

# DEVELOPING

# A DATA PERMANENT FOR CO-OP ALL PARTNERS



Open Data  
Manchester

CarbonCo-op

# CONTENTS

Executive summary	<b>2</b>
Introduction	<b>3</b>
Choosing a 'data cooperative' model	<b>4</b>
Considering a 'single' and 'multi-stakeholder' approach	<b>4</b>
Reframing data as 'work'	<b>4</b>
Methodology	<b>5</b>
Understanding the types of data collected by Carbon Co-op	<b>7</b>
Data processing within a data cooperative	<b>8</b>
Mapping data flows	<b>9</b>
Designing the consent mechanism	<b>12</b>
Developing a governance system	<b>15</b>
Conclusion	<b>20</b>
Next steps	<b>22</b>

# EXECUTIVE SUMMARY

From July 2020 to March 2021, Open Data Manchester worked with the Carbon Co-op, its members and the wider community to understand whether data cooperatives can:

- **empower** people to take greater control of their data
- **offer** choice about how data is used
- **build** trust and confidence in the process of sharing data

Better data sharing, which can be achieved through a cooperative data structure, can improve how organisations work, and how they work with other stakeholders, which can save money, improve services and create new value. But, like other formal cooperative structures that have helped people work together to get things done for almost 200 years, they take time and patience to set up, and to maintain.

Trust is vital for cooperatives, and is no less vital in structures that agree to cooperate through the pooling of data resources, and while scale often presents opportunities for greater efficiency, size can eventually see a loss of trust, too. However, it is limited resources and technical understanding that may yet be the data cooperative's biggest barriers to sustainability, particularly for those civic-minded organisations that may be most minded to explore them.

For all but one of the months of this project, Manchester was in 'lockdown' due to the COVID-19 pandemic, which had an obvious effect on our ability to run our co-design workshops as planned, but they did go ahead in the virtual world, and we are looking forward to developing this model further.

# INTRODUCTION

This project responds to a call from the Open Data Institute for research and development initiatives focussed on creating **trustworthy**, and **financially sustainable**, infrastructure for ‘data institutions’ in key sectors of our society, the **Data Access Stimulus Fund**.

**Data cooperatives** are one of a number of such data institutions defined by the Open Data Institute, of which it says they: “enable members of the cooperative (organisations or people) to share data with others, with data governance decisions made by those members”.

With access to the experience of **Carbon Co-op**, a Manchester-based, community benefit society that has spent 15 years developing decarbonisation schemes for its members, Open Data Manchester set out to examine the power that these data-powered ‘new mutual organisations’ could have to transform the energy sector, as well as understanding their limits.

Energy is clearly a priority sector for any country – but in the 21st century, governments around the world are responding to the need to rapidly decarbonise in light of stark warnings about what will happen if we continue our rapid consumption of finite, and polluting, fossil fuels. Carbon Co-op, and the many diverse organisations operating in this space, represent an alternative to large, centralised efforts to meet this challenge.

The concept of data cooperatives as a mechanism to look after people’s data has been around for a number of years – they are not limited by sector, or the nature of the organisation – but successful examples include Barcelona’s **Salus Co-operative** and the Swiss **Healthbank**, which both pool and share data at scale to advance medical research.

Organisations like Carbon Co-op could be ideal places for the development of such a data model in the energy sector, because of their close relationship with their members, enhanced by high levels of democratic accountability and governance. Through member-aligned purpose and close stakeholder relationships, participation and trust are nurtured, all necessary for the successful development of a data cooperative.

However, the viability of the Salus Co-operative and the Swiss Healthbank comes in part from their size, so the question is, do small, needs-based, member-led organisations like Carbon Co-op have what they need to take advantage of what a data-cooperative model could offer?

From July 2020 to March 2021 Open Data Manchester co-designed and tested the components of a model to help us understand how data could be used more effectively to further the aims of such organisations and their members. We have worked to identify and understand data-subject motivations, data flows, and the mechanisms that would allow a data cooperative to use and share data more effectively, returning value to the cooperative and its data subjects.

