OPEN:
DATA:

COOPERATION

OpenData
Manchester

OPEN:DATA:COOPERATION

In 2014/15 Open Data Manchester explored whether cooperative structures could enable the supply of anonymised open data through the creation of mutually beneficial personal data stores. The idea of the data cooperative came out of an ongoing conversation between people within the cooperative movement and the open data world about the role of cooperatives, and the possibility that they could rebalance what many perceive as asymmetric relationship between data subjects (people with personal data) and data users (people who use data to develop services and products).

It was thought a cooperative structure would have a number of benefits around the creation of trust and consent, empowering members through the right information and tools, building trust through mutual ownership and creating a model for representation through data. This is the report of the workshops carried out in Manchester and Berlin.

Julian Tait 2015

BACKGROUND

Our modern technologised societies exist on data. Mostly these data are invisible and unknown to us. The services that we interact with, the daily transactions that we make and the way we negotiate through our everyday generate data, building a picture of who we are and what we do. In the age of the Quantified Self there is a growing trend for self monitoring allowing us to track what we do and how feel when we do it. These data are valuable. Aggregated they enable organisations to predict, personalise and intervene seamlessly and sometimes invisibly.

Even for the most technically literate, keeping track of what we do and don't give away is daunting. Personal Information Management Services (PIMS) are starting to emerge offering people the chance to stem the unbridled exploitation of personal data by both public and private organisations whilst also creating monetary rewards for their users. Many of these commercial organisations seek to act as a brokerage service for personal data. The creation of data cooperatives that can act as PIMS have the potential to empower individuals to have more control over their data, creating value for themselves and their communities, and for people to have more of a say in the services that are built.

The sensational revelations by Edward Snowdon shone a spotlight on the personal data that is collected through the IT software and hardware infrastructure that we rely on today. Although highlighting that we unintentionally give away a lot, it perhaps hasn't built a wider popular discussion around protection and usage of personal data. It is inevitable that as the awareness about the data that we produce rises there will be a demand for services that give people more control.

PIMS offer to deliver monetary value to users, but how much value is up for debate as there are differing methodologies to quantify it. OECD research found that the value of a single street address was $0.50 whereas a military record might be $35[1]. Value is also context dependant - data about someone exhibiting behaviours that might indicate a large purchase might be deemed more valuable by companies that manufacture or sell that item. The value of PIMS has been put at £5 billion for decision support services and £11.5 billion for personal data management and life management services per annum [2] making the economic case for the development of these service compelling.

Data cooperatives are starting to emerge that have a broader social and ethical outlook than simple monetary transaction. The Good Data [3] which allows people to control data flow at a browser level with benefits going to social causes and the Swiss-based Health Bank [4] where personal health data is aggregated for the advancement of medicine, are examples of this. As the principles of data custodianship for social good become understood there becomes an opportunity for more to emerge.

STRUCTURE & REPRESENTATION

Data cooperatives have the ability to represent the interests of data subjects and data users. Cooperatives come in many flavours, traditionally coming out of the needs of the membership who subscribe to them. Structures of these cooperatives have generally been organised around a single class of member - workers, producers, consumers, etc. Single class structures create an equitable environment for members but can tend towards self interest and even though often bound by the notion of common good, the mechanism for the creation of the common good or commons is seldom explicit.

Internationally the creation of new forms of cooperatives that explicitly express the development of common good, across multiple classes of stakeholders are more abundant. Social co-ops in Italy [5] and Solidarity coops in Canada [6] often provide services such as health and social care, education as well as community infrastructure projects.

The ability to have multiple classes of stakeholders within a data cooperative has the potential to create a more equitable environment for both data users and data subjects to exchange data. The influence of different classes within the organisations could be managed by fair distribution of voting rights with a user such as a research organisation having the same voting rights as a data subject. Michel Bauwens founder of the P2P Foundation talks about the creation of these new forms of cooperatives, and how they can build a commons both material and immaterial [7]. This commons would be subscribed to by other commons creating entities and licenced to non-commons creating organisations. This suggests a federated relationship between such organisations where commons is shared could exist. But the challenge would be how to define the exchange within this system and if a cooperative contained both producers and users how does this affect the production of commons?

