Building an ethical energy data co-operative

Open Data Manchester CIC / Carbon Coop

Workshop dates:

Introduction to Data

1. What is data and what does it tells us about where we live?
   Thursday 17th September 2020, 6.30pm – 8pm
2. Data about ourselves and how to look after it
   Thursday 24th September 2020, 6.30pm – 8pm

Designing a Data Co-operative

3. Designing an ethical way to share data (workshop 1)
   Thursday 1st October 2020, 6.30pm – 8pm
4. Designing an ethical way to share data (workshop 2)
   Thursday 15th October 2020, 6.30pm – 8pm

Managing a Data Co-operative

5. How do we manage a data cooperative? (workshop 1)
   Thursday 19th November 2020, 6.30pm – 8pm
6. How do we manage a data cooperative? (workshop 2)
   Thursday 21st January 2020, 6.30pm – 8pm

Project context

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. By creating organisations that are trusted intermediaries for people's data more and better data could be shared, providing greater value for individuals, their communities and wider society.

Mutually owned organisations such as cooperatives have by their nature a close relationship with their members enhanced by high levels of democratic accountability and governance. Many cooperatives are created out of a specific need, such as housing, the need for ethically sourced goods or childcare, and it is through collective action and close stakeholder relationships that their purpose is maintained and accountability and trust are nurtured. Applying the attributes of mutuality 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.

The concept of data cooperatives has been around for a number of years, but for small needs-based, member-led organisations the ability and capacity to understand, investigate and develop a data cooperative model is limited. Therefore we are developing a prototype model in partnership with an existing Carbon Coop to understand the potential value of the data that its members create through its activities.

A number of different types of data relating to energy use and performance are collected by members and the cooperative itself through its day to day operations. These data range from power usage from smart meters, home energy performance assessments, specific electrical appliance data from water heating to electric vehicles and other ad hoc data.

By pooling member's data there is potential to develop new ways of using energy more efficiently, returning value for members, the cooperative and wider society.

Workshop 1: What is data and what does it tell us about where we live?

Thursday 17th September 2020, 6.30pm – 8pm

Link to slides: 200917 CAC-DCO Capacity Building 1.pdf

Roll call:
Sam - Open Data Manchester
Julian - Open Data Manchester

Agenda:

- Introductions / Introducing the Data Coop project
- Who we are
- What, exactly, is data?
- What is open data?
- Being responsible with data
- Exploration: Data about where we live

--

Question: What is data to you?

(please write below!)

Information
Information with/for a particular purpose
Information in general
A version of reality
Knowledge > info >data
Information

--
Have you used/accessed open data?
- Yes
- Yes
- Yes - but only to see what's there rather than to use it
- Yes lots - but then we are open data manchester :)

What data have you accessed?
- ons data, gm mapping, geographical admin area data, other maps
- Waste arisings data (for environmental campaign)
- UK energy stats
- FiT data
- Weather data!
- Satellite data, open street map data
- EPC data, gas consumption data, English Housing Survey data and similar
- Traffic data at work
- ONS, NHS Fingertips, TfGM data, Eurostat
- Air quality
- Air quality and crime
- Census data, bmd data, tithe maps, wills, mariners registers - family history!

What data would be useful to have access to?
- More post-occupancy data following retrofit (to decarbonise housing), including temperature, energy data and indoor environment impacts, plus householder perception and experiences
- It might be worth linking the GM map to the postcode level energy data that was issued as "experimental" for 2018. I am including the links below but would add a word of warning about quality. For my postcode there are 39 houses and data for 39 gas meters but only for 37 electricity meters.
  https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/853776/Experimental_postcode_domestic_gas_statistics_2018.csv

  https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/853784/Postcode_level_standard_electricity_2018_postcodes_A_to_L.csv

  https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/853785/Postcode_level_standard_electricity_2018_postcodes_L_to_Z.csv

  https://assets.publishing.service.gov.uk/government/uploads/system/uploads/attachment_data/file/853786/Postcode_level_Economy_7_electricity_2018.csv

- I have emailed BEIS the data for 2019 will be released 22nd Dec or 28th Jan
- Its also worth noting how the statistician deals with the issue of low numbers that could breach GDPR rules " *The postcode level data excludes: Postcodes that are disclosive are defined as such if they:*
  - contain less than 6 meters
  - the mean and median consumption equal 1 or if median consumption is equal to 0
  - the median consumption is equal to 0 and the number of meters is less than 12 If a postcode is determined to be disclosive, it's consumption is included in the partial postcode, as long as it is itself not disclosive.

For example, if NR11 3DZ is disclosive, it would be included in NR11. Aggregated partial postcode consumption (eg: "NR11") only includes consumption from disclosive postcodes data and not total consumption in that area. Any disclosive meters at the partial postcode level have been removed from the dataset."

- Substation level electricity time-varying demand (professional interest!)
- Royal Mail PAF file
- apis to access my ovo and hive data
- Local information about services, for example that help to put cases together for planning/infrastructure services and health provision - so I can engage with evidence in public consultations

What barriers have you come across?
- Ethics and personal data queries, researchers not working openly in the past - much research data could be of value for future decision-making
- comparing data from different sources because of different parameters
- Lack of standardised data collection protocols and sector specific repositories
- Metrics changing over time (policy changes) - so hard to compare historic data
- Quality of crowdsourced data
- Data doesn't exist, is ancient or impenetrable to a 'lay' person
- Not granular enough for my local area
- Mixing of type of summarisation making it difficult to compare data eg quartiles v means v medians
- Data publishers not being transparent about the quality assurance process they went through plus what might be missing from their data (this is the case where there are data standards!).

--

Exploration: Data about where we live

[Exercise sheet](#) - feel free to make a copy and write on it!
Any interesting/surprising finds? (write below)

- Data doesn't often measure what you think it measures - e.g. 'unemployed' includes students so not an indicator of economic precarity

What questions do you have at this point?

Who do we see as potential members of the data co-operative - energy companies, energy customers, community groups?
What do we think those people/organisations need to do that a data co-operative could help with eg buy power / save money / reduce consumption / build a smart home integration?
Do all smart meters record the same attributes?
What other attributes might be useful eg from external sources?

Potential scenarios to explore

What happens to data if a meter or service provider changes - would it still be comparable over providers and over time?

- Ideas from Kate: Digital footprints and GDPR (both I would love to learn more about!), do any 'data logbook' type apps exist so people can track where they think their data is shared and how it is generally used, at any time?

Workshop 2: Data about ourselves and how to look after it

Thursday 24th September 2020, 6.30pm – 8pm

Link to slides:
https://drive.google.com/file/d/1LJifkt9ZFmR7VR8U9Hl3X2e1SVe3iAy5/view?usp=sharing

Roll call:
Sam
Julian

Agenda:

- Introduction / catch up
- Energy data exercise
- Discussion: our personal data
- Presentation: data about ourselves
- Tools for protecting our data
- Exploring some scenarios

Energy data Profiles

What could be going on in the scenarios below?