A data cooperative should empower people to make informed choices about how their personal data, and the data they create, gets used, so it is essential to develop, and maintain, trust and confidence in the systems of data collection, pooling and processing. This is built on

the sharing of relevant information with data subjects or members, to ensure that both the technical and human processes for data management and decision-making are understood.

### *Choosing a ‘data cooperative’ model*

The development of new forms of organisation to steward data on behalf of **data subjects** has the potential to create fairer and more equitable uses of personal data, and the other data we create. By developing organisations that are trusted intermediaries for people’s data, more and better data could be shared, providing greater value for people, communities and wider society.

Such **stewardship organisations** protect the interests of the data subjects that have entrusted them to manage their data, in a way that returns value and minimises harm to the subject. They also have a role in enabling better re-use of data through maintenance of data quality, creating efficient processes for data sharing and better continuity of data supply.

A number of data stewardship models are emerging, such as **data trusts** with a fiduciary duty to the data subjects, and **personal data management services** that have a contractual relationship with the individual. Mutually controlled data organisations, such as **data cooperatives** and **data collectives**, however, offer the greatest opportunity for data subjects to have more direct control over how data is collected, pooled, processed and shared.

There are many types of cooperative already in existence, with most designed to address specific needs, such as worker cooperatives, credit unions, producer cooperatives and housing cooperatives.

Applying the attributes of mutuality that these organisations foreground to a data stewardship organisation offers the potential to create an ethically focussed, needs-based, accountable entity that can create value from the data in a way that best matches the collective aspirations of the membership.

### *Considering a ‘single’ and ‘multi-stakeholder’ approach*

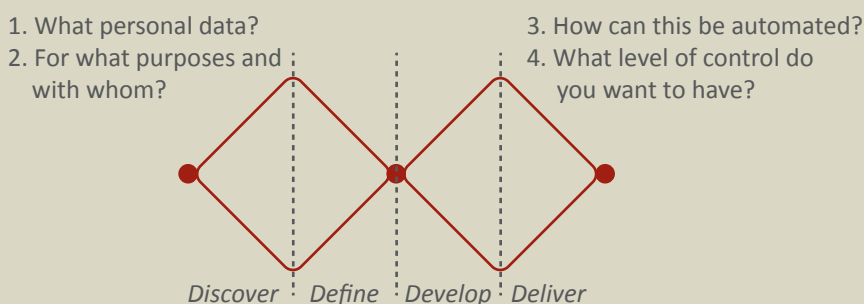
We identified two cooperative models that we believe are of interest in the development of data cooperatives, that of a **single-stakeholder**, ‘worker cooperative’ of data subjects, which has a flat structure and no recognisable hierarchy, and that of a **multi-stakeholder** cooperative, with both data subjects and data users as members.

Although the multi-stakeholder cooperative model offers the intriguing possibility of decisions that balance the interests of both data users and subjects, this project looks at the flat, worker cooperative model, where the membership is made up of data subjects with equal rights, and they make or delegate decisions as to the running of the organisation.

### *Reframing data as ‘work’*

Reframing **data as work**, or an **output of labour**, may help us establish a relationship between the data subject and the data they create. Given that many of the activities we do in our daily lives produce such data, an understanding of the value created may empower individuals, both within and beyond the scope of this project.

This project intended to follow a classic ‘double-diamond’ design methodology, where the work programme went through a discovery, design, development and delivery process. But, in reality, due to the complex and interconnected relationship between the consent mechanism and governance system, and the COVID-19 pandemic restrictions dictating that co-design work took place online, this created a slightly more prescriptive workflow, where design had to be done in advance of the workshops and presented to those in attendance.



*Fig 1. ‘Double-diamond’ design methodology used in the project*

Through a series of co-design workshops, Carbon Co-op members and other interested parties were able to guide the development of the data cooperative model, including:

- shaping the consent mechanism and governance system
- suggesting amendments
- highlighting concerns

Before this co-design process started, two capacity building workshops were offered, so that there was a base level of knowledge regarding different types of data, data rights and relevant legal requirements. This was followed by consent mechanism development workshops and, finally, data governance (see full programme below).

