

Data Sharing & Data Trusts

Workshop @ Data for Policy, 10th June 2019

Schedule

- Session 1: Why is data sharing important? 13:30 – 14:30
- Session 2: Data sharing and trust 14:30 – 15:50
 - Break 15:00 – 15:20
- Session 3: Increasing access to data while retaining trust 15:50 – 16:50

Goals

1. Recognise the benefits of and obstacles to data sharing
2. Define why you may or may not want to share data
3. If you do want to share data, develop the big picture of how to go about it

Session 1: Why is data sharing important?

Johanna Walker, University of Southampton

Why is Data Sharing Important?

Johanna Walker

University of Southampton

Data for Policy, London, 10 June 2019

Johanna Walker

- PhD: Open data, data sharing and innovation
- Data sharing WP leader Data Pitch (H2020)
- Data protection by design: Building the foundations of trustworthy data sharing – Data for Policy, London, June 10-12
- Personalised user engagement with TV: a case study – Critical Perspectives on Data Governance, Southampton, June 29
- A framework for data sharing for open innovation - Open and User Innovation, Utrecht, July 8-12
- Decision-making processes in data trusts - WomENcourage, Rome, September 16-18

Data Sharing Economy

“If favourable policy and legislative conditions are put in place in time and investments in ICT are encouraged, the value of the European data economy may increase to €739 billion by 2020, representing 4% of the overall EU GDP.” [Building an European Data Economy]

1. Promoting the re-use of public and publicly funded data

Open Data and Public Sector Information Directive (April 2019)

2. Assessing the need for further action on access to and re-use of private sector data

Key principles to be considered so as to make data sharing a success for all parties involved, in B2B and B2G situations

3. Removing data localisation restrictions: the free flow of data

The Regulation on the free flow of non-personal data May 28th (non-personal equivalent of GDPR)

DP Direct Economic Impact

28 data challenges
(total)

17 organisations sharing data
(total)

€4.6M invested
(total)

€6M financial impact
(17/18)

229 applications
(total)

47 companies
(total)

14 countries
(total)

69 jobs created
(17/18)

Data Sharing

Closed platforms can be set up by core players in a data sharing environment as well as by independent intermediaries;

Open platforms target a wide and unknown range of participants, they become “data marketplaces”, just as Ebay is a marketplace for tangible goods;

Between entirely “open” and “closed” platforms are those which enable sharing between specific players, while the “data sharing club” is open to new entrants if they fulfill certain requirements.

The Data Sharing Economy: On the Emergence of New Intermediaries

Richter, H. & Slowinski, P.R. IIC (2019) 50: 4.

<https://doi.org/10.1007/s40319-018-00777-7>

Drivers and barriers

DDI key enabler of growth and jobs (AI)

Internet of Things

Non-rivalrous

Data ‘owned’ by the entity that holds it

Reluctance

Challenges to solve

Incentives

Broader than industry sectors

Conditions

Match supply and demand

Trust

Ecosystem examples

- Data Sharing Incubators and Accelerators: eg Data Pitch, European Data Incubator
- International Data Spaces: eg Fraunhofer, Deutsche Telekom
- BDVA-PPP I-spaces - host Closed as well as Open Data from Business and Public sources for data driven innovation
- ODI data trust pilots



Data Pitch – a data sharing case study

Johanna Walker, University of Southampton, UK
@ms_j_walker

London, June 10th 2019, Data for Policy

From open data to the free flow of data



Data Pitch context

- The digital revolution is built on data
- Removing obstacles to data mobility could add 4% to the EU GDP by 2020 (Deloitte)
- Data Pitch delivers a programme that showcases the benefits of open, data-driven innovation
- We bring together established businesses and startups/SMEs to meet challenges with shared data



[Source: Deloitte]

