The benefits include:

- innovative downstream use of data that does not require facilitation or management from the publisher;
- societal benefits delivered as other organisations build products and services using EA data; and
- easier cross-governmental use of data, for example the National Audit Office (NAO)'s flood risk data visualisation tool.

Case study appendix

Case study

Forecasting enabling pollution reduction

Northumbrian Water has developed analytics to forecast flow at sewage pumping stations, using existing datasets spanning sewage pumping data and historic asset failures. Following a hackathon in 2016 where historical asset failure, flowmeter and level monitor datasets were shared, predictive analytics were developed to provide lead indicators for potential future asset failures. Drawing on the insights generated from these analytics, Northumbrian Water were able to better forecast and prevent sewage related asset failure. They saw a 70% reduction in pollution originating from sewage pumping station assets from 2016–2017 and following this initial reduction this level has been sustained beyond 2017.

Wessex Water is developing intelligent sewer monitoring algorithms to identify issues and allow for proactive intervention before sewage spills from storm overflows take place. Following an open data trial, intelligent sewer depth monitoring has detected a number of partial blockages before they become performance issues and reduced unnecessary alarms during heavy rainfall by 97%, increasing operational efficiencies.

Case study appendix

Customers' expectations and diverse needs

Customer service

Customers' interests are evolving, and their expectations are growing in terms of customer service.

> Ofwat's Out In the Cold review of water companies' responses to the 'Beast in the East' weather event of 2018 indicates the need for access to accurate customer data, and places emphasis on this for the failure of communications, which were 'often not targeted, timely, clear or helpful' (p7).

Sharing customer data in a controlled way can have direct benefits for individual customers through additional products and services, building relationships with third parties, and ways of engaging with customers. An example of using data to engage with customers is Severn Trent and Thames Water's web apps that allow customers to check for known or frequently reported problems in their areas when reporting leaks or other issues.

The industry would not be starting from scratch or be the first to embark on this journey with customer data. We have seen the emergence of new products and services through customer data sharing in open banking, which allows third parties secure access to customers' banking and other financial data. For example, NatWest's 'Housemate' app in partnership with Experian[7], allows housemates to split bills directly and to share rent payments to build individual credit scores, giving customers financial autonomy.

Vulnerable customers

We must also make sure customers' diverse needs, and particularly those in vulnerable circumstances, are properly met. Using data sharing to understand more about customers' circumstances can help ensure support can be provided where appropriate.

7. Open Banking Expo 2021 NatWest launches 'Housemate' app (accessed 21/09/2021).

Our Listen Care Share campaign explores how water companies can best support customers struggling to pay, respond to their information needs, and go the extra mile for vulnerable customers and those experiencing mental health issues. Through this campaign, we identified access to data as a key way to improve outcomes for customers by reviewing data that is collected, water companies can use it to understand more about those who are receiving or missing support. We expect companies to improve their communications with customers through making better use of data to deepen understanding of customer needs. We see data sharing as integral to identifying those in need of support.

We have seen tentative steps in customer data sharing in water to ensure that support is provided. The Water Priority Services register collates personal data and information about customers in vulnerable circumstances, which is then used to provide additional support and may be shared with trusted partners and third parties to facilitate this support.

Case study

Safeguarding the vulnerable

Priority Services Registers (PSRs) allow customers in vulnerable circumstances to record personal data with their water company, ensuring that they are offered additional support. PSR customers may receive a range of extra support, such as help with meter readings, advance warning of supply chain interruptions, accessible communications and a dedicated customer advice hotline. Data may be shared with partners or third parties, such as the British Red Cross, providing additional assistance for vulnerable customers in incident situations. Data sharing between the PSRs of water and energy companies has undergone a successful pilot scheme[8], and there are plans for further integration.

Similarly, the opt-in WaterSure scheme provides financial assistance to metered customers who have both high essential water usage needs and are receiving support through benefit or tax credit systems. Data sharing from the Department for Work and Pensions (DWP) on those receiving benefits enables the successful operation of WaterSure across the water companies.

Case study appendix

8. Data sharing between United Utilities Water and Electricity North West in a pilot scheme.

Case study

Affordability

United Utilities utilises open banking technologies to verify customer income in real-time, improving the accuracy and efficiency of customer affordability assessments. Conducting real time assessments has allowed United Utilities to streamline its process for determining customer eligibility for reduced-rate social tariffs, reducing time and effort costs.

Beyond water, the emergence of open banking has helped to foster a culture of innovation around providing support to vulnerable customers. **Nationwide's Open Banking for Good** challenge brings together businesses, civil society and government departments to create and scale solutions to improve financial capability in the UK through creating tools and services that might enable customers to better manage their money.