## Introduction to Data

1. What is data and what does it tell us about where we live?
2. Data about ourselves and how to look after it

## Designing a Data Cooperative

3. Exploring a consent mechanism
4. Challenging the consent mechanism
5. Linking consent to governance

## Managing a Data Cooperative

### 6. Exploring data governance

The development process and outputs of the workshops were recorded in the project development logs, which can be found here: [\*\*http://bit.ly/ODM-DCOPD\*\*](http://bit.ly/ODM-DCOPD)

# UNDERSTANDING THE TYPES OF DATA COLLECTED BY CARBON CO-OP

To understand the data being collected by the Carbon Co-op, a data audit was undertaken. Datasets held by the organisation were identified, along with information about location, format, granularity and the presence of any personal data. As well as giving insight into the variety of data held and processed, this knowledge was used to inform discussions in the co-design workshops.

Much of the data generated by Carbon Co-op and its members is kept in separate, cloud-based databases and systems. Many of the data aren't consistently structured or described, meaning that additional processing would be needed to make the data interoperable for sharing.

It is not possible to publish details of individual datasets, due to operational and commercial sensitivities, but below is a summary of data types identified.

## **Organisational**

Carbon Co-op collects a wide variety of data relating to the management of the organisation, and the projects and services that it provides.

Personal data is kept for statutory purposes, such as members' registers and the delivery of services to its members. This can relate to members' involvement with specific programmes, equipment loans or help-desk queries.

## **Sector-specific**

The undertaking of home retrofits by Carbon Co-op can be a highly invasive and capital-intensive procedure for householders. A huge amount of data is collected for these purposes, relating to the fabric of the home, types of appliances, heating and lighting configurations, along with topographic and behavioural data. This allows for the benchmarking of the thermal performance of the home and for making any recommendations about specific interventions that need to be made.

A number of cooperative members are involved with programmes where they need to gather near-real-time data about energy use or ambient conditions within homes. This can range from half-hourly smart-meter readings, through to high-resolution sensor data that can allow diagnostic assessment or evaluation of particular programmes. Ultimately this data may be used to control participating members' energy-intensive applications, such as electric-vehicle charging or water-heating systems.



# DATA PROCESSING

## WITHIN A DATA COOPERATIVE

### **What are we willing to share?**

Through the workshops we were able to get an understanding of the types of data that people might be willing to share, with whom and in what scenarios. What these exercises reveal is that different people value the same data differently and that data has context-specific value.

Naturally, people would gladly share sensitive data in a medical emergency, but not if they were signing up for a website subscription, which indicates that there is a direct relationship between the 'threshold value' of an individual's data and the context in which the data is transacted.

### **Context-specific value**

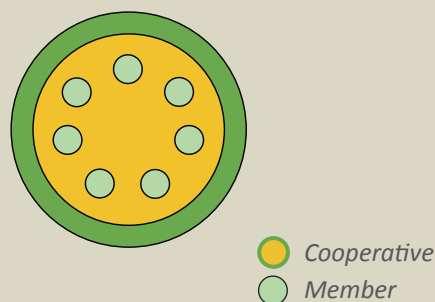
Many cooperative organisations like Carbon Co-op are based around specific issues and concerns, with an explicit alignment of members to those issues. There is potential that such organisations could easily take advantage of the opportunities presented through better sharing of data, given that trust already exists through these aligned aims, underpinned by confidence in the governance and decision-making structures.

Many of the Carbon Co-op members said they would freely share data if they could be confident that this sharing aligned with the cooperative's aims, which in the case of Carbon Co-op is related to decarbonising the energy sector, promoting energy efficiency and reducing the amount of greenhouse gases in the atmosphere.

# MAPPING DATA FLOWS

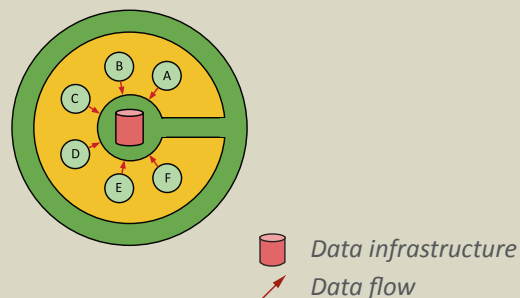
Mapping the different data flows that could take place within the cooperative and with external, third parties is central to understanding the different types of transactions and value exchanges that exist, or could exist.