The Data Pitch project

A European open, data-driven innovation programme

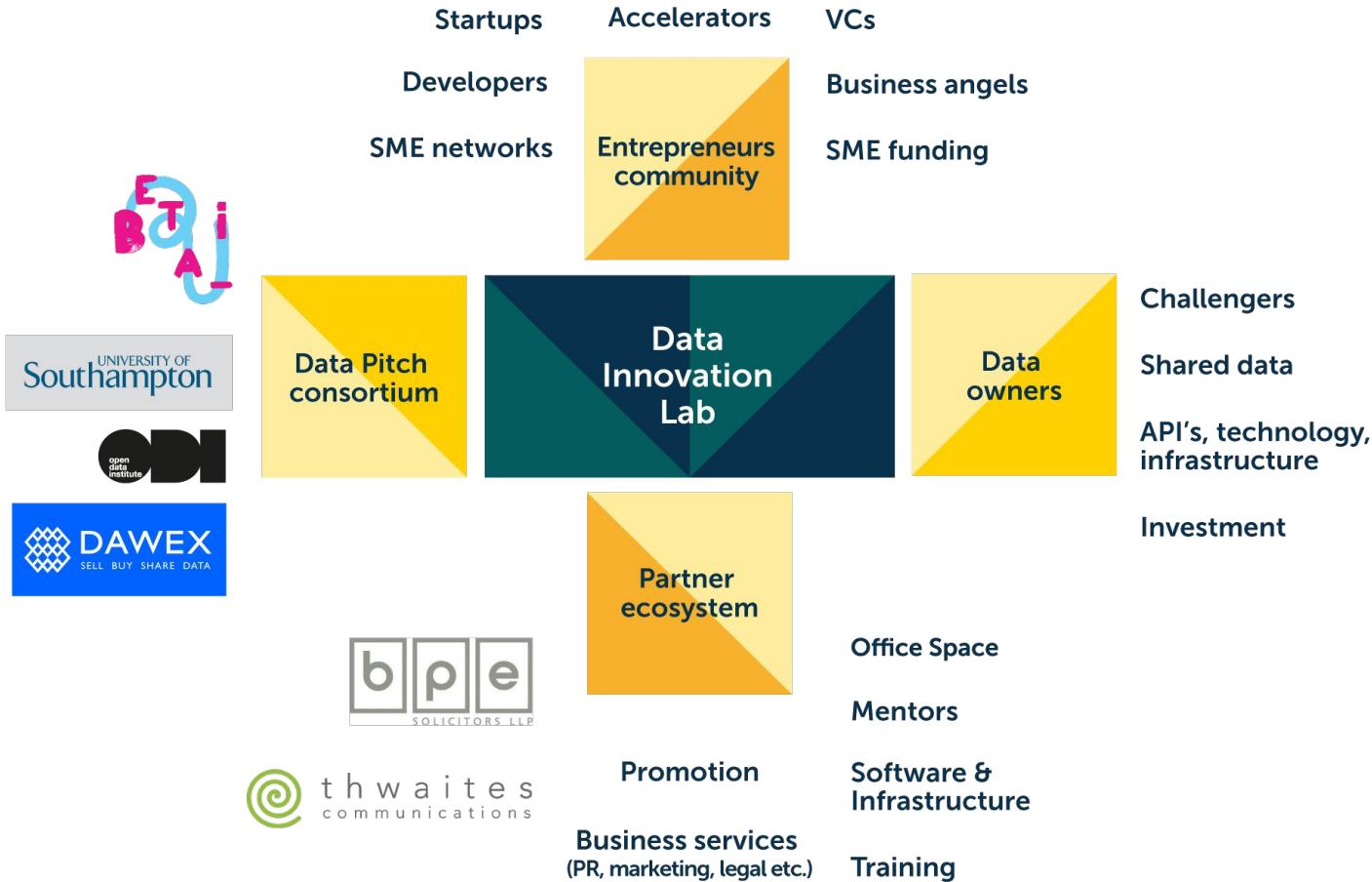


- Consortium of four organisations from UK, France, Portugal
- Delivers an open innovation programme centred on data shared across organisations
- Funded by the European Union's Horizon 2020 Research and Innovation Programme
- 3 years: 01/2017 – 12/2019
- Budget: 2.75M consortium; 4.89M SME fund



A European data ecosystem

Innovators, entrepreneurs, data providers, enablers



Our innovation process

Broad engagement, transparent & fair evaluation, real data challenges, startup support



1st
round

1st July
2017

1st October
2017

1st February
2018

This week
😊

Launched July 2018

Our value proposition

Tailored for strategic groups of stakeholders



Data Sharing

Funding

Promotion & networking

Datathons

Network of European SMEs

Data-driven business models

Business & technical support

Cross-marketing

Network of European SMEs

New collaborators

Promotion & networking

Tracks and challenges

Three tracks, see datapitch.eu/challenges



1. Data providers

- SME solves DP's problem using DP's data
- IP stays with SME. SMEs grants DP a non-exclusive license
- Data Pitch manages relationship with DP
- Contract template available

2. Sectoral

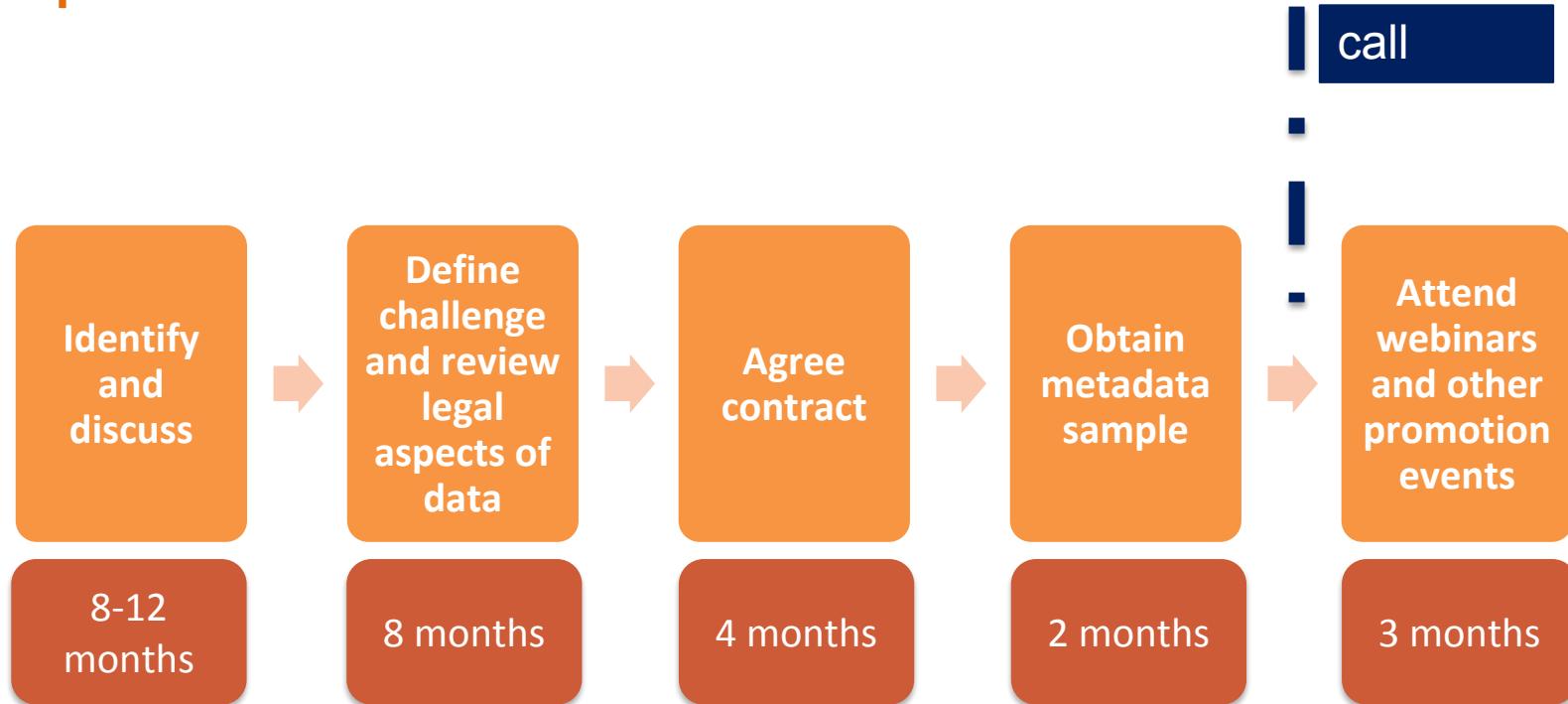
- SME solves sectoral problem using data of their DP
- IP stays with SME
- SME manages the relationship with DP
- DP requires proof of access to data and carries out legal and privacy checks

3. Open

- Only for truly transformative ideas.