Case study appendix

Affordability

Affordability concerns have been amplified by the impact of Covid-19. One third of households in England and Wales already sometimes struggle to pay their household bills[9], and this figure may rise as the long-term impact of the pandemic becomes clearer. As we have described in PR24 and beyond, there is also unlikely to be the same scope for bill reductions from falling underlying financing costs that we saw in PR14 and PR19.

In other industries, sharing customer data in certain circumstances can provide more personalised customer experiences whereby having a better understanding of customer's financial health, companies may provide flexible or personalised billing or the use of data portability across different sectors (as highlighted in the example of NatWest's 'Housemate' app above).

Transparency and trust

Open data can also create a more transparent and trusting relationship with customers. Companies that embrace open data and provide information and insight into their operations can in turn gain the reputational benefits from doing so even where open data has the potential to expose issues.

The greater transparency of open data has the potential to reveal a much richer picture of company performance (both good and bad), to inform companies themselves as well as their stakeholders, policy makers and regulators. It therefore provides significant opportunities for companies and regulators to better identify and drive improvements and impact and to reduce potential detriment. Digitalisation of the water sector means that data collection and sharing must also be in line with existing competition law obligations.

9. Online nationally representative survey of 2,100 bill payers in England and Wales. Conducted for Ofwat by Panelbase (fieldwork 26 March to 1 April 2021). See Listen, Care, Share: Water customers' experiences during Covid-19, p9.

Case study

Surfers Against Sewage

Surfers Against Sewage developed and run the Safer Seas and Rivers Service app, which uses open data from the Environment Agency and eight major water companies to alert water users when bathing waters are safe or unsafe for human use. The app covers nearly 400 locations across the UK, has 51,000 registered users and has sent 1.3 million bathing water quality alerts since 2019.

The benefits include:

- forecasting data and alerts can support users to understand when it is safe to use UK bathing waters; and
- the increased availability of EA datasets has led to creation of user-friendly products, allowing the public and water users to access the conversation. Data releases alone might have proved too technical for many users of the app.

Case study appendix

Open data can be used by a diverse range of audiences, as well as to develop apps and services for customers which provide information and insight to them on their water and sewerage services. This provides a way for customers to actively engage with data and information on areas that matter to them, such as the impact upon the environment.

Open data can also assist companies in engaging with a broad range of stakeholders and open up avenues of partnership or cooperation. Through increased transparency and building trust, open data has the potential to reframe traditional company relationships, encouraging a more collaborative future approach.

2. Unlocking the best outcomes from open data

There are many benefits to open data. However, achieving open data in a considered and sustainable manner requires more than just data releases.

Through engagement with stakeholders, we have identified key enablers for unlocking the best outcomes from open data:

- a strong data culture and the development of **capability and skills**,
- improved **collaboration** on open data across the industry; and
- established data **infrastructure**.

In this section we provide more detail on these key enablers. We have also provided a series of case studies throughout this report and in the appendix, which highlight some of the open data activities already underway within the industry.

In our engagement, it was clear that delivering open data will require significant effort and there are huge opportunities but there will also be challenges to overcome. Further challenges are also likely to be identified as open data develops. They are not unique to experiences in other sectors, but they are real and will need to be worked through. We are confident that the industry is well placed to address these.

We have reflected in this section what we have heard so far about these opportunities and challenges around the three enablers set out above. We observed that the weight placed upon the challenges varied significantly between companies.

In seeking out opportunities and addressing these challenges, the industry must ensure that it avoids **limiting or preventing the use of data inappropriately.** There is great potential for companies to work together to share insights and ideas on how to overcome challenges and those that may be encountered in the future.

Strong data culture and the development of capability and skills

A strong individual company and industry-wide vision and strategy is essential for delivering open data. Strong senior leadership ensures strategic focus, and that data and analytics capability is developed.

Several companies showed a strong commitment to open data and had undertaken open data initiatives. We observed that these companies also had **data strategies and plans in place for maximising the use of data and analytics within their businesses, including open data.** These companies were also building and nurturing data and analytics capability within their organisations. Strong senior leadership support for adopting open data and building data capability was also noted.

Several companies had undergone a significant internal data transformation and told us that the establishment of data systems, management and governance were delivering new insight and driving performance improvements through implementing and using open data.

Companies that are less advanced in terms of data capability and skills reflected this in their position on open data. Those companies were at earlier stages in their thinking around open data and typically have not participated in an active fashion in cross-industry initiatives. A number of companies had strong advocates for open data, but this ambition was not reflected in organisational strategies and plans.