Five potential data flows were identified that could be managed by the data cooperative, each bringing different value to the members, the cooperative itself and third-party data users.



*Fig 2. Energy cooperative*

The energy cooperative contains and is managed by its members.



*Fig 3. Flow 1 - Member to data cooperative flow*

Flow 1 is the simplest of the data flows handled by the cooperative, where data is shared with the data cooperative for internal use. Data is held and processed by the cooperative, and could be used to improve service provision.

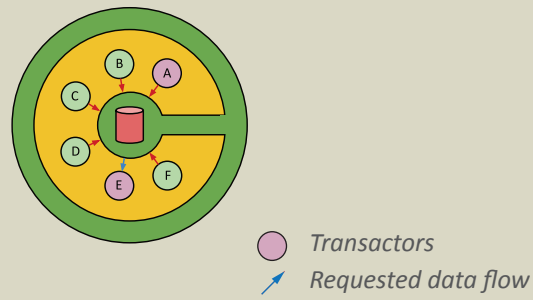


Fig 4. Flow 2 - Intra-cooperative flow

In Flow 2, we see data being shared between cooperative members, facilitated by the cooperative. This could be done through a member granting access to certain data, such as for benchmarking, or for understanding the performance of a particular energy-saving intervention.

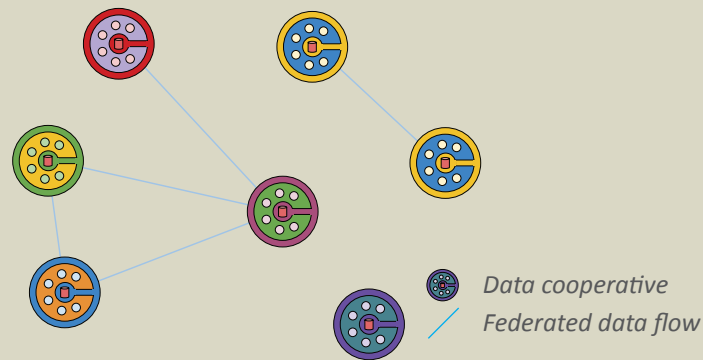


Fig 5. Flow 3 - Federated data flow

A federated flow (Flow 3) could occur between organisations that have the same aims and data-governance processes. In the energy sector, this could be energy-generation cooperatives sharing data with energy-consumer cooperatives, with the data enabling load balancing, which matches supply with demand, or facilitating direct energy supply.

The third-party flow (Flow 4 - Fig 6.) is a more traditional data-sharing or licensing agreement. This would be data shared with organisations that might not have direct alignment in aims, but could offer other types of value. These relationships could be with a broad range of organisations, from academia to private-sector businesses.

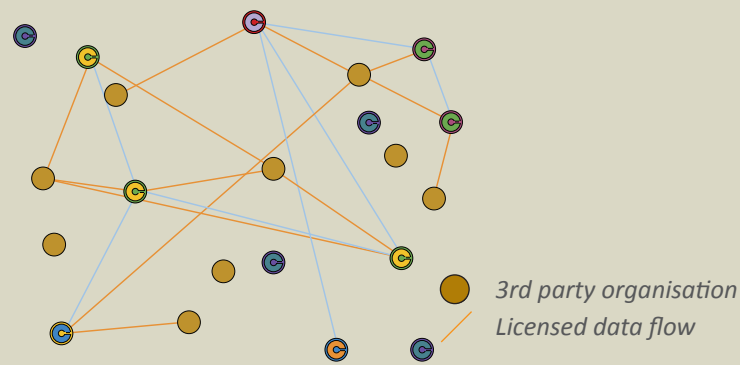


Fig 6. Flow 4 - Cooperative to third-party data flow

The last identified potential flow (Flow 5) is the open-data flow that creates a common, open asset available to everyone.