Example
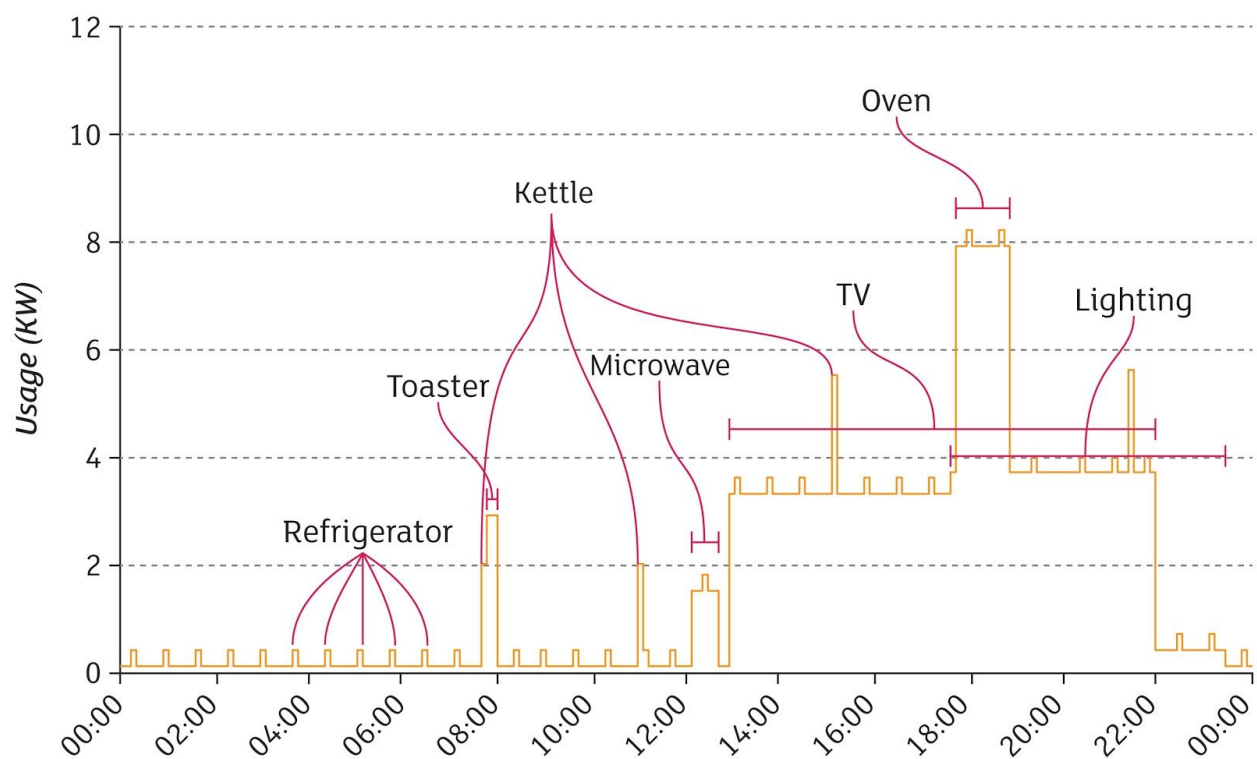
Anomaly 1
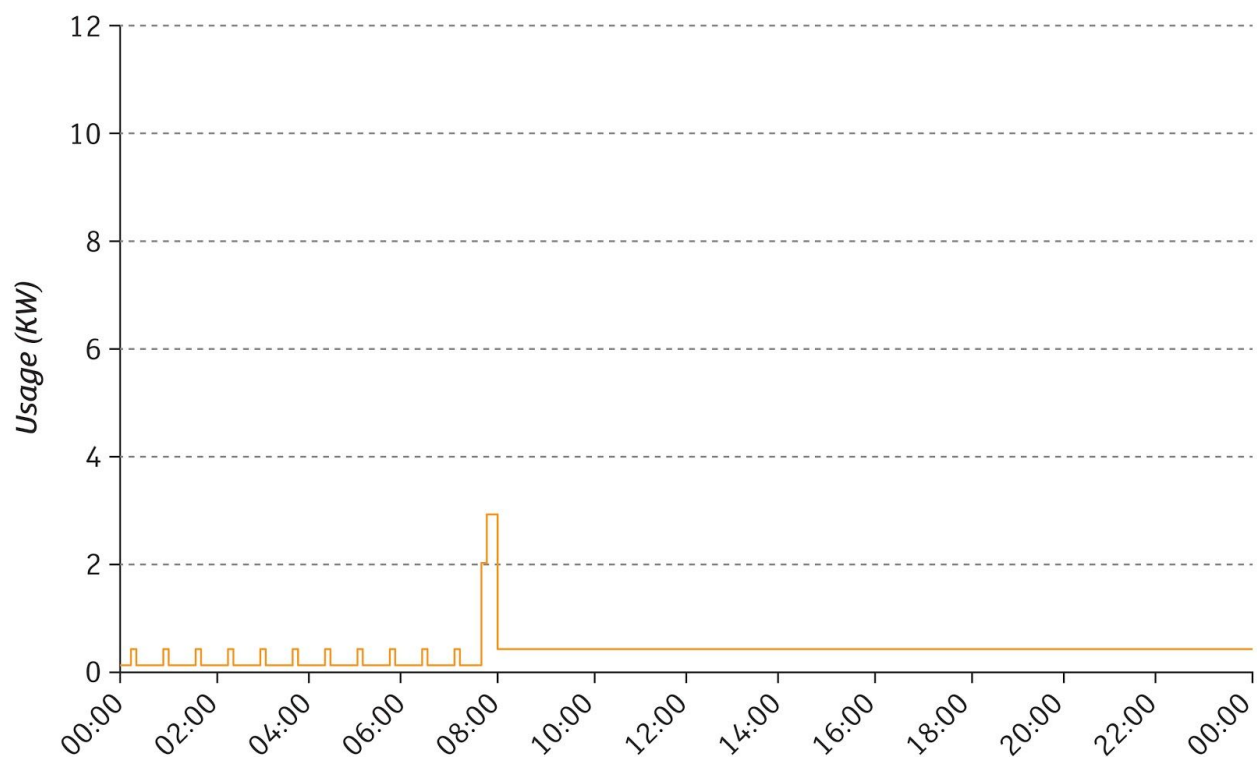
On holiday but maybe the fridge cools down, or somebody pops in to feed the cat! (opening the fridge door!)
 Had breakfast then went off to stay with family (left a light on for security, but turned off fridge)
Gone away, left a light on?

Anomaly 2
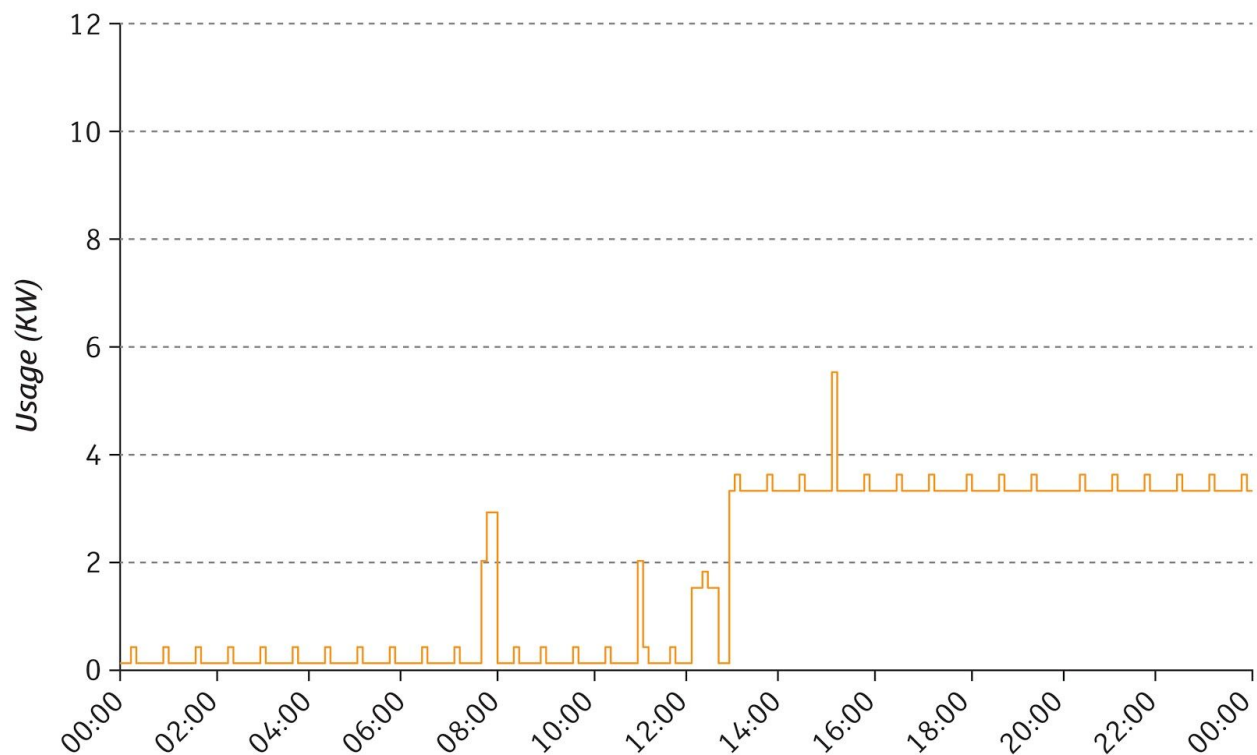
TV marathon - they get a takeaway and binge watch something (or fell asleep in front of it, or worse).