Costs

As well as costs in building skills and capability, there may be costs of data cleaning, processing, storage and hosting to consider. Furthermore, once data sets have been opened, there is a continuing cost in updating these. A number of companies considered that these costs are likely to be easily outweighed by the operational efficiencies gained.

Smaller companies[10] in particular cited relative costs and limited specialised data skills as potential inhibitors to opening data. However, a few smaller companies also recognised that they might be some of the biggest beneficiaries of open data, drawing upon larger cross-industry data sets to develop stronger insights and build internal capabilities. In particular, one company pointed to the development of Artificial Intelligence (AI) algorithms that increase their precision with the size of datasets that they can access.

One company suggested that opening data sets could reduce its ability to tap into the commercial value of its data reducing the potential savings it would be able to pass onto customers (for example if data could be used to negotiate better contracts, then those economic efficiencies can be delivered to customers). The company also recognised that the loss of this benefit might be outweighed by wider benefits (transparency, efficiency, innovation) of open data.

Reputational implications

Very few companies directly cited the potential reputational risks of opening data, although a number raised how opening data sets might result in increased criticism of companies and misinterpretation of their practices. However, other companies actively welcomed increased transparency as a positive reputational benefit.

10. From the 17 main incumbent household companies.

Case study

Automatic sewer defect detection

Dŵr Cymru is harnessing artificial intelligence and image detection techniques to automatically detect sewer defects across their networks.

Dŵr Cymru was already collecting hours of CCTV footage from within its sewer pipe network and manually reviewing this footage to search for defects, such as cracks or fractures in pipe walls. It used artificial intelligence and image detection techniques to train a model that would be able to automatically review video footage and alert upon detection of a potential defect. The first model was able to accurately detect nearly 80% of defects correctly, freeing up time previously spent by operators manually reviewing video footage and allowing them to use their skills more rewardingly to maintain the network for customers. Dŵr Cymru is now working on refining this solution to increase accuracy ahead of roll out across its operating area.

Open datasets from other water companies could strengthen the reliability and accuracy of AI models such as this. Differences in geographical areas and pipe specifications would need to be noted and accounted for in this specific example.

Case study appendix

Improved collaboration on open data across the industry

All companies recognised that to deliver the best outcomes from open data, collaboration across the industry (and more widely) is essential. In many instances, companies recognised that they did not individually have all the data necessary to address a particular industry-wide challenge. Equally, many companies recognised that the value to others of open data may be in accessing industry wide data rather than individual company data.

Working together to collect, share and use data can help to address challenges, improve decision making and adapt to changing environments. Working collaboratively can also facilitate the linking and merging of datasets to create larger datasets, which would enable deeper analysis through research or machine learning.

Commitment

Several companies also mentioned the level of commitment required from all parties and the need for contribution to be fair and equal. Reticent companies withholding data sharing but potentially benefitting from others doing so was raised as a concern.

Companies also pointed to open data collaborations cross-sector including across utilities, regional relationships as well as across the supply chain, and with public bodies. Regional relationships or platforms for collaboration, such as Data Mill North, allow for the geographic nature of company data be utilised, tapping into existing location data and location specific insights. The cross-sector collaboration we have observed is often tethered to regions, locales or challenges, as explored in the case studies below.

Case study

Cross-sector collaboration

Bristol Water worked responsively with the NHS to provide real time monitoring of water supply to the Covid-19 Nightingale Hospital in Bristol. Chemical and physical monitoring of the site ensured that the hospital was able to operate securely, and establishment of an early warning system meant that Bristol Water were prepared in case of supply interruptions.

Data Mill North is an ongoing collaboration between Leeds City Council and other stakeholders, spanning both data providers and data users. **Yorkshire Water** has published customer meter datasets, drinking water quality data and leakage data through Data Mill North, generating engagement from data scientists and third parties. Other data publishers include government departments, city and borough councils, community organisations and start-ups.

Case study appendix

We heard that identifying the use cases for open data through engaging with external users can lead to the emergence of previously unconsidered relationships. Engaging on needs and requirements beyond the industry can trigger new innovations and initiatives and these conversations will be important in unlocking the full potential of open data. In considering collaborative opportunities, companies should engage with potential users beyond their existing relationships, as well as increasing transparency and publicising their open data releases.

Established data infrastructure

All companies raised points around the importance of data infrastructure. We have defined data infrastructure to include datasets, the technology, training and processes that collect, maintain and use the data to gain insights and solve problems. We have set out below what we heard regarding:

- data quality and completeness;
- common standards, metrics and governance;
- data platform; and
- prioritisation.