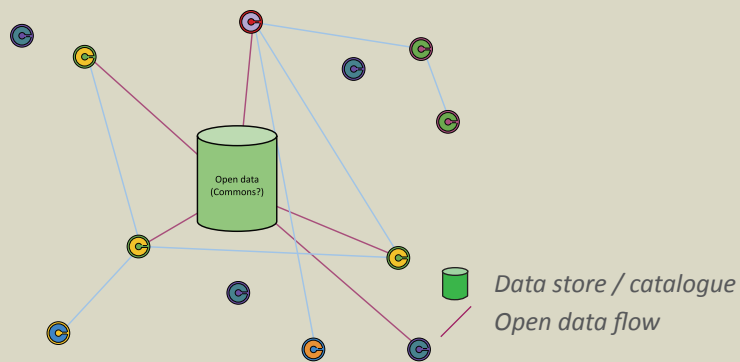


Fig 7. Flow 4 - Cooperative to open data flow

Data released as open data here could likewise further the cause of the data cooperative – so in the case of the energy cooperative, it might be aggregated energy consumption information, or other data that could create better awareness of energy cooperatives and low-carbon solutions.

# DESIGNING THE CONSENT MECHANISM

## Who is responsible?

Considerable time was spent looking at the role of consent, and the need to give people control over the choices made about how and where their data is used.

Our original thoughts were that the data cooperative would essentially be a flat structure – similar to a workers’ cooperative – where every member would have equal rights and control of the cooperative, be a data subject, and have certain responsibilities regarding how the cooperative was run. The cooperative may have officers who carry out duties on behalf of the members, but essentially the members are in control.

But trying to understand who is ultimately responsible for data sharing – acting as the ‘data controller’ – within this kind of data cooperative is a challenge. Also, the need for ‘consent’ within an organisation whose *principal purpose is to collect, pool and share data on behalf of its members* is moot.

There are a number of lawful bases for the processing and sharing of personal data under the General Data Protection Regulations (GDPR) with the data cooperative being able to use ‘legitimate interest’ rather than ‘consent’ as a basis – as the members have all agreed to the purposes of the organisation by joining the data cooperative.

However, even though the lawful basis for sharing members’ data may exist outside the need for requiring explicit consent, and the members are aligned with the aims of the cooperative, the assumption that everyone is equally happy with the same data being shared is a dangerous one.

In one of the consent workshops, we found that people value the same data differently and, even though members might be aligned to the purposes of the cooperative, there may be myriad different reasons for that alignment.

So, while the data cooperative can *lawfully* share its members’ data without further consent, that doesn’t mean that it should – the data cooperative’s purpose is to *help people regain control* over their data. Data cooperatives must, then, create an environment where individual members have control over their data beyond a lawful consent process, to enable members to have greater agency over the data being shared, which further enhances trust.



Persona or archetype permissions offer a more efficient way for people to consent to the sharing of data. It reduces the burden of consent choices by members agreeing to a set of behaviours that most align with their outlook, so any future data shared would be on the basis of those agreed behaviours.

This offers benefits for the data cooperative, as it would allow quicker data-sharing decisions to be attained, but care must be taken in the design and governance of these archetypes, as poor design can build in bias and assumptions, and they need to be checked to ensure that they are working as expected. The decisions made based on them would need to be open to scrutiny, and how they are managed and adapted over time would need to be transparent.

### Traffic light consent:

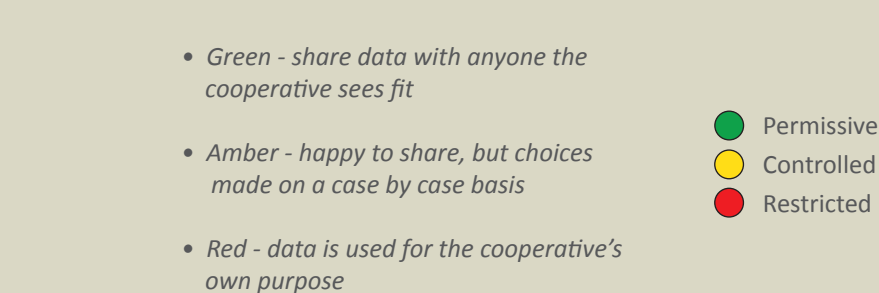


Fig 10. Traffic light consent mechanism

The final, traffic-light consent mechanism is a broad consent mechanism that allows data to be shared on the basis of three options. The most permissive 'green' option allows the cooperative to share the members' data how it sees fit. Operationally this is the most efficient, as there is no need to consult with individual members regarding how the data is used – although ultimately the data cooperative will have other specific aims and overarching governance mechanisms

With the 'amber' setting, the individual can opt for granular consent for specific datasets, allowing the control of more sensitive data, such as a name or address.