Tv left on

Turns an electric heater on

Electric heating system (not copying - had the same idea! :))

Anomaly 3
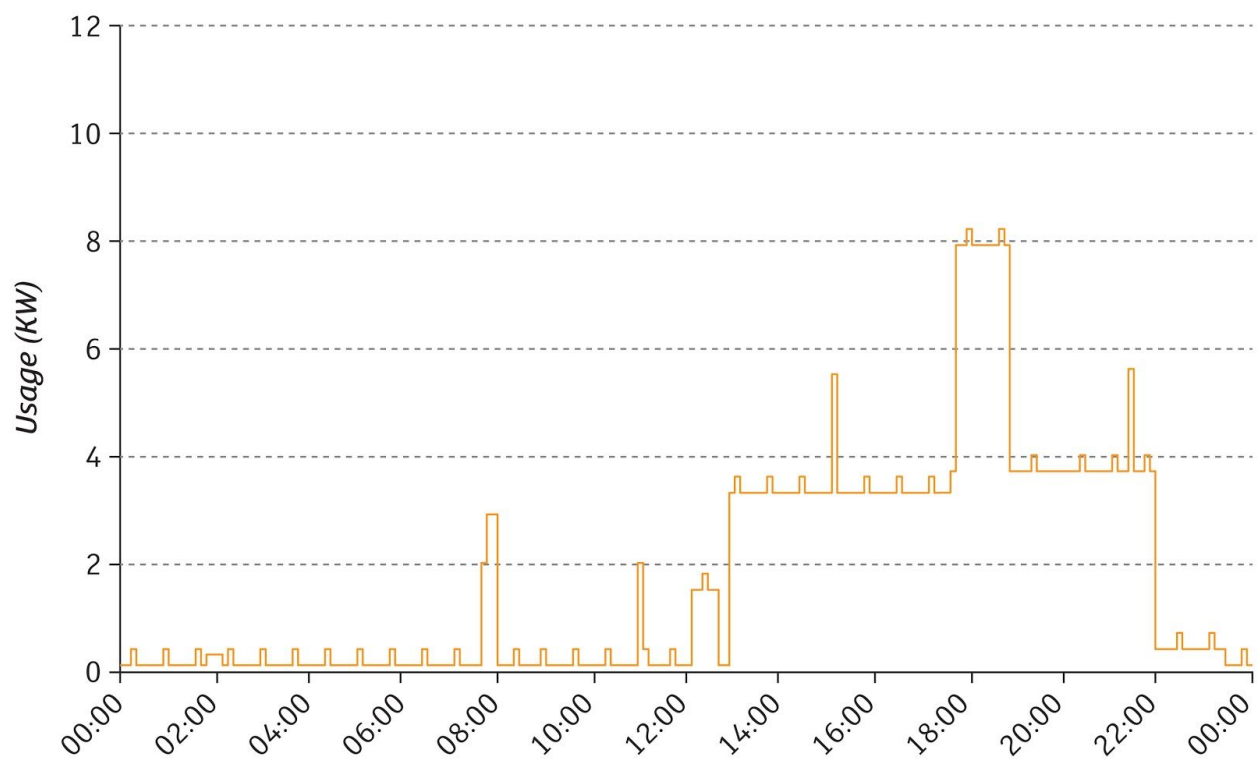
Evening baking
Dinner in the evening.

Anomaly 4
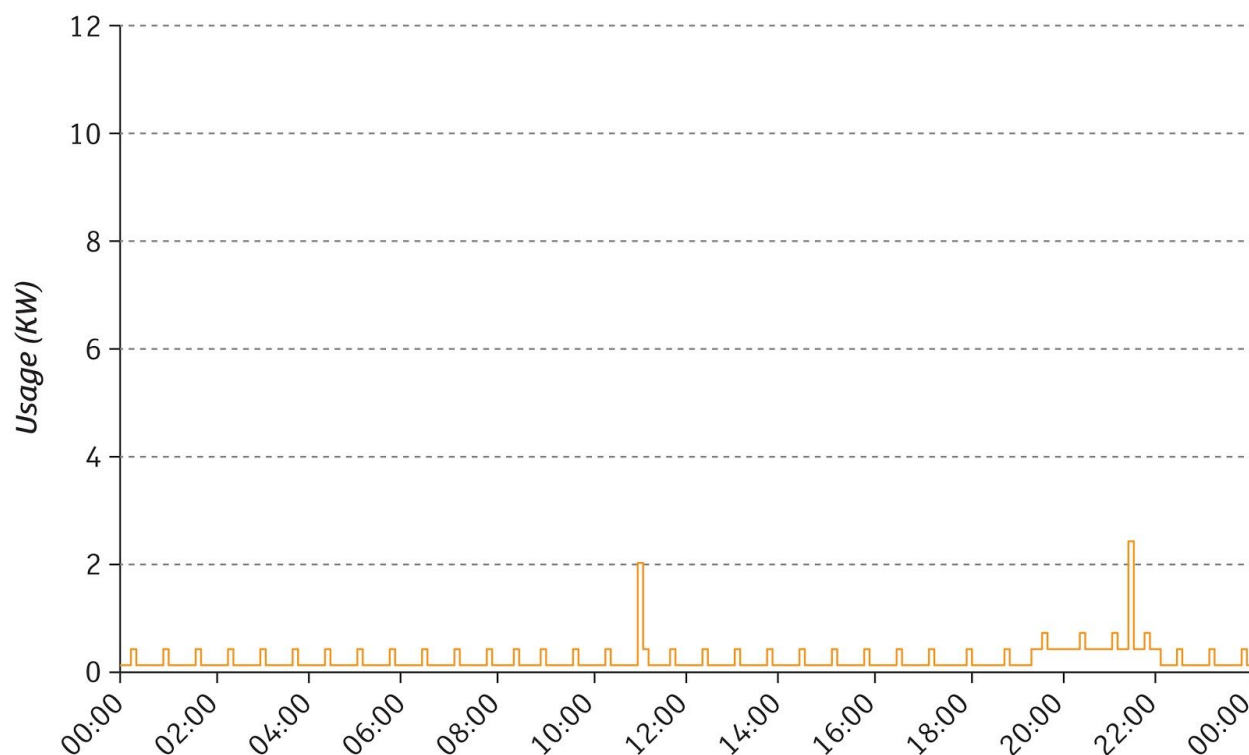
Someone is on holiday, their neighbour comes round twice a day to check things out (and helps themselves to a pot noodle in the evening…).
Summertime - very low energy user? Or a caravan??

Discussion Questions:

Q1 What smart devices do you own? How do you use them?

Mobile - internet (e.g. checking email, finding places).

Hive for 'heat by' a time. Supposedly smart TV.

Smart phones - use for banking, mapping, emails etc

PV inverter - sends power import & generation to server in China & controls domestic battery.

Use of phone to augment memory

2. What online services do you use? Can you remember what personal details were asked for?

Have input heating schedule into Hive.

Email/Social Media A/C: Date of Birth, name, location.

Social media, tracking family and friends, learning language, listening to music

Banking: Name, Address, DoB…

Learning apps - music, exercise etc

3. Have you ever been part of a data breach? How did you find out? How did it make you feel?

Yes. Found out because they emailed me, but but I also go to https://haveibeenpwned.com/ . Made me feel annoyed - in one of the cases I WROTE to them asking if they had been hacked, they said no and then later...Never worked out if I was right/whether it was co-incidental.

Yahoo, and others. Have been emailed and advised to change password. Feel uneasy.

Email address put publicly online - really annoying! Also, was involved with *causing* a data breach.

No but suspect my information has been shared

Email spam - weird email senders I have no idea how they got my email!

Tools for protecting our data

Ghostery - https://www.ghostery.com yes
AdBlocker - https://adblockplus.org yes
Have I been pwned - https://haveibeenpwned.com/ yes
Disconnect https://disconnect.me/
Duck Duck Go - https://duckduckgo.com
TempMail - https://temp-mail.org/ browser plug-in on firefox that generates a disposable 24h email address

Please add your own! What do you use?
PrivacyBadger
Google-free Android OS (from /e/ foundation - can [buy a Fairphone with it pre-installed](#))
ecosia
https://vpn.mozilla.org/
OKFreedom as a VPN

Scenarios - exploring personal data

Choose a scenario from the list below, and discuss in your group the following questions:

1. What are the benefits of each scenario?
2. What are the potential dangers?
3. What can you do to protect yourself?

Scenario 1: You sign up to a new car insurance provider and they ask you install a black box in exchange for a lower premium.

Benefits: Lower premiums, more sensible driving and drivers. Logs safer driving. Acceleration, speed.

Potential dangers: People who work out of hours shift patterns are penalised. Value is not necessarily returned.

What can you do to protect yourself: Don't insure with companies that offer this service. Not convinced that these actually are cheaper

Scenario 2: You've been concerned by a rise in crime in your neighbourhood, and your family member gives you a new smart doorbell.

Benefits: see who is at the door before answering. Can see who is at your house when you are not at home. Catch the fly tippers! Catch scammers

Dangers: No idea who sees the data / pictures. Could miss warning of local danger if you didn't answer the door and there was e.g. a gas leak on the street. Profiling. Exposing friends/family/suppliers privacy too.

Protections: ?!