Defining data provider challenges

Covers technical, legal, innovation and marketing aspects



Greiner Packaging International

1



SMART MANUFACTURING

Harnessing IoT data for tomorrow's smart factories

Challenge identifier: DPC3-2018

Proposed by

Greiner Packaging International GmbH (GPI). GPI manufacture and markets plastic packaging solutions for food and nonfood industries.

Greiner Packaging International

2



Background

Description

Data

Expected outcomes

Expected impact

Description

GPI has recently invested in extensive sensors across 3 manufacturing plants. We are now looking for ways to utilise this data, along with our existing data, to best support and enhance our business. We are looking to develop applications which span across traditional boundaries of manufacturing, logistics and supply chain (perhaps even sales), providing data-informed solutions to better coordinate these processes, make them more efficient, and platforms supporting these new processes.

We are particularly interested in solutions that:

- Define new relationships between data to provide new insights and understanding;
- Discover new business opportunities and improve production efficiency;
- Help ensure the integrity of interactions within supply chains;
- Integrate data of different modalities (sensors, acoustic data, historical records, thermal maps) to produce useful products and services;
- Use data to predict maintenance needs and create more efficient servicing and repair services; and
- Predict and help optimise consumption and stock level, including in multi-country operational scenarios
- Capture and interpret data to produce answers to commonly asked questions and reduce human intervention;
- Are able to integrate poor, inconsistent or fuzzy data or information and provide interfaces that communicate key findings and effectively engage users.

Data

Examples of data include but are not limited to:

- Production orders
- Logistics (order process)
- Sensor data in production (machine, energy consumption, cooling water)
- Environmental data (shop floor temperature/humidity)
- Quality data (product properties, scrap rate)
- Failure cases of machines
- Maintenance and usage history

More detailed information about the data can be found in our data [catalogue](#).

Greiner Packaging International



3

- Data Provider in the second call of Data Pitch;
- The company has a position dedicated to advance data driven innovation;
- Use the sensor data from three of their manufacturing plants to develop solutions that enhance the business in terms of manufacturing, logistics, supply chains and sales;
- Breadth of opportunity/Capacity to fully commit to the innovation process;
- Matched with five innovators who are now working on advancing different areas of their business.

Greiner Packaging International



idatase

Country: Germany
Challenge: Smart Manufacturing

Reducing the risk of human error to reduce costs and to ensure more IoT initiatives succeed



OBUU Tech
Country: Spain
Challenge: Smart Manufacturing

A customisable solution for stock provisioning optimisation in aeronautics, infrastructure and railway maintenance organisations



Informed Actions
Country: UK
Challenge: Smart Manufacturing

A software solution for social housing organisations to better manage resources by predicting problems such as fuel poverty, damp and mould, and tenancy churn



Mammoth Analytics Limited
Country: UK
Challenge: Smart Manufacturing

Empowering non-technical users with the tools they need to easily analyse raw data and make better business decisions



Renvis
Country: Greece
Challenge: Smart Manufacturing

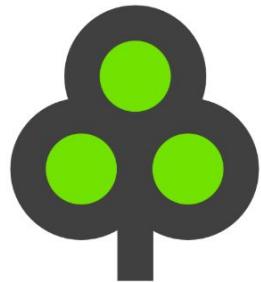
A smart manufacturing solution that gathers real-time data, maps decision making and suggests actions

Cohort 1 startups

142 applications, 11 countries, 8 challenges



predrepair_



Frosha



mojo



RECOGNAI



Data
Moove
data experience



RADIOBOTICS



NextQuestion



transformative



pharmawizard



A.i. Poli

Discuss!

- What data would you want to share?
- In which role?
- What would you need to enable this?

Session 2: Data sharing and trust

Kieron O'Hara, University of Southampton

Jack Springman, Ctrl-Shift

Data Sharing and Trust.

Kieron O'Hara
10 June 2019

The Problem of Trust

- NOT to increase trust



... Is To Align Trust and Trustworthiness

- Trustworthiness => delivery of commitments
 - A virtue
- Trust => belief in trustworthiness/reliability
 - An attitude
- Trust w/o trustworthiness => fraud
- Trustworthiness w/o trust => opportunity costs
- Can we trust all and only trustworthy people?
- Far more discussion of trust than trustworthiness

Trust Gaps

- Trust = confidence in the trustee
 - Communication a prerequisite
 - What is the trustee committed to?
 - What do each of the parties think she is committed to? (And do they agree?)
- Trust = confidence in the system
 - What will happen under various contingencies?
- Uncertainty in data sharing
 - Liability
 - Reputation
 - Ownership/control

Stakeholders

- Data subjects (if any)
 - Will my data be misused? Will my privacy be compromised? How can I hold third parties accountable? What benefits might I gain? Will I be affected by non-personal data (e.g. profiles)?
- Data owners
 - Will I lose commercial advantage? Am I liable for actions of third parties? Am I liable for low quality? Do I have consent to share? Have I anonymised effectively? How can I get return on my data acquisition investment? Whose IP?
- Data consumers
 - What is the provenance of the data? Can I gauge its quality? Is consent in order? Can I reidentify data subjects from anonymised data? Can I integrate datasets successfully? Whose IP?
- Sectoral players
 - How is my reputation affected by actions out of my control (cf. Cambridge Analytica)?
- Society
 - How can social value be realised? Will data sharing help entrench commercial oligopoly/monopoly profits?

Result?