The final, restricted 'red' option allows the data cooperative to use the chosen data only for its own internal purposes. As a member, choosing this option for certain kinds of data would inhibit the cooperative from returning greater value from it, and used too frequently, may beg the question as to why the member has joined the cooperative in the first place.

# DEVELOPING A GOVERNANCE SYSTEM

Although data governance covers all aspects of the data lifecycle – from collection to destruction – we are concentrating on how the data-sharing decision-making takes place.

As identified in the development of the consent mechanism, the process of consent is an ethical one, giving cooperative members control over how their data is used, rather than a lawful basis for data processing and sharing.

This control has to be realised within the data-governance process, so that members' consent choices are implemented, and that trust and confidence is maintained within the cooperative.

The data-governance function is a relatively simple one – a member's data is shared if their consent choices match the data user's organisation and purpose of data use – as outlined in the diagrams below.

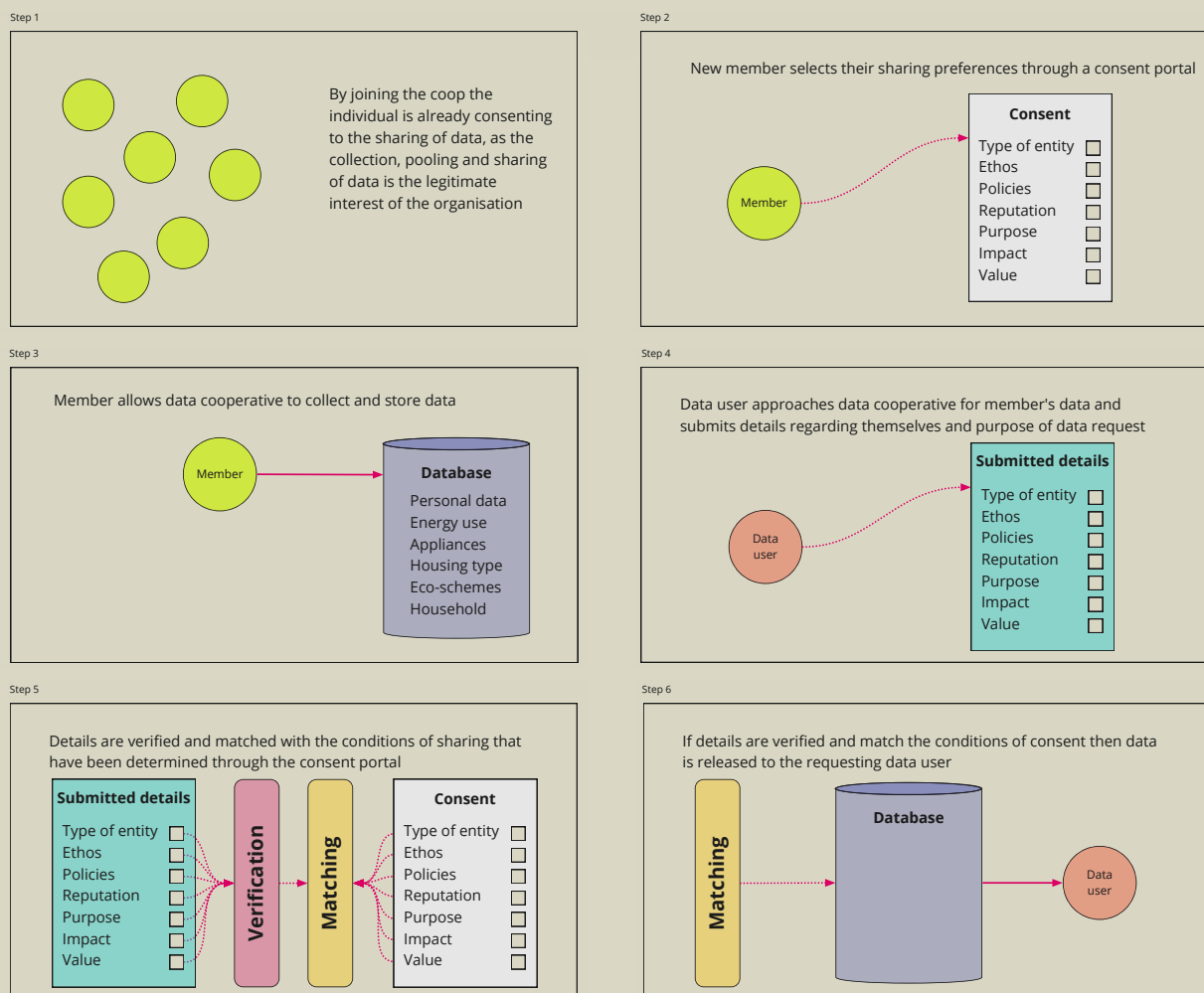


Fig 11. Simplified data sharing process



Although the function is simple, the underlying process is much more complex, because coding abstract, subjective terms, such as 'ethos', 'purpose' and 'value' is challenging, especially if you want to create a process that is fair and transparent.

Developing a persona or archetype-based consent mechanism that will allow you to simplify the different 'options' of consent available to members would help on the consent preferences side. However, assessing not only the purpose of a data request, but the organisation too, is also challenging.

The only comparable assessment process that we found has been implemented by *Ethical Consumer* magazine. This is a cooperative that assesses the 'ethical credentials' of retailers and consumer products through a scoring mechanism based on ratings across: five categories, 19 areas and 300 subcategories.

The scores are then published through its magazine and website. This process is labour intensive, and a team of researchers use a range of reports and insights to calculate the score.

### **Version 1 – assessing a data request**

Version 1 (Fig 12.) of the data governance process for a data cooperative expanded this type of approach to include a number of assessment stages, as shown in the diagram on the next page. A high resolution version can be found here <http://bit.ly/ODM-DCORGV1>