Scenario 3: You've got a new smart TV and it's asking you to log in to WiFi.

Scenario 4: A new dockless bike scheme has started in your area. To use it to have to install their smartphone app.

1. Advantage: Very convenient esp if it works between cities!

2. Danger: Bikes will be GPS tagged so you will be tracked on the bike and also between cities. Also it quite an exclusive way to run a bike hire system

3. Mitigation: just don't use it. Get your own bike or hire from a normal company.

Benefits: Bikes conveniently located, don't have to use docking stations. Don't need your own bike for getting round time

Dangers: Personal information needed - banking details, date of birth, name, age. Tracking can reveal a lot about you.

By GPS blocker, but that might not help as they would know where the bike was picked up or left. Don't use the bikes. Use someone else's phone

Scenario 5: Your dog keeps running away! You buy a GPS tracker collar for your dog.

1. Find your dog prevent it causing trouble for you

2. Dog tracker will track you when you walk it - could allow anyone else in your household to track you! E.g. suspicious/controlling spouse (is this where 'dogging' comes from?)

3. Tracker only activates when off lead?

Scenario 6: You want to change energy suppliers and you browse a price comparison site.

Benefits - saving money. Very difficult to compare lots of suppliers otherwise.

Danger: personal data being shared with energy suppliers and possibly being sold to them. Some of them will create a personalised price for you based on historic energy data, which exposes a lot of data to advertisers!

Protection: enter as little personal data as possible and fake it if it doesn't alter your benefits?

Scenario 7: A friend on Facebook sends you a quiz to determine your intelligence

Benefits: distraction.

Dangers: Attention-hoover. Profiling.

Protection: UN-friend such "friends"! Disable cookies before hand.

Reflection

1. Has anything from tonight surprised you?
2. Will you do anything differently?

That Facebook Algorithmic Factory diagram. (plus this one for me - Kate!)

Yes - If I get a(nother!) fairphone Ill get that Google-free option. (hooray! NB you can also port /e/ onto an existing Fairphone 2 and 3. And other Google-free OSes are available - Rich)

I will look at Goshtery already have adblocker.

We have spent time thinking about the actual data but the meta-data is probably just as valuable.

Not hugely surprised by anything

Won't do much different - I'm usually willing to trade some sharing of my data for the benefits it brings

The story about the gentleman's pacemaker data getting used to convict him of arson was a shocker. (oh yeah this one for me too - Kate!). This session has restored my concern about my PV data (includes 30 second-resolution energy consumption) being sent to the cloud - but setting up an alternative will be a right palava.

I had not thought much about smart phones being so connected to Google - I like to use second hand phones but maybe I will look at that fairphone option too (NB you can port /e/ onto many many types of phone and they sell second hand refurbished ones with it preinstalled)

I am going to explore the tools for protecting my data - thanks all! Most are new to me. I think I will generally be a bit more alert to data sharing. I thought I was fairly cautious but could be more mindful for sure.

The google-free Android phone! If it's true i'll def get a fair phone when my current one packs in.

Newstory regarding pacemaker

https://www.cbsnews.com/news/mans-cardiac-pacemaker-data-led-to-arson-charges/

Workshop 3: Designing an ethical way to share data

Thursday 1st October 2020, 6.30pm – 8pm

Link to slides: TBC

Roll call:
Sam
Julian
Richard - Carbon Co-op

Agenda:

- Overview/recap
- How the next two workshops will run
- How sensitive is this data to you?
- What would you be willing to share and under what conditions?
- Summary and next steps

How sensitive is this data to you?

Data Types Jamboard - Group 1 (Sam)
Data Types Jamboard - Group 2 (Julian)

What would you be willing to share and under what conditions?

Questions:
1. What would you be willing to share?
2. What questions does this raise for you?

Group 1:

Scenarios**:**
1. You want to change electricity supplier
2. There is a project within the coop that wants to measure energy use.

3. A large electricity provider offers you £50 gift voucher for information

- I just wouldn't bite! Wouldn't feel inclined to give any info because it's probably worth more than £50 unless they could justify a good social purpose.
- Depend on who they are!
-

4. You have responsibility for looking after an elderly relative who lives in their own home and are considering using assistive technology.

- Disaggregation of data. Wouldn't be worried with sharing monthly energy use, but granular 'every 5 minutes' not comfortable
- However, this could be life-saving, as per last week's slides
- What level of consent do you need to get?
- Are you the person who makes the judgements on their behalf?
- Capacity of the elderly person is important - power of attorney.
- Willing to share my personal contact details as they would need to get to me and also with the people who are providing the care.
- Rather than making raw data, shouldn't it be translated into alerts etc? So the actual data isn't being shared.

5. You've just retrofitted your house and you have access to energy performance data.
6. You are the proud owner of a new electric vehicle and are trying to work out how to use it most efficiently
7. Carbon Coop wants to work with other energy cooperatives and share data
    - This links to 2
    - Half-hourly usage, but that's measure of how close I (Richard) am to Carbon Co-op projects and trust in the other (nonprofit) cooperatives.
    - Anonymised
    - But it would give a lot of detail away (occupancy)
    - Approximate location - maybe postcode
    - It would depend on what you could deduce from that location data.
    - If it's less than 12 properties, they are excluded (see Ian's links on P4).
    - Heating systems, house types, metadata of the house.
    - But when you include these data, it becomes less anonymised.

8. Carbon Coop has negotiated for lower energy prices from a supplier

Group 2:

Scenarios**:**
1. You want to change electricity supplier
Size of bills, size of house, opt in regarding.

Try to guess your consumption
A lot of data is given away
How much energy I consume a year - that is what I would like to give
Balance of what you use in the evening over what is used in the day
Customers as meter points
Takes friction out of choice

2. There is a project within the coop that wants to measure energy use
   Have my data - but not with my personal details
   Member to member can create gamification
   Need to know why the data was shared more widely
   Feedback loop - show me what you've done with it.

3. A large electricity provider offers you £50 gift voucher for information
   All the false data that they want
   Dependency and need leads to people giving away more data
   They need other people's email address
   False email address

4. You have responsibility for looking after an elderly relative who lives in their own home and are considering using assistive technology.

5. You've just retrofitted your house and you have access to energy performance data
6. You are the proud owner of a new electric vehicle and are trying to work out how to use it most efficiently
7. Carbon Coop wants to work with other energy cooperatives and share data
8. Carbon Coop has negotiated for lower energy prices from a supplier

Consumption data treated as personal data - due to potential burglary
But no issues regarding saving carbon
Don't want to be behaviorally analysed
Washing machine spin cycle
Value exchange - eg give this if you can give me lower carbon - be clear about the purpose and benefit of what I get if I give
Sale of companies - give me the right to opt out with my data

Workshop 4: Designing an ethical way to share data part 2

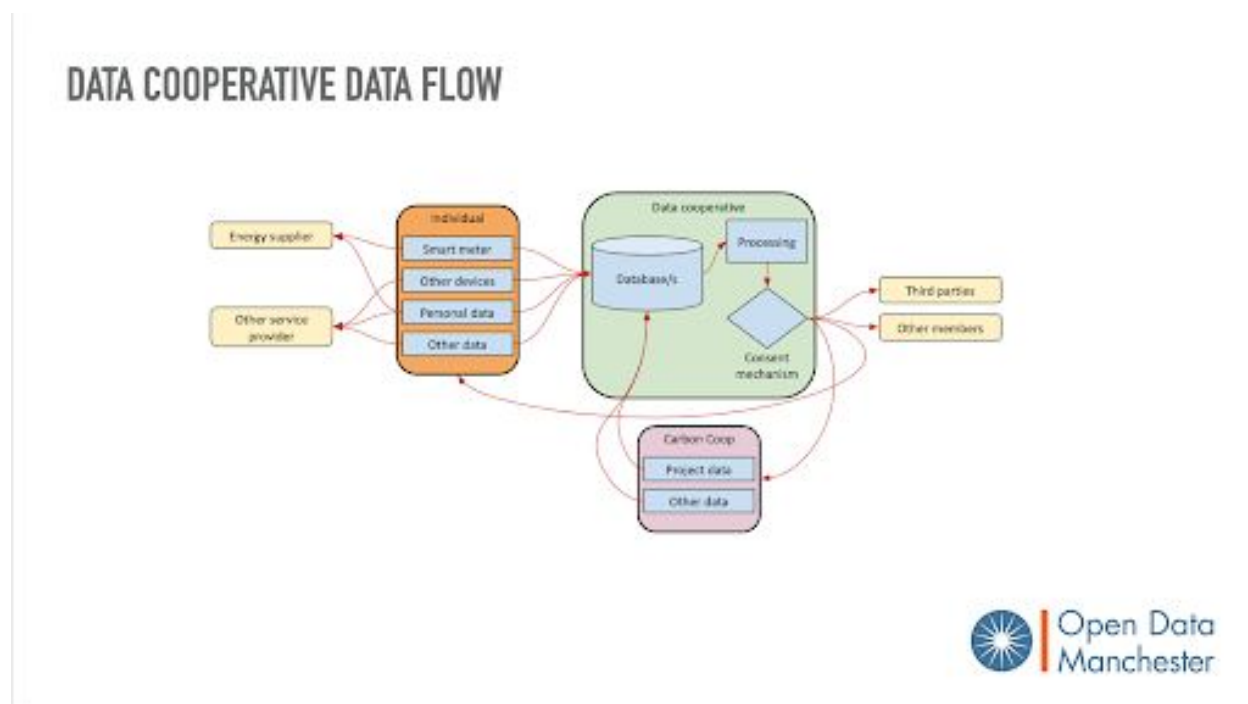Thursday 15th October 2020, 6.30pm – 8pm

Link to slides:

Roll call:
Sam
Julian

Agenda:

- ● Introduction / catch up
- ● The data cooperative data flow
- ● Data processing
- ● The consent process

Data Cooperative Data Flow

Feedback

- Challenges of getting that data? Ie some devices might have commercial benefit so will they allow you sharing it with the data coop?
- Could the data cooperative be service - rather than my boiler etc sending data off to its own server/service.
- There must be a consent mechanism between the individual and the data coop?
- Ian
  - Is there an implicit consent - For example functional cookies on a website that won't work without them
  - Thinking back to my days running a clinical research business a couple of points come to mind
    - Informed consent - how do you ensure that the data subject has actually understood what they are consenting to? The guidance may have changed now but in the 1990's we had to provide information leaflets at reading age 9, ie mid primary school or The Sun level.
    - In clinical research it is a requirement that consent can be withdrawn at any time.

Data Processing flow
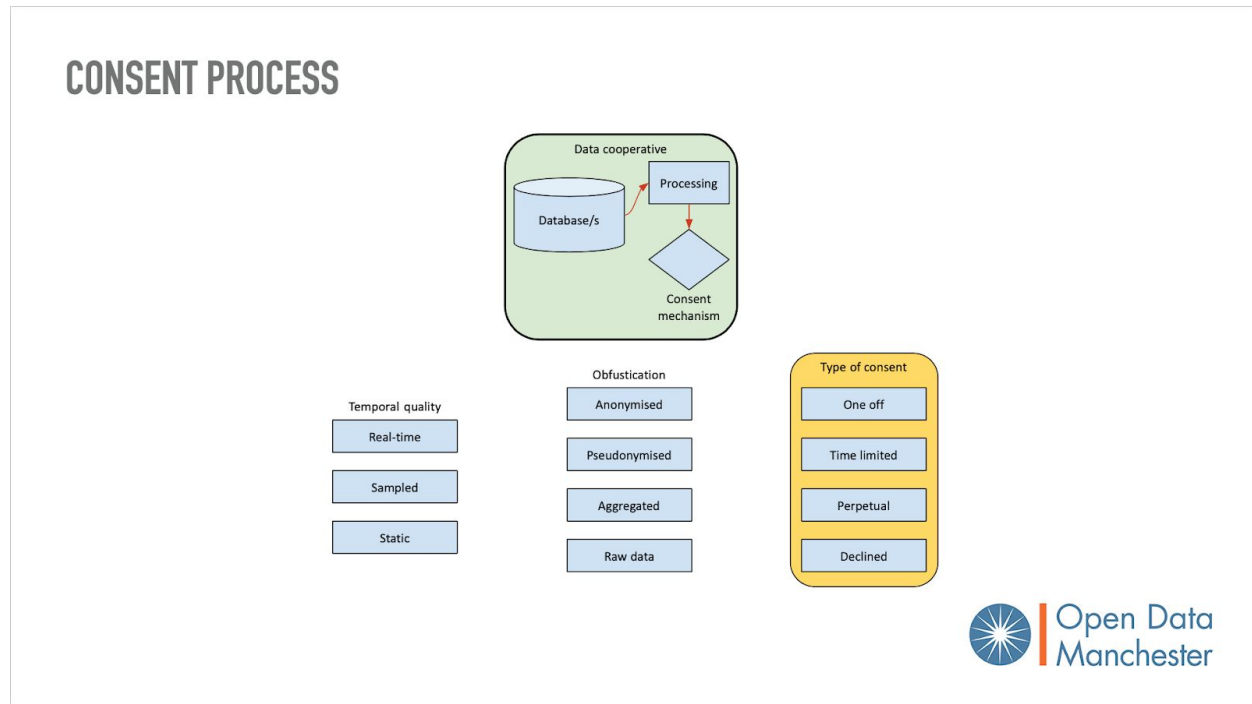
DATA PROCESSING
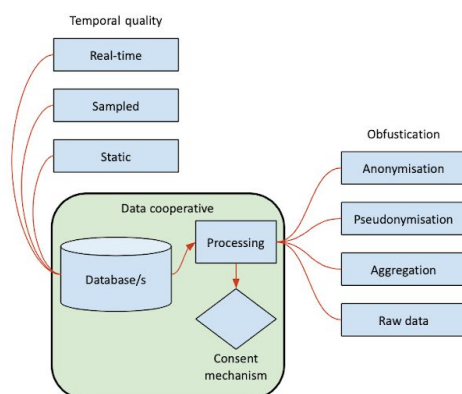
Feedback

- How are you addressing the idea that there is no such thing as "anonymisation" (i.e. that its not really possible, because people can reverse it)?
- Risk for the data coop in handling the sensitive data.
- Hierarchy of obfuscation - is aggregating data in the right place?

Consent Mechanism Process flow

Consent Mechanism Feedback

Scenarios

Explore each scenario using the following as a guide:
- What data would you choose?
- What temporal quality would you allow?
- What level of obfustication?
- Would you be happy with the data cooperative making those decisions on your behalf?
- What would you need for the coop to make those decisions on your behalf?

Group 1 Scenarios

1. **Carbon Coop want to use your energy data for research into the success of a retrofit project.**
Need to know the starting state of building - one off
What work was done - one off
Air leakage test data - one off
Energy data - gas data and electricity - real time - would question the amount and frequency of data.
Data from embedded sensors - temperature and humidity.
Details of energy using appliances
Would want to have a way of saying or consenting on who access this data - want a mechanism regarding approach from X to give consent
Would want a bullet point list
Would want to agree for that data.
I am allowing the coop to hold that data would I be comfortable. Nervous about access.

2. **Shell Energy are doing some research into local grids and the viability of a planned local hydro generation facility and want to use your energy data.**
Sod off
They don't need this data
They only need aggregated demand
Depends how local
What about energy mix
They would have to justify why they would need that data
If company was genuine and trying to develop local solutions
More valid question would be the type of technologies that you are actually connecting on that circuit.

3. **A third party want your contact details and energy usage to sell you energy saving devices.**
Buying hours of heat
Energy as a service
Do I trust these people to have my contact details, and I thought or knew they had good intentions

4. **A fellow coop member wants to compare energy usage of people with similar houses so that they can improve their energy efficiency.**
Trust and sharing works within a small group
If stranger would want to know a bit more about it first
Depends what they want to know about.
No if they are creepy

Intention is important
Appropriate granularity
Would be interesting if the data was used for research

Your reports how was your data used or accessed over the last month

Group 2 Scenarios

1. **Carbon Coop want to use your energy data for research into the success of a retrofit project.**

House metadata (static)
Internal temperature sensors & energy consumption
Would want to require anonymisation in any project outputs
Time limited to length of project? Might need longer term data for contextualisation

Depending on what your health conditions are, you might think about different things - so time limited makes sense, anonymised.
Because this is the first time something like this is done, you might want to start with the least level of obfustication and then see what they produce based on what they produce with what you give - before making any further decisions.

Granularity of data important - high level energy use isn't much use at all. You'd need things like external temperatures, how many people in the house… this years will be so different from last years!

How far out do you want to go? Ie what suppliers you used? That will affect the retrofit surely?

Knowing the purpose for which they're going to use the data is very important. Would need consent for all individual items! Could end up with a huge truth table!

As a novice, i would want something like 'if you give me this, this and this, we can give you this'.

Purpose - ie you care about decarbonisations so would you be happy for the coop to share this data regardless on a project that matches your aims?