<https://www.allbusiness.com/asset/2015/09/business-risk-concept.jpg>

Ethical Data Stewardship

- Regulation is not enough
- Foster human flourishing (BA/RS data governance report)
 - Practical wisdom, not box-ticking
 - Sensitivity to context
 - Map the trust gaps in the specific context
- Virtuous data stewardship
 - Care
 - Prudence
 - Respect
 - Trustworthiness
 - Conscientiousness

The Role of Institutions

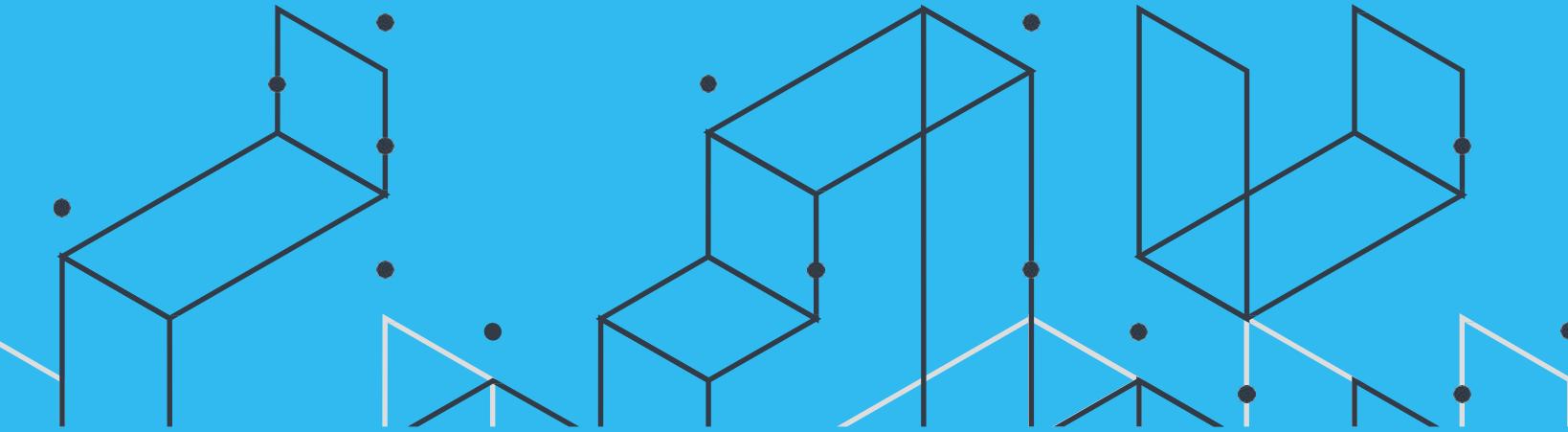
- Institutions for trust – certifying trustworthiness or reliability
- Economies of scale
- Loci for communication and engagement
- Manage expectations
- Define the commitments of trustworthiness
- The role of a data trust
 - Establish what trustworthy data stewardship consists in

Disclaimer

- Texts, marks, logos, names, graphics, images, photographs, illustrations, artwork, audio clips, video clips, and software copyrighted by their respective owners are used on these slides for non-commercial, educational and personal purposes only. Use of any copyrighted material is not authorized without the written consent of the copyright holder. Every effort has been made to respect the copyrights of other parties. If you believe that your copyright has been misused, please direct your correspondence to: kmoh@soton.ac.uk stating your position and I shall endeavour to correct any misuse as early as possible.

Report Available

on 17th June at www.ctrl-shift.co.uk



CtrlShift

Break!

Please be back by 15:20

Discuss!

- Who needs to trust the data you want to share?
- What are the risks if this trust is lacking?
- What do you need to build and maintain this trust?

Session 3: Increasing access to data while retaining trust

Peter Wells & Jack Hardinges, Open Data Institute

The different approaches to increasing access⁺ to data while retaining trust



Jack Hardinges, Peter Wells

@jhardinges1 @peterkwells
theODI.org



Contents:

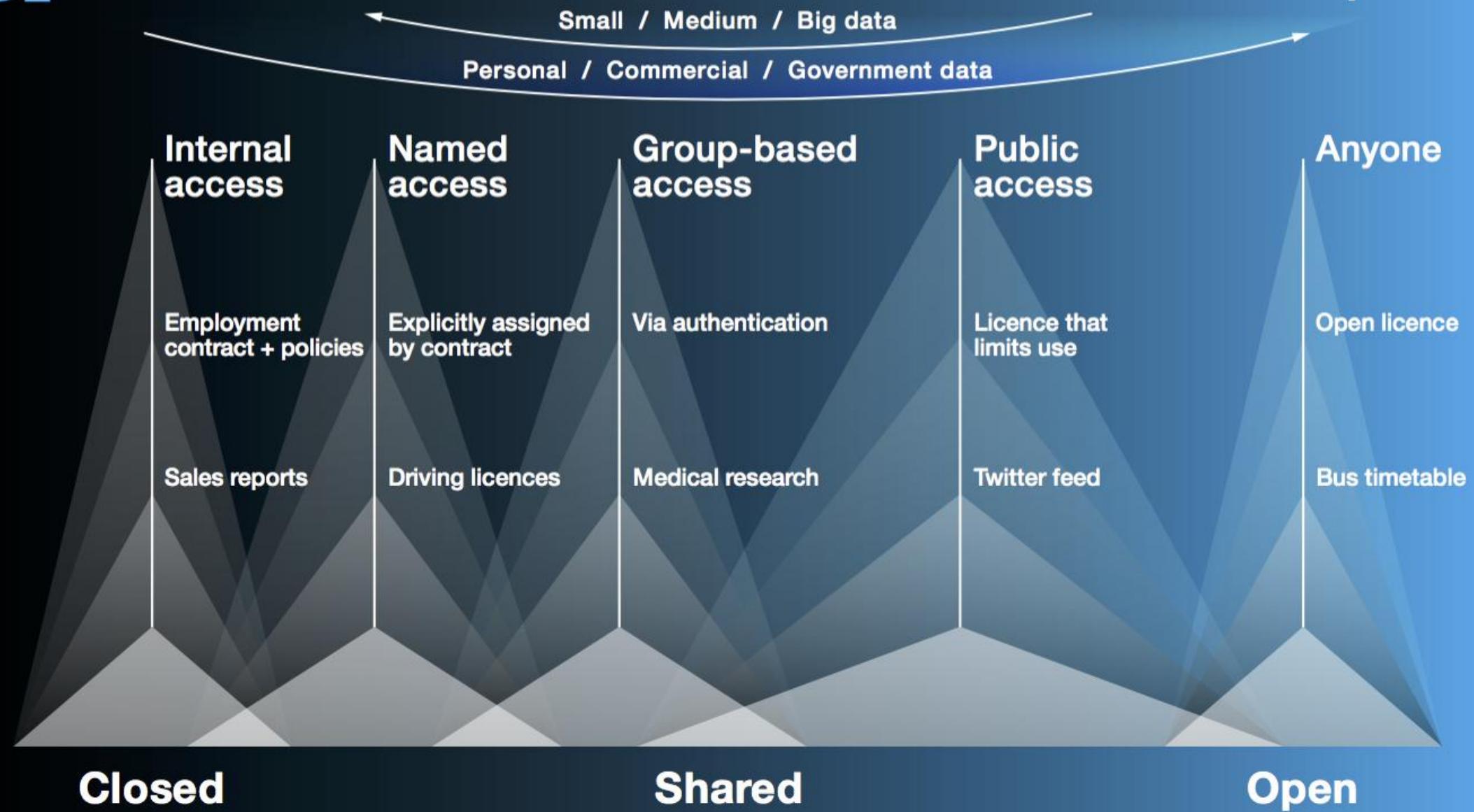
1. Increasing access to data while retaining trust (PW, 20 mins)
2. Case study - what is a data trust and what has the ODI learnt from studying them (JH, 10 mins)
3. Discussion (All, 20 mins)