It will be important that data infrastructure is designed to be as open and accessible as possible, while protecting data which should be private and ensuring security issues are addressed. Data infrastructure components should also be designed and supported to benefit all potential users. Engagement with stakeholders during development will be critical.

Data quality and completeness

Ensuring good quality and complete data is essential to deriving strong insights from open data. We heard from companies that data used for regulatory purposes is generally in a good format (and quality assured), but that there remains variability in the quality of data sets across companies more widely. Many have only recently digitalised their reporting processes or have not yet done so.

Case study

Data standards

FIWARE4Water is an offshoot of the European Commission funded FIWARE. As part of this initiative, **South West Water** has been working to establish consistent standards and analytics for smart meter data, allowing for comparability across utilities and datasets. South West Water focus on deriving value from smart meters for customers in a way that links back to the FIWARE platform, therefore allowing for interoperability and data exchange across other smart meter datasets and datasets from other sectors.

Case study appendix

However, as we have stated in PR24 and beyond, companies will need to resolve known deficiencies in their data, and this will be important for opening data. We can also learn lessons from the opening of the business retail market and the quality of the data companies hold.

A number of companies also pointed to opening data as a way to resolve known issues in data, with one suggesting that external data scientists can identify deficiencies in data and help to resolve them. Other companies suggested that while data quality is an ongoing area for improvement, there is in fact value in opening data sets even if concerns remain about the quality of the data, provided appropriate contextual explanations are included. This could enable external groups to derive insights or could simply be in response to recognition that some datasets may be useful even if not of the best quality or completeness (in particular historical datasets).

Common standards, metrics and governance

Beyond ensuring quality data, comparability across water data sets is also key to deriving value from open data. Establishing common standards and metrics could allow for easier cross-sectoral comparability, while shared approaches to data governance could allow for easier interoperability and the protection of private data.

Most companies supported a shift towards common standards, definitions and reporting, however leaders in open data emphasised that support for commonality should not present a barrier to initial data sharing and opening. Technological solutions could provide a work-around for integrating early data releases and extracting insight from disparate data sets.

Wider points raised were around ensuring that data standards are unbiased and fit for purpose and that they work for the end user. Two companies felt that there was a regulatory role for Ofwat in this area.

Some consideration should be given to the frequency of data releases and the need to factor in data assurance. We also observed one company engaging across the sector and internationally to establish common standards, indicating that wider impact should be considered.

Data platform

Most companies supported the concept of a platform as it could ensure a co-ordinated approach in areas such as access rules, standard setting, as well as data governance and ethics, however early investment could result in path dependency or risk leaving some companies or data streams behind.

One company is leading a collaborative industry-wide open data initiative alongside 11 of the 17 water companies and other specialist collaborators, which we welcome as significant progress in developing a shared strategy and approach (see case study below). This initiative has highlighted that it will be necessary to experiment and look at new technologies and ideas to reach an optimum solution and we support this view.

Prioritisation

Companies expressed that the pace of industry-led standards and platform should not hinder their ability to experiment, open and access data sets and we welcome this. We heard that encouraging an environment where it is safe to experiment and learn from those experiences even where they are not successful is important.

In addition, almost all companies expressed a preference for opening data sets incrementally, although different data sets were raised as potential start points. Prioritisation of data sets is something that we would expect to see strong industry-wide collaboration on, focusing on outcomes and building on engagement with wider stakeholders.

Prioritisation of opening data sets should also consider the overarching public benefit, in line with Ofwat's public value principles. Companies should consider where and how they can collaborate with others to optimise solutions and maximise benefits (principle five), seeking to align stakeholder interests and enable water companies to deliver the best value for, customers, communities and the environment.

Case study

Northumbrian Water and 'Stream'

Northumbrian Water set out its ambitions for unlocking more value from data in its Annual Performance Report for year ended 31 March 2021, including its development of 'Stream'.

Stream is a collaborative industry-wide open data initiative. 11 of the 17 main incumbent water companies have joined forces to take forward open data across the water sector, with an initial strategy exercise identifying 'critical and high value opportunities for the sector' and building an 'opportunity roadmap'.

Through working together to create a shared strategy and approach to open data, companies are 'identifying answers to the challenges of what information is shared, how it is shared, and crucially how it is kept safe'[11]. Companies have tested three initial proofs of concept aligned with strategic goals within the sector, spanning data housing infrastructure, novel algorithms to flag algal blooms and predictive asset alert systems.