2. **Shell Energy are doing some research into local grids and the viability of a planned local hydro generation facility and want to use your energy data.**

Assumption that everyone wants 'good healthcare' but what's the extra value that they're getting from you? It's all about trust and motivations of the person receiving that data.

Who are you empowering? If they're still involved in fossil fuels, you're still empowering them! I'd want to look at the whole company, ie what taxes do they pay, how do they treat their staff etc.

Down to the ethical policies of the data coop. They would have to make it clear where they stand on even contemplating working with companies like this.

3. **A third party wants your contact details and energy usage to sell you energy saving devices.**

Depends on the 3rd party! We'd make stringent requirements on the company involved.

Anyone trying to sell me things, wouldn't want to be involved with sharing my data. I wouldn't want them contacting me.

Green Building Store - triple glazed windows and new doors etc, company I trust and they're recommended. So potentially share data once for a consultation, but not allowed to use data for any other purpose. Not for marketing ie wouldn't want them coming back to me in 3 months time saying 'i can see your energy usage has changed so you might like this…'

Because I'm doing stuff on my house at the moment, I'm interested in energy saving devices, i might do it now. Normally I'd say.

If they came to me, it'd be a red flag.

They wouldn't need your energy usage.

It needs to be something that the members of the coop ask for, rather than giving consent to the data coop to simply share.

Could the data coop accept referral fees? Discounts for the people they contacted?

It should be us getting in touch with people, not them getting in touch with us.

4. **A fellow coop member wants to compare energy usage of people with similar houses so that they can improve their energy efficiency.**

Instinct is to think you'd be ok if you knew them…
Lots of potentially sensitive data at risk.

Aggregated data would be happy with i.e. cited as an example of a sem detached house.

Hive do something similar, give monthly update showing comparison with other houses in your area, but it's meaningless.

Would want to know that they're accessing this data on a peer to peer level, but perhaps time limited. If it was some i know! If it was someone i dind't know, i wouldn't know their purpose

Would still want to time-limit access.

Individuals aren't tied to the same GDPR rules - "personal or domestic" activity excluded, so if you grant it to another individual you have lost control of it!

Would need the coop to involve us all in telling us what the results are.

GDPR: https://gdpr-info.eu/recitals/no-18/ (not sure if this is an official link), but from ICO: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/exemptions/ & from EU: https://ec.europa.eu/justice/article-29/documentation/other-document/files/2013/20130227_statement_dp_annex2_en.pdf (10pp).

Exploring different flows

Mapping the four data flows that the Coop is potentially in control of.

1. **Intra cooperative flow**

What are you using this data for?
Conditions for using data
Time period of access - could set by the user
Difficulty of sharing within the cooperative because hard to maintain anonymity

2. **Member to cooperative flow**

Is the flow already on?
Prerequisite for being part of the cooperative - if the member has already consented to be part of the coop is their consent to share data with the coop already explicit?
Trust between member and cooperative
Coop looks after data - but isn't shared
Coop custodianship of all data - so needs to be secure
Member coop permission to:
- Store and pool data
- Process data
    - Anonymisation
    - Aggregation
    - Analyse - under the constraints of legitimate purpose

- ○ Is it legitimate interest? - What do other data organisations do?
- ○
- Share data - based on consent

3. **Cooperative to Cooperative flow (ethically aligned organisations)**
   Can only be based on consent
   > What type of data can be shared

4. **Cooperative to 3rd party flow (anyone else)**
   > Can only be based on consent
5. **Open Data?**

Technological infrastructure would have to be robust enough to manage these consents

What level of intervention is needed.

1. **Consent**
   Granular consent everything on a case by case basis

2. **Persona based permissions systems**
   Preferences based on personas or archetypes

3. **Traffic light system** ie
- green - share data with anything the coop sees fit
- orange a) happy to share but want to choose on a case-by-case basis.
- Orange b) happy to share with members al the time, but 3rd party on case by case basis.
- red - only for the coop's own purposes.
-

Ie meet once a month, and coop governors / nominated data committee discuss requests.
Basci duty of care - Default position would be no personally identifiable data would be shared - on all levels
Exception - individual data subject would have to consent.

Workshop 5: How do we manage a data cooperative?

Thursday 19th November 2020, 6.30pm – 8pm

Link to slides:
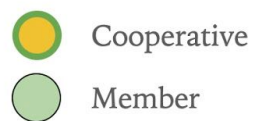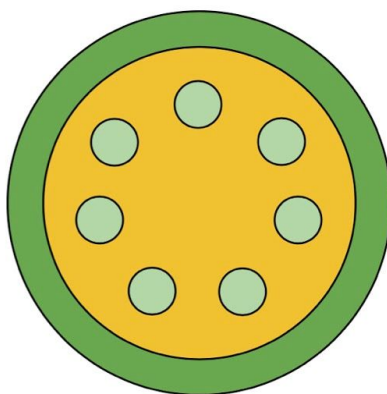
Roll call:
Sam
Julian
Richard
Ian

Agenda:

- ● Introductions
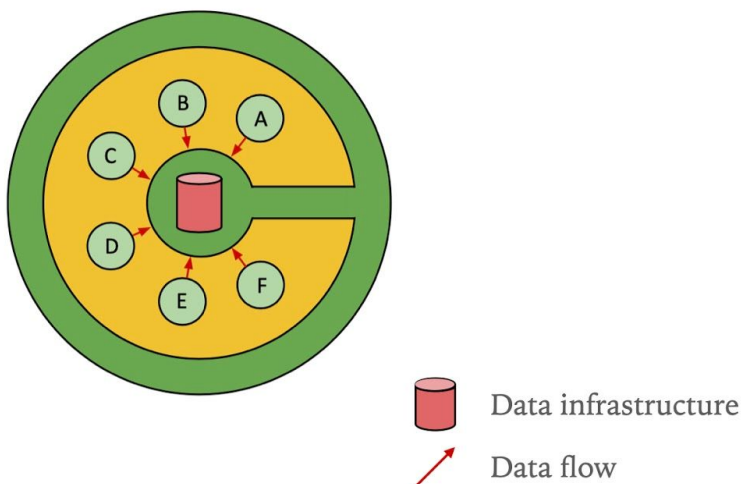- ● Project recap
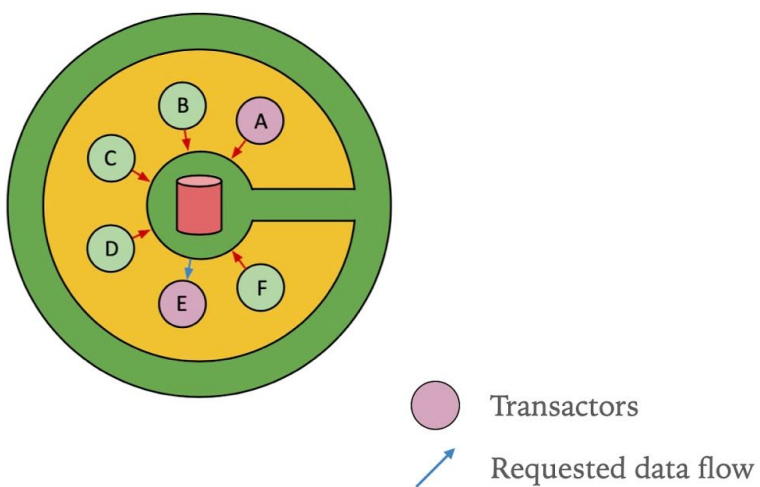- ● Data flows
- ● Consent Mechanism