Our mission

**We work with
companies and
governments to
build an open,
trustworthy data
ecosystem.**

Our vision

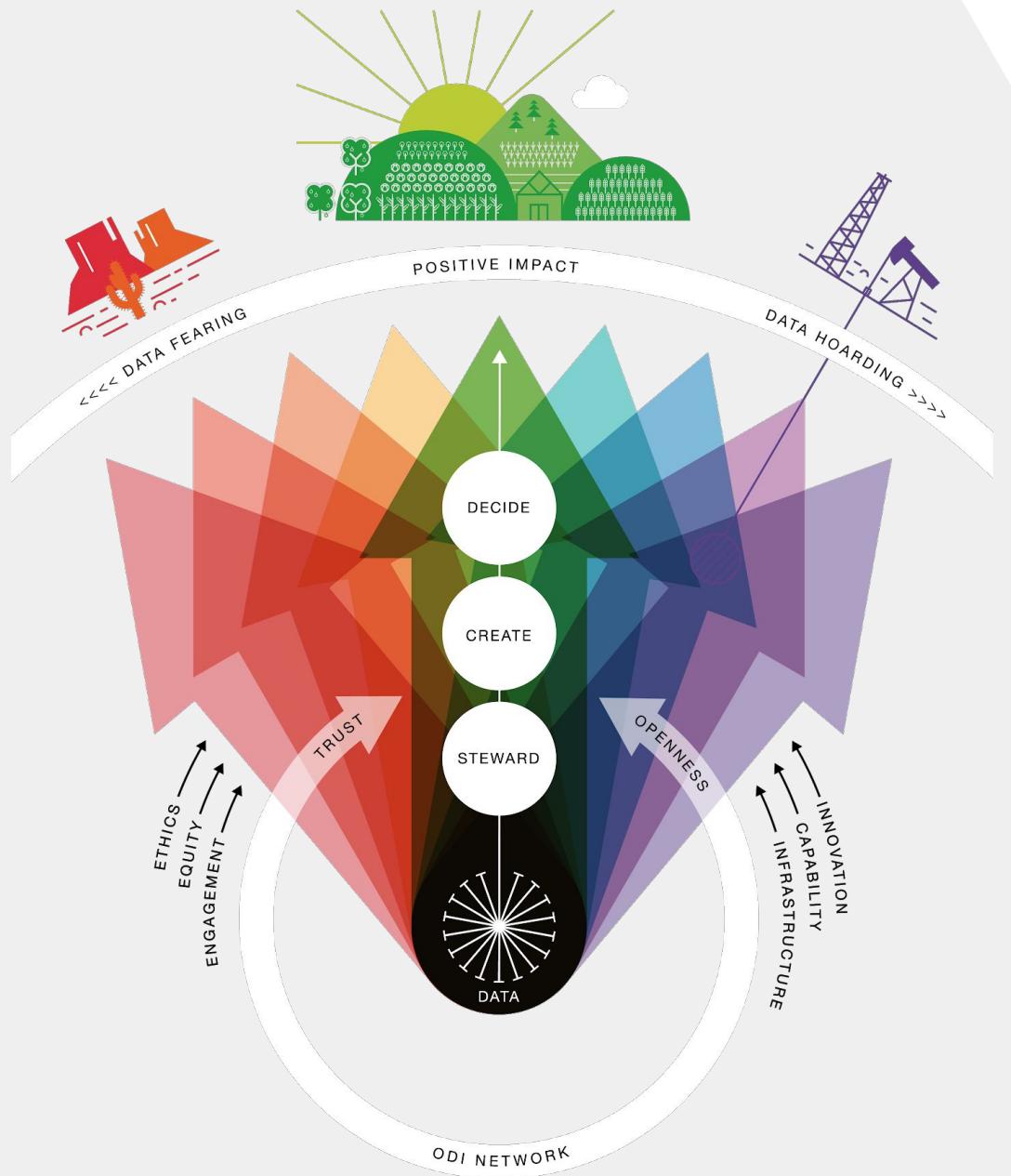
**We want a world where
data works for everyone.**



An aerial photograph of a multi-lane highway interchange. The roads are dark grey asphalt with white dashed lines. The interchange features several ramps and a central circular overpass. The surrounding area is densely populated with green trees and bushes, creating a pattern of green and grey. The perspective is from above, looking down at the complex network of roads.

data is
infrastructure

We need to increase access to data while retaining trust



Increasing access to data + • •

Research question

Which data access models give people increased access to data while retaining trust?

Project goals

- ▶ To understand existing and new data access models
- ▶ Bring together data stewards and enable them to share learnings/collaborate