Each step would refer to a ranked table determining the score at each particular stage, creating a weighted score that would be fed into a governance algorithm, which would then decide if data is released or not. Requests that fail to meet a threshold at a particular stage would be rejected.

The assessment tables and the weightings of the assessments would be created through a collective process within the cooperative, with the hope of creating a shared understanding of how requests are decided on.

## **Version 2 – clarifying the model**

The data governance process was presented at a workshop, prompting a number of discussions:

- The automated decision-making by the data-governance algorithm made people uneasy that certain decisions may be made without human oversight or accountability

- The data governance process had, not only to be transparent, but legible so that people could understand how decisions were being made

- The data-governance process would need to be tested manually until people had confidence that it worked

- The system would have to be periodically audited to make sure that it was functioning as expected, with tweaks made to the weighting of each assessment

- The original process conflated the organisation requesting the information, with the purpose of the information being requested, creating inefficiency for ‘trusted’ organisations making ‘new’ requests for data sharing

- The relationship between ‘purpose’, ‘ethical purpose’ and ‘value’ was difficult to differentiate as they are closely related, demonstrating the challenge and importance of defining ‘what we mean’

Following the workshop, we stripped back the governance process so that it had three assessment stages:

1. ‘Data requirement’ assessment check against data and format availability, as in Version 1
2. ‘Requesting organisation’ assessment
3. ‘Purpose of data request’ assessment

This approach created efficiencies by enabling a verified organisation to make multiple requests without going through an organisation assessment each time. Like before, assessment was a scoring process, but this time we looked at defining the organisation and the purpose differently.



Fig 13. Organisation assessment scoring

This process dispensed with assessment tables and included more dimensions, which again would be collectively determined by the cooperative. The output from the scoring would determine a decision, and although this decision could be made using an algorithm, it could just as easily be a manual calculation.