Data Flows

Energy Cooperative

ENERGY COOPERATIVE

Flow 1: Member to Cooperative Flow

FLOW 1 – MEMBER ↔ COOPERATIVE FLOW

🔴 Data infrastructure

↗ Data flow

Flow 2 - Intra Cooperative Flow

FLOW 2 – INTRA COOPERATIVE FLOW

🟣 Transactors

↗ Requested data flow

Flow 3 - Inter Cooperative (Federated) Flow

FLOW 3 – INTER COOPERATIVE (FEDERATED) FLOW

Data cooperative

Federated data flow
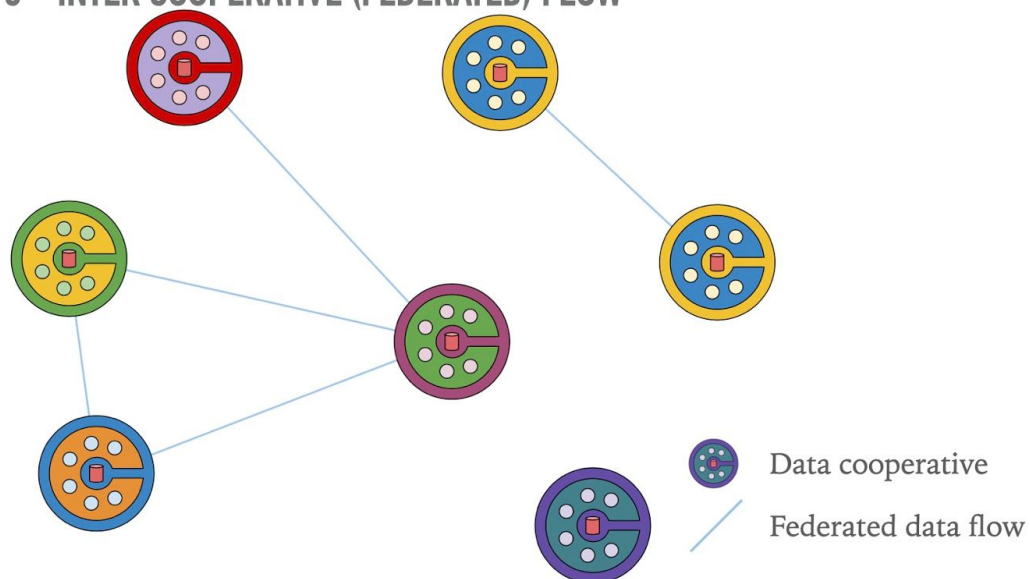
Flow 4 - Cooperative to 3rd Party Flow

FLOW 4 – COOPERATIVE ↔ 3RD PARTY FLOW

3rd party organisation

Licensed data flow
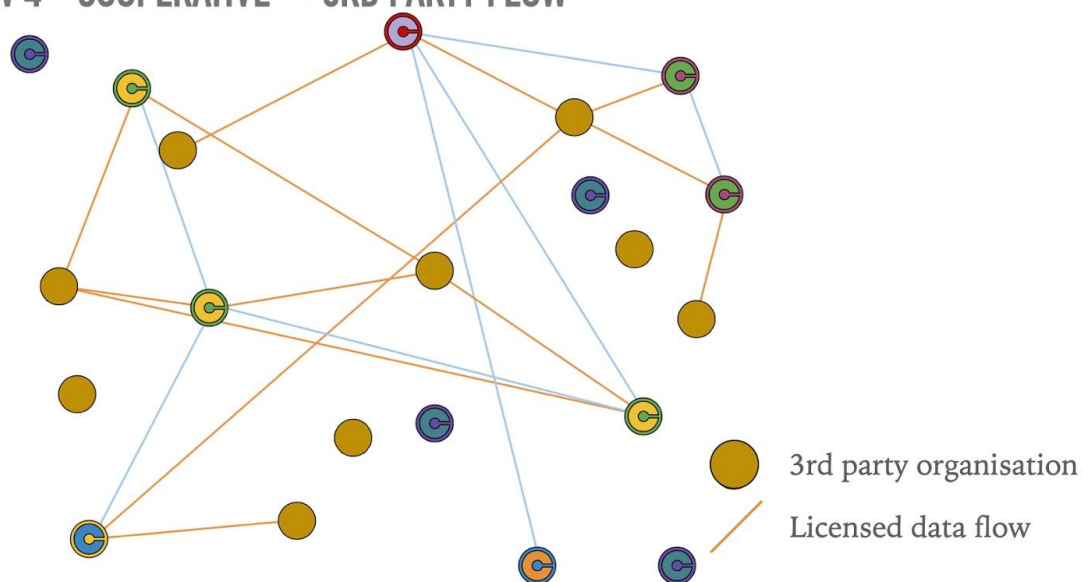
Flow 5 - Cooperative to Open Data (Commons) Flow
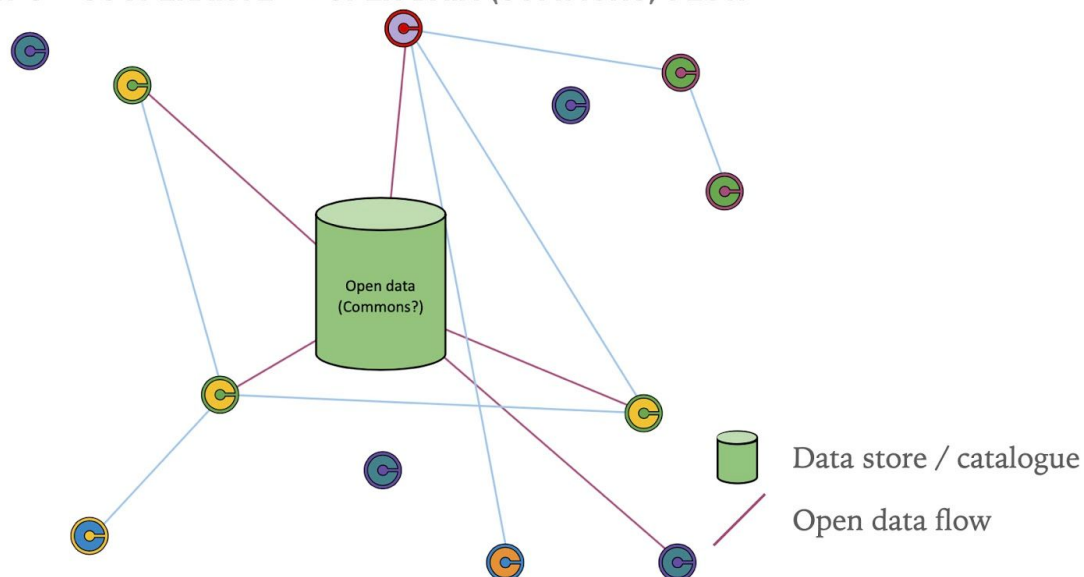
Discussion on data flows

- If the data was shared with an academic 3rd party who then wanted to turn it to open data, when/how would that consent be collected?
- Academic 3rd parties - could they publish the datasets alongside publications?
- Having a set of pre-configured consents would allow a co-op to know what it could share with a third party/academia. "We have 73 members who are happy to share anonymised data with third parties and to have it published in aggregate on an open basis".
- "It's the agreement to an outcome or purpose of sharing that feels important in gaining consent. So really important that those are clear." (from zoom chat)
- Medical data co-ops:
    - Savvy Co-op: https://www.savvy.coop/
    - Midata: https://www.midata.coop/
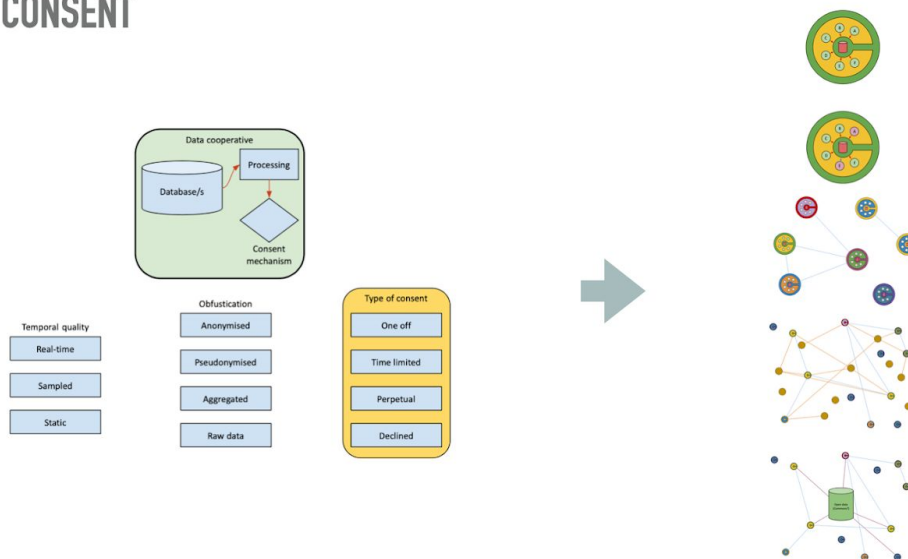
Consent Mechanisms

Granula Consent

GRANULAR CONSENT

- Similar to website cookie preferences used on websites
- Every type of data is given explicit consent to be shared
- Choices are highly customisable
- Consent can be withdrawn at any time