Collaboration to solve sector-wide challenges and explore the solutions and opportunities that open data might provide indicates that the companies involved in Stream are increasing their open data maturity and are considering how the benefits offered by open data can best be unlocked.

Other company participants are Anglian Water Group, Thames Water, Severn Trent, SES Water, United Utilities Water, South West Water, Dŵr Cymru, Southern Water, Yorkshire Water. Also Scottish Water.

Case study appendix

11. Northumbrian Water Limited Annual Performance Report for year ended 31 March 2021, p40.

3. Expectations for delivery of open data in the water sector

We are confident that the industry is well placed to deliver open data and has the capability, expertise and incentives to do so. This is supported by the commitments the industry made in the UK 2050 Water Innovation strategy.

We know that it will take some time to fully implement and that some companies are more advanced than others. However, **we want to see the water industry keep pace with the wider digital economy and we want to see measurable progress in delivering open data over the next 12 months.**

In section two, we identified through our engagement with companies and stakeholders' key enablers for the successful delivery of open data. We would like to see progress against each of these areas as part of a holistic approach:

- a strong data culture and the development of capability and skills;
- improved collaboration on open data across the industry; and
- established data infrastructure.

It is for the companies to determine how they do this and to take the action necessary. However, based on what we have heard and the experiences within the industry, we consider that positive indicators of progress within the next 12 months could be (but are not limited to):

- companies developing open data strategies, including developing a clearer understanding of their data sets and a roadmap for the delivery of open data, building skills and capability, and setting long term goals. These strategies should be endorsed at CEO and/or Board level;
- evidenced collaboration between companies and stakeholders, as well as beyond the water sector;
- formalisation of this commitment to collaboration through a new or existing open data network or working group, including independent members to hold the industry to account for delivery of open data and to provide expertise in what is a fast changing and specialised area;
- continued work on an industry-wide open data initiative to identify and develop the enablers to open data, including skills, infrastructure, and data governance practices; and
- more experimenting and trialing open data, developing a more mature understanding of the benefits and the work to be done to realise these over the long term.

We would like to see each company individually and collectively consider how they will evidence their progress against these key factors and keep their stakeholders engaged and informed.

“We will ensure that we open data and share knowledge and technology to avoid duplication. We will seek to share data freely to encourage innovation. We will create change at pace, seeking to transform the sector”

– UK 2050 Water Innovation strategy

Non-household retail market

As stated above, while this document is primarily focused on the 17 incumbent regional monopoly companies, we know that there is also a significant amount of market data in the non-household retail market, including data in CMOS. We consider that many of the benefits and this approach open data could – and should – also apply to such market data (or at least subsets of it). We would therefore like to see market participants work collaboratively to explore the opportunities and benefits that a more open approach to market data could present to business customers and the environment.

What is Ofwat's role?

We want to ensure that an **open and trustworthy data ecosystem in the water sector is established.**

We want to see the value for customers, communities and the environment delivered through open data and we want to help realise those benefits.

We will do this through ongoing engagement, in our consideration of how we regulate going forward, and the expectations we set the companies.

The next 12 months will be important for establishing some of the key foundations for open data, including progressing the industry wide collaboration initiatives we have seen emerging. We will **evaluate the progress the industry is making in autumn 2022 against the key enablers identified above.**

We are **confident that the industry is well placed to deliver open data**, but should we find that little or no progress is being made, we would revisit this position and consider the formal tools we have to ensure the delivery of these goals in the future.

4. Next steps

We want to ensure that the industry harnesses the full potential of open data.

We look forward to seeing the progress the industry makes and we will evaluate this progress in autumn 2022. We will in the meantime continue to engage in the work that is now happening, what is being learnt and what benefits open data in the water industry could bring for companies, customers, communities and the environment.

Please get in touch via Twitter using the hashtag #H2Open or through our web page, where you can also find out more about what companies are doing and view additional resources.

Ofwat (The Water Services Regulation Authority)
is a non-ministerial government department.
We regulate the water sector in England and Wales.

Ofwat
Centre City Tower
7 Hill Street
Birmingham B5 4UA
Phone: 0121 644 7500

© Crown copyright 2021

This publication is licensed under the terms of the Open Government Licence v3.0 except where otherwise stated. To view this licence, visit nationalarchives.gov.uk/doc/open-government-licence/version/3.

Where we have identified any third party copyright information, you will need to obtain permission from the copyright holders concerned.

This document is also available from our website at www.ofwat.gov.uk.

Any enquiries regarding this publication should be sent to mailbox@ofwat.gov.uk.

OGL