User research

Discovery

Control + access

Trust

Opposition to transparency

Business cases

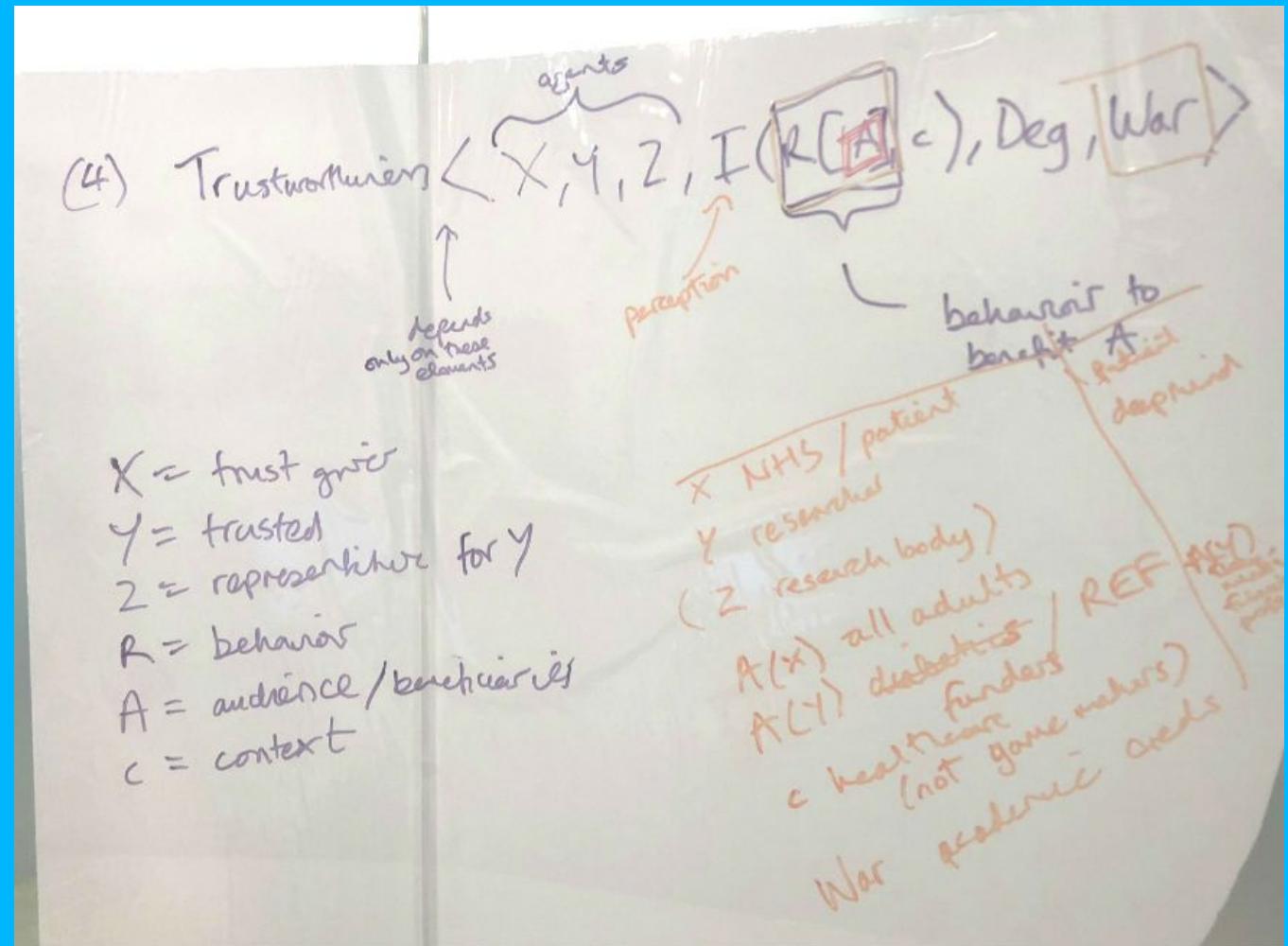
Personal data



Trust

A general definition of trust

(Kieron O'Hara, 2012)



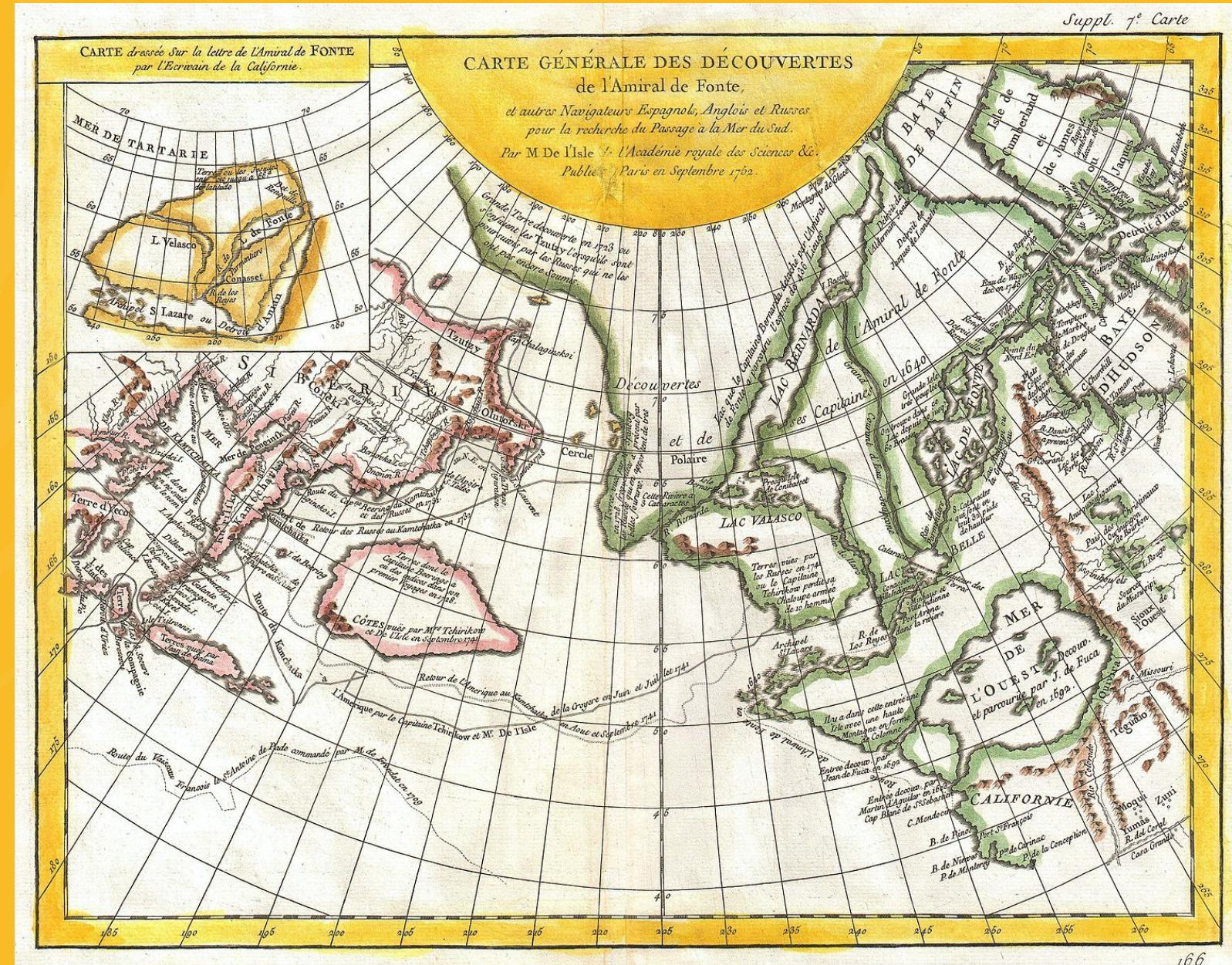
Taxonomy

Better
Understanding the
Responsibilities and
Rationales
Informing
Trustworthy
Options
for Sharing"



Eric Molina, "hands down the best california burrito i've ever had"

Travel



Data access map

There are a number of approaches.