GRANULAR CONSENT

Personal / Archetype based permissions

PERSONA / ARCHETYPE BASED PERMISSIONS

- Proxy based system where consents are based on the characteristics of a fictional persona or archetype
- The persona/archetype may have ethical values that are a subset of the cooperative's
- Users identify with the persona/archetype that matches their own and accept their recommendations
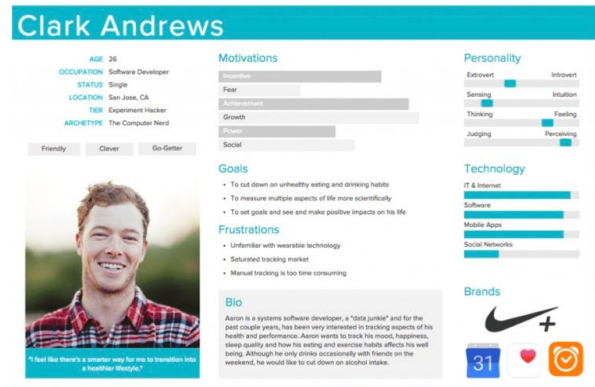
Not like the above

PERSONA / ARCHETYPE BASED PERMISSIONS

Persona A

Persona B

Persona C

Traffic Light Consent

TRAFFIC LIGHT CONSENT

- **Green - share data with anyone the coop sees fit**
- **Amber - happy to share but choices made on a case by case basis**
- **Red - data is used for the cooperative's own purpose**

🟢 Permissive

🟡 Controlled

🔴 Restricted

TRAFFIC LIGHT CONSENT

🟢 Permissive

🟡 Controlled

🔴 Restricted

Discussion

Group 1

Questions: (answer for each mechanism: Granular, Persona and Traffic Light - 10 minutes on each)
1. What are the pros and cons?
2. Would you be happy sharing your data via this mechanism? Why/why not?
3. What would you need from the coop to trust that your data is being looked after in this mechanisms?

I think it is sort of inevitable to end up in granular
Benefits of mixing the persona and traffic lights
Journey can help made simpler than the decision making actually is. Reducing friction
Conditional and unconditional trust
People would join a data cooperative as they would want to share data
Stuff that is almost totally anonymous such as temperature outside our homes
PV is almost anonymous, doesn't tell you when you are in or out
Enter into a commercial relationship with a business and not another type.
Persona would enable different types of sharing and reduce friction
Building trust - looking at cookies and make assumption that if they have everything on - go somewhere else
Persona ethos may pigeonhole people. Some people would want to know.
Disruption of third party flows in small coops
Signing up gives a whole heap of permissions.
Obligation on the cooperative to make sure everything is understandable and that people can make.
Paper prototyping to create a process that would enable diverse groups to make decisions
Drill down to things that a generally anonymous
Nervous about the location of house, but there are ways to anonymise this.
External audit of the decisions made each year

Group 2

Questions: (answer for each mechanism: Granular, Persona and Traffic Light - 10 minutes on each)

GRANULAR

4. What are the pros and cons?
    a. Con: is very complicated - granular consent has many dimensions to it, impractical to use this alone & to control all possibilities.
    b. Coop needs to set the minimum?
    c. Pro: Most general solution, most flexibility
    d. On case by case basis just isn't practical
    e. Can't reasonably ask every member about every new partnership
    f. Gives you full control. Pro for the users. But would complicate for the coop itself.
    g. Withdrawal - would you expect data to removed all on those bases? Is it even possible once data has been aggregated elsewhere etc?
        i. No
    h. In research if people withdraw we usually state that their existing shared anonymised data may not be retracted (this is stated when they provide informed consent at the start of the study)

5. Would you be happy sharing your data via this mechanism? Why/why not?
    a. Yes, good for the security-conscious (and time- and patience-rich) user who can deal with detailed controls. (The sort of people who read EULAs) (i.e. like me)
    b. I would like the option, but in practice I would use it rarely because it would take too much time to configure everything. I would not like to be forced to think about granular consents for everything all the time (see also: having to decide about cookies for every website every time I go there; it's an obstacle to what I'm actually trying to do when visiting a website)

6. What would you need from the coop to trust that your data is being looked after in this mechanisms?
    a. Periodic check-up from the coop, summarising your consent choices?
    b. Aggregated data as a baseline

PERSONA

7. What are the pros and cons?
    a. Simplified in several dimensions - choosing from a menu of options a lot easier
    b. Greater variety of personalities
    c. Pro: should be quick to find a persona with like-minded values - Con: as long as they are well designed personas and they find one they fit within (could maybe be able to select 2+ and apply permissions from each)

       d.   Con: might be hard to effectively communicate what each persona means to a member who's having to choose between them (who is 'Andrew Clark' and am I like him??)

       e.   Pro: set of personas can be tailored to the expected consent patterns of members in this particular coop

8. Would you be happy sharing your data via this mechanism? Why/why not?

       a.   Yes, provided there was a good range, it would still allow me to protect what I most cared about - ppl would err on the side of caution (or slightly more abandon potentially) and fall back (/jump up) to the nearest persona.

9. What would you need from the coop to trust that your data is being looked after in this mechanisms?

       a.   Personas don't change over time (or, at least, once I've chosen a persona, it won't change under my feet)

       b.   Periodic check-up from the coop, summarising your consent choices?

       c.   Fully transparent details of what each persona means

TRAFFIC

10. What are the pros and cons?

       a.   Simplified in several dimensions

       b.   Could be more than 3 levels?

       c.   Con: Would this be more complex to manage from the Co-op's perspective? (for amber)

11. Would you be happy sharing your data via this mechanism? Why/why not?

       a.   Bit too coarse-grained for me

12. What would you need from the coop to trust that your data is being looked after in this mechanisms?

       a.   Periodic check-up from the coop, summarising your consent choices? Or who your data has been shared with recently

A regular statement or summary of want consent you've given, what data has been shared, to whom and what purpose

      That's likely to be hard with large sets of time-series data. Perhaps the report could just include consents given, and the data flows which were in operation during the statement period (e.g. "during the last 3 months, you agreed to share your meter data with other coops; during the last 3 months we had data sharing partnerships with coop A, coop B, and coop C; your data may have been shared with zero or more of them")

Other notes

-   What about 'withdraw consent' as part of the mechanism process?

- Discussed a bit in 'GRANULAR' section above
- Persona and Traffic light are almost 'on top' of granular - you still then have the option to go granular - personalise your persona!
- Whichever one is selected could perhaps be trialled and reviewed?

Workshop 6: How do we manage a data cooperative? Part 2

Thursday 21 2021, 6.30pm – 8pm

Link to slides: n/a

Roll call:

Agenda:

- ● Introductions
- ● Project recap
- ● The Data Governance Mechanism
- ● Questions for discussion
- ● Next steps

The Data Governance Mechanism

Link to Miro board: https://miro.com/app/board/o9J_IXnTADM=/

Group 1 thoughts / comments

1. As a prospective member of a data coop, would you be happy if your data was shared with 3rd parties?
2. If not, why not? / What's missing?

Group 2 thoughts / comments

1. As a prospective member of a data coop, would you be happy if your data was shared with 3rd parties?
2. If not, why not? / What's missing?

Not clear whether this process is something that could be automated, or whether this is a process that would be followed by a human administrator - if this were being implemented as a business process you'd do it all on paper/ad hoc a few times and then work out what would be beneficial to automate. A huge software project otherwise, would require massive capital investment to develop!

Does the Ethical Assessment (4) need to be where it is? Could it go earlier?

Fundamentally, should (/can) we have algorithms/software making ethical decisions? We are uneasy/opposed to this as a suggestion! We would have confidence in software to make very qualitative judgements on these 'scorings'. But definitely a role for software to handle requests & ensure that all data required in captured. And the preferences database seems like it would only be practical through software. Creating a typology of purposes would be a major undertaking.

Easy case vs more difficult cases - a filtering function
What we're saying here is quite complex, how would you describe this to people ie party 1, who aren't experts? If we can get to that simple state, then can help us with what we need to make those more difficult cases

What is the value of the data that you are releasing?
Are you being paid for it? Quid pro quo? - perhaps part of the value assessment?
Where is money? What will pay for this, this is potentially an expensive piece

Smart Energy Research Lab - UKDS, UCL etc: https://serl.ac.uk/ - could provide a framework to learn from - also how is it funded? Is there an argument that if academic institutions can run something like this alongside Government that community led groups could be granted funding to do similar with more action/community benefit focus

There would need to be some sort of contractual agreement as well, presumably? Is there a stage missing where the requesting party enters into some sort of contract or agreement.

Questions for discussion

1. **Mechanism for scoring**: If we're automating this process, there would have to be a score that comes with each assessment. When there is a threshold reached, then it allows data to flow.

   The reason there is a scoring mechanism is that there are different flavours of the same data ie energy monitoring that is very granular or sampled or static.

   How might we do this?

2. **Creating typologies:**

How might we do this? Is there anything out there we could draw upon? What potential issues might there be?

3. **What other questions do you have?**

Next steps

- Further work on the mechanism
- Potential extra workshop
- Report for the ODI
- Event on 30 March, details TBC.