# CONCLUSION

The Data Access Stimulus Fund has provided Open Data Manchester with the opportunity to explore the potential of data cooperatives, supported by Carbon Co-op, along with a co-design process that we believe is vital for transformative, infrastructure projects to succeed.

Those who attended the workshops, although mainly people involved with other kinds of cooperative organisations, found the concept of the ‘data cooperative’, as a means to collect, pool and share data, understandable. It broadly appears to offer members greater control, choice and confidence regarding data sharing.

But, while data cooperatives offer the potential to share data through an aligned purpose, considerable work must be done to create, and create trust in, governance processes and decision-making systems. Although a data cooperative likely has a ‘legitimate interest’ for data processing and sharing, we know from these workshops and the wider work that we do, that data sharing is context specific and different people have differing motivations.

For members to trust data-cooperative governance, the processes need to be transparent and legible, with regular auditing. Members must have a say in the underlying data-sharing assessment criteria, yet it would remain a challenge to create a streamlined decision-making process that is both subjective and fair, with regard to ethical assessment. Automating the data-sharing decision-making process would be more efficient, but is problematic, only potentially overcome if the processes are transparent and legible, and automated when members have confidence in the output.

Specific concerns about the consent model created here related to the potential use of ‘personas or archetypes’, done to reduce the decision-making burden on the cooperative and its members, but which cannot be created without work to achieve and maintain trust in them.

It was identified early on in development that members might not want economic return, but value returned in other ways, so subscription models could be explored. There is also potential for a data cooperative to have a closer relationship with an already existing cooperative and act as a quasi-independent data-management function.

Yet, sustainability of these organisations will remain one of the biggest questions – there is a cost burden regarding the creation and operation of data cooperatives. They are a legal structure with members and statutory obligations. Registers have to be maintained, and filings and accounts submitted. Data also has an infrastructural cost too, such as the creation of the database infrastructure, along with software development and various data-transformation functions.

There is also the human burden of maintenance of the cooperative, making sure the database is operating properly, that governance processes are adhered to and general administration.

If the cooperative is larger, then scale would perhaps allow the costs to be absorbed from revenue obtained through data transactions. The issue of trust versus scale is one that cannot be ignored for data cooperatives – but it presents an interesting opportunity for further research – how can data cooperatives achieve trust at scale?

## NEXT STEPS

Open Data Manchester will continue working with Carbon Co-op to explore the development of a data cooperative model, either as an independent cooperative or a function of the existing one. Over the coming months, we will be prototyping the governance model, as this was recommended by the design group and they would need to see how the governance model works under various scenarios. This is not only to see if it would work, but also allow members to see how it would work.

Further work needs to be undertaken to broaden engagement so that the data cooperative model is understandable and relevant to all members, and to challenge assumptions made during the design process.

Further areas for development are streamlining some of the process – which will come from the prototyping – and exploring how similarly aligned cooperatives can federate data. We are committed to further development of the model and on the 30 March 2021 we are hosting ‘How to Build a Data Cooperative’, which will showcase this work, alongside international examples of best practice, and encourage community members to participate in further developments.

## **Open Data Manchester CIC**

Open Data Manchester is a not-for-profit formed from a diverse group of open data advocates in 2010 that supports organisations, wherever they are in the world, to release data and help people use it.

Whether working in partnership or on standalone programmes, we promote good data practice through expert advice, strong advocacy, participatory events, state-of-the-art research, technical support and interactive training.

Profits from our commercial work go into the community, where we host events and create resources designed to help more people use data more effectively, and ensure the sustainability of our work.

## **Carbon Co-op**

Carbon Co-op is an energy advocacy and services cooperative based in Greater Manchester. They have pioneered community-led retrofit services and established a reputation for innovative, energy-systems approaches. They have an engaged, active and motivated membership, built around their mission of carbon reduction and climate-change mitigation through the creation of new energy services and new delivery models, for example, in facilitating deep retrofits for domestic householders.

To complement these, they have also developed a suite of digital tools and services that they make available to members. This includes offering smart-meter access and home-energy-efficiency assessments, to which they are adding a demand-side-response (DSR) capability as part of their OpenDSR, BEIS-funded programme. They develop technologies with an open-source philosophy, and are advocates for open standards and interoperability.