Would a data cooperative necessarily adopt these newer forms of distributed and commons creating structure? There appears to be a consensus that commons creating, multi-stakeholders cooperatives are positive, but they come with increased complexity. Can individual circumstances especially when dealing with communities based around sensitive issues, create an environment for sharing beyond a single class of stakeholder? A single class cooperative may seem to be a simpler, immediate solution for a community of people who have specific needs and issues and where strong trust relationships need to be maintained.

SCALE

Data cooperatives have the potential to work at scale generating and trading in bulk, high worth data as well as forming around smaller communities of interest, such as around a particular health issue, to draw down or negotiate for a better service.

Creating a critical mass of data subjects that would allow the data cooperative to operate at scale would be challenging. Marcos Menendez from The Good Data sees that for PIMS such as themselves would need to create a minimum data subject base of around 500,000 people to be viable. There is potential for data cooperatives to partner with organisations or charities with a similar ethical outlook to build the data subject base.

It may be easier to form cooperatives around a single issue such as diabetes - where diabetics pool their data for particular research causes. The value of such an organisation is that it can help create a more informed decision making process with views of the data subject being strongly represented.

Within a multi-stakeholder model the service provider might also be part of the data cooperative such as a local authority or other public sector organisation. In the current economic climate where multiple personal data sets are held across different public services a data cooperative model could also be used to enable consent across multiple organisations

Making the purpose of the data cooperative understandable is key. Although single issue cooperatives are relatively simple to understand, the representation of data at scale may be challenging. Data cooperatives could act as a platform that builds consent and allowing the representation of personal data across a broader portfolio of interests such as research, service delivery or collective purchasing.

TRUST & CONSENT

Trust and consent should be the foundations on which PIMS are built and data cooperatives have the potential to create both. Mutuality offers an opportunity - especially with a multi stakeholder model - to represent the interests of all stakeholders from individual data subjects to data users - creating an environment of mutual understanding and trust.

The benefits of enhanced trust between the individual data subjects and data users could enable better data and context to be created by data subjects. Through understanding the ways that the data is being used and trusting that the data user understands the needs and concerns of the data subjects has the potential to create a more enlightened and responsive relationship. Even without data users being part of the organisation, the data cooperative would be able to take on the role of trusted representative which in turn could create consent.

Informed consent across all data subjects in a cooperative could be challenging. It would be easy for a data organisation to empower those that already have knowledge and agency to maximise their data, but the data cooperative should have an interest in empowering everyone.

DATA LITERACY

Raising the level of data awareness amongst cooperative members would create more informed decision making, but this task would need to be delivered in a sympathetic and nuanced way where the needs and concerns of the data subjects are addressed.

Ultimately some people may not engage because of service dependency, lack of choice. or a perception that it isn't relevant or useful to engage.

For a data cooperative to represent its membership and control the flow of data it needs to have legitimacy, know and understand the data assets of the membership, and have the authority to negotiate with those data assets on the members behalf.

Decisions around data sharing and understanding the potential consequences are difficult and complex. As an intermediary the cooperative would need to ensure that individual members were able to give informed consent. Data literacy goes some way to achieving this but also mechanisms need to be created that can allow people to have agency over the way that their data is used.

CONSENT

Can one organisation be representative of the broader range of ethical positions held within a membership structure? For practical reasons the data cooperative might have a high level ethical policy but individuals within the cooperative may make data sharing choices based on their personal ethical standpoint. This could be enabled by proxy or preset data sharing preferences. The alternative could be to have smaller federated or distributed niche organisations that have specific restrictions on data reuse.

There exist many mechanisms for the creation of consent. These by and large create the environment for proxy voting in decision making processes. A mechanism such as Liquid Feedback [9] - popularised by the Pirate Party, where an individual bestows voting rights to a proxy who aligns to their position, with the 'liquid' element allowing proxy rights to be revoked at any point. Other mechanisms might follow along the lines of the Platform Preferences initiative developed by W3C [10], which sought to create privacy policies that could be understood by browsers - ultimately considered too difficult to implement. A potentially easier solution might work on the basis of preset preferences based on trusted individuals or the creation of archetype or persona based preferences that people can select.