We need to make it easier for data explorers.



Some areas
of the map
are
well-explored



Collabrador island

Some need
more
exploration



A federation of data explorers?

UoS, Datapitch,
Govlab, Sean/Keith,
Sylvie/Neil, data
stewards, data
requestors, big firms,
WEF, DG Cnect,
governments
&c &c &c



Contents:

1. Increasing access to data while retaining trust (PW, 20 mins)
2. **Case study - what is a data trust and what has the ODI learnt from study them (JH, 10 mins)**
3. Discussion (All, 20 mins)



Independent report

Recommendations of the review

Published 15 October 2017

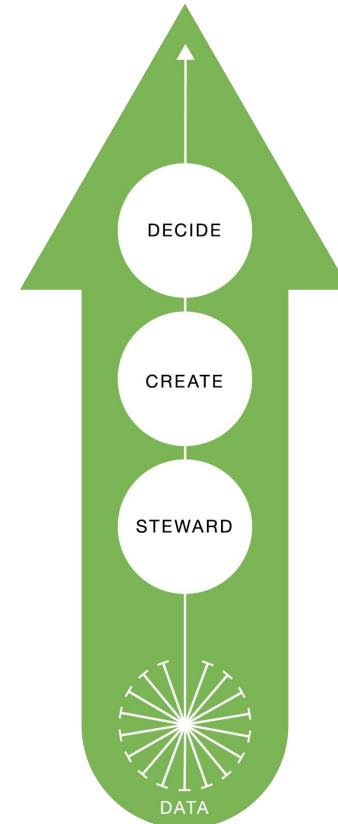
Recommendations to improve access to data

1. To facilitate the sharing of data between organisations holding data and organisations looking to use data to develop AI, Government and industry should deliver a programme to develop Data Trusts – proven and trusted frameworks and agreements – to ensure exchanges are secure and mutually beneficial.

- A data trust as a repeatable framework of terms and mechanisms.
- A data trust as a mutual organisation.
- A data trust as a legal structure.
- A data trust as a store of data.
- A data trust as public oversight of data access.

Data stewards

Decide who has access to data, under what conditions and who can benefit from it.



What is a data trust?

A legal structure that provides independent stewardship of data.

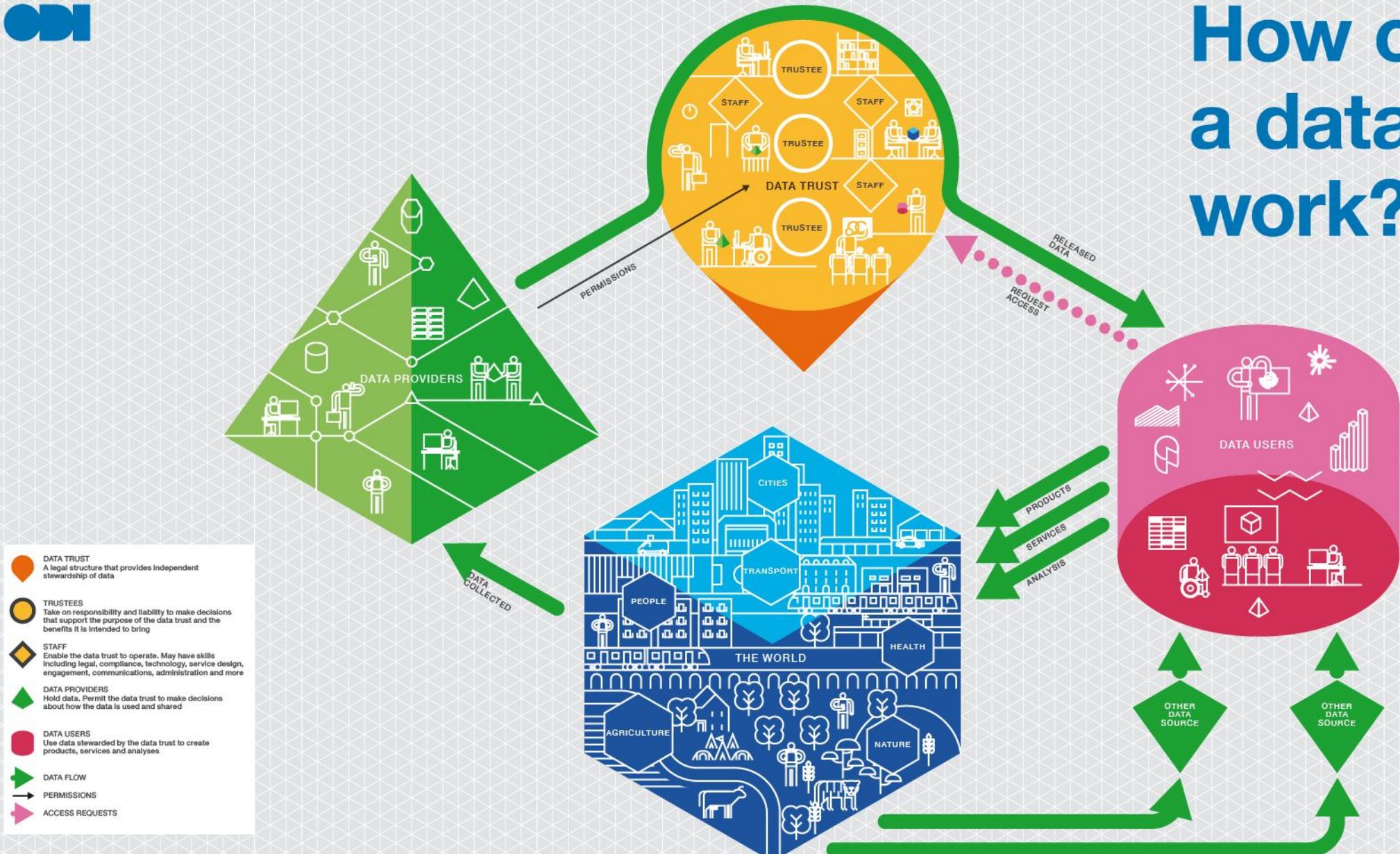
With data trusts, the organisations that collect and hold data permit an **independent** institution to make decisions about how that data is used and shared for an agreed purpose.