EQUALITY

How would the argument for greater individual data rights be made when service providers see that personal data mediated through their products as their intellectual property? Work has been done through the midata initiative [11] and the developments of personal data passports - where individuals grant rights to organisations to use the data for delivery of service. UK Government has supported this initiative, but has backed away from underpinning the programme with changes in legislation. The lack of regulatory enforcement may limit the efficacy of any initiative that seeks to grant individuals' rights and agency over their data.

At present there is a certain level of cynicism around voluntary codes of practice where power imbalances exist between stakeholders. The lack of legislation might also create a chilling effect on the ability of data cooperatives to gain the trust of their membership due to their inability to totally control the flow of data. Existing UK data legislation does give data subjects rights to access personal data held by external organisations through Subject Access Requests. A data cooperative could act as a proxy for individual members automating regular Subject Access Requests. This model is being explored by Our Data Mutual [11] in Leeds, UK. There are challenges with using Subject Access Requests at present. Organisations can charge up to £10 for each request and although provision of the data in digital format may be specified, responses usually take the form of reams of paper print outs with responses taking up to 40 days.

It has been mooted by the UK Government that the cost of Subject Access Requests will be reduced - potentially to zero and that organisations will be compelled to supply the data in digital format. This would go a long way to making the process of automated Subject Access Requests viable but in an ideal world data should be pushed rather than pulled.

SUPPLY

A challenge that all data cooperatives would face would be how they maintain a relationship with their membership so that services based upon, or value that is extracted from the data is not subject to unforeseen supply-side problems. If a data cooperative represented its membership and entered into licensing relationships with data users on behalf of its membership, what would be reasonable for a data user to expect, especially if data subjects had the right to revoke access to data at anytime? With larger scale data cooperatives this may not be too much of a problem as scale has the potential to damp down unforeseen effects.

The Good Data proposes to get around these issues by only holding data for a limited amount of time essentially, minimising disruptions in data supply by creating a buffer. It may be necessary for the data cooperative to create terms and conditions for data subjects to minimise sudden supply-side issues.
Smaller data cooperatives, especially ones that are created around single issues may have difficulty in engaging in activity that requires service guarantees.

Developing a mechanism for federation, cumulatively creating data at scale might be a potential solution, but creating a federated system of consent may be more difficult to achieve. As suggested previously economic activity might be a low priority for such organisations where the main purpose might be to represent members and create the environment for informed service provision.

The challenge facing federated data cooperatives and how they interact is undefined. It has been noted that building distributed and federated systems is difficult, and that centralised systems persist due to operational efficiencies. The advent of alternative forms of ‘blockchain’ transaction could enable distributed organisations to coexist using ‘rules based’ or algorithmic democracy. But alternative transaction systems and currencies often face challenges when they interface with dominant and established forms of currency and value. How data cooperatives could practically use these new mechanisms for exchange needs to be explored.

SUMMARY

1. Data Cooperatives are owned by their membership and therefore are more accountable;

2. have the potential put a halt to the over collection of personal data through representing data subjects and advocating on their behalf;

3. can create value for their membership;

4. can form around single issues or scale with many data subjects;

5. can become representative and be used to create change;

6. could help their membership to understand how data is used - data literacy;

7. can liberate personal data on members behalf through Subject Access Requests;

8. can encourage better data and context to be produced by data subjects;

9. build trust and consent within the organisation and

10. can be a blend of open data and personal data organisations

REFERENCES

[1] *Exploring the economics of personal data: A survey of methodologies for measuring monetary value (2013)*
[2] *Personal information management services - An analysis of an emerging market (2014)*
[3] *Website - https://thegooddata.org*
[4] *Website - https://www.healthbank.ch*
[5] *Website - http://www.caledonia.org.uk/vanek.htm*
[6] *The solidarity cooperative in Quebec: a new formula in North America - Jean-Pierre Girard (2002)*
[7] *Website - http://p2pfoundation.net/*
[8] *Website - http://liquidfeedback.org/*
[9] *Website - http://www.w3.org/P3P/*
[10] *Website - http://www.ourdatamutual.org/*
[11] Website - *https://www.gov.uk/government/news/the-midata-vision-of-consumer-empowerment*