The **trustees** of the data trust take on responsibility for how data is used and shared and take on some liabilities. They must ensure these decisions support the purpose of the data trust and the benefits it is intended to bring.

While data trusts cannot take the form of ‘trusts’ in a legal sense, they use **legal structures** and forms that take their inspiration from them.



How could a data trust work?



Related examples

- Genomics England Access Committee.
- Office for National Statistics Secure Research Service.
- Administrative Data Research Network / Administrative Data Research Partnership.
- NHS Health Research Authority Confidentiality Advisory Group.
- ‘Safe Havens’.
- various clinical trial data sharing platforms, including Clinical Study Data Request and YODA.
- METADAC.
- ICES Data Repository.
- ...

Pilot projects to understand:

1. The types of challenges that data trusts could be used to address, as well as their limitations
2. What ‘a legal structure that provides independent stewardship of data’ looks like in practice
3. How data trusts could be built and the process made more repeatable

Pilots



Civic
Data about electric vehicle parking spaces and data collected by heating sensors in residential housing



Food waste
Food waste and sales data



Illegal wildlife trade
Image and acoustic data

Pilot methodology

A multi-disciplinary team undertaking (over three months):

- user research
- legal analysis
- designing decision-making processes
- technical architecture assessments
- data governance
- assessing the viability of implementing data trusts



**What challenges do
organisations face in
increasing access to
data?**

Challenges to increasing access to data

- A lack of evidence of the business case for sharing data
- Concerns about the reputational impact of providing access to the data
- A lack of resources to share or make wider use of the data
- Confusion around ‘data ownership’, rights and control
- A lack of standardisation
- Limited data literacy and skills



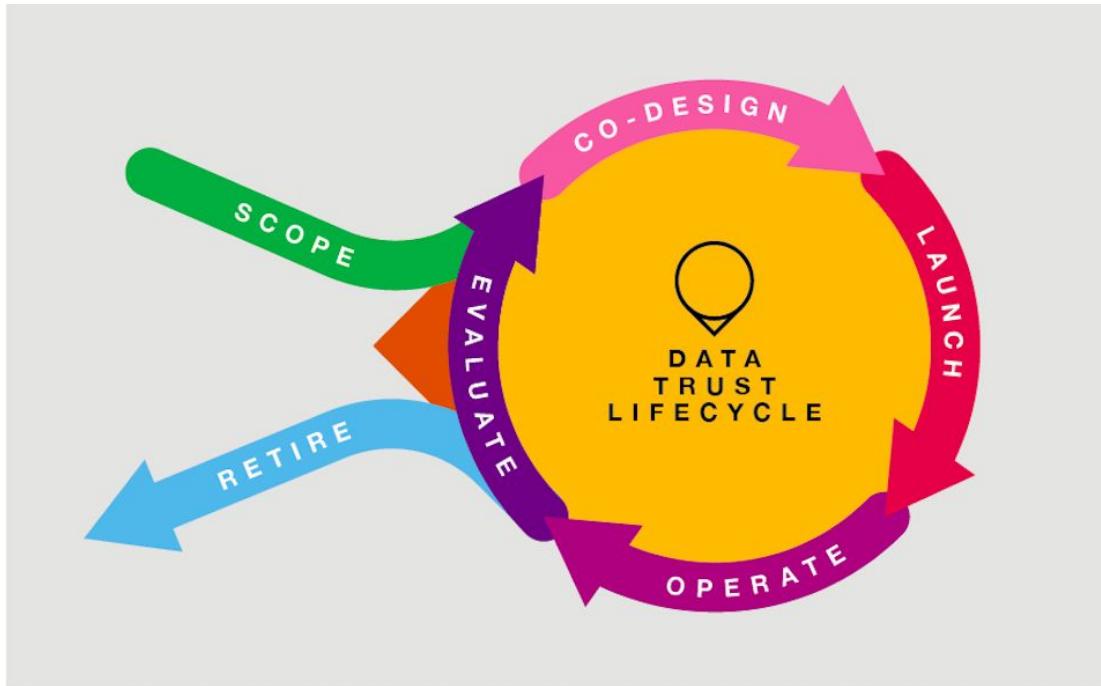
What challenges do
organisations face in
building data trusts?

Challenges organisations face in building data trusts

- Confusion related to the term ‘data trust’
- Making comparisons between data trusts and other approaches
- Determining independence
- Adopting a sustainable funding model
- Making decisions that are open, participatory and deliberative
- A lack of maturity in the surrounding ecosystem
- Demonstrating trustworthiness
- Using data trusts to steward personal data
- Avoiding technology-first solutions

Can we help
organisations to build
data trusts?

Building data trusts



Scope

Establishing what the problem is

Co-Design

Designing a data trust collaboratively

Launch

Building and implementing a data trust

Operate

Maintaining and building on the initial structure

Evaluate

Assessing effectiveness, including external evaluation

Retire

Retiring a data trust if it has served its purpose

Recommendations

For data holders exploring data trusts

- Ensure data trusts are an appropriate data stewardship approach to meet your goals
- Be clear about what you mean by the term ‘data trust’ and use a different term if necessary
- Use and improve the data trust life-cycle the ODI has developed
- Engage a multi-disciplinary team
- Be trustworthy in how you collect, use and share data
- Adopt principles that strengthen data infrastructure
- Work in the open and make use of existing networks



Contents:

1. Increasing access to data while retaining trust (PW, 20 mins)
2. Case study - what is a data trust and what has the ODI learnt from studying them (JH, 10 mins)
3. Discussion (All, 20 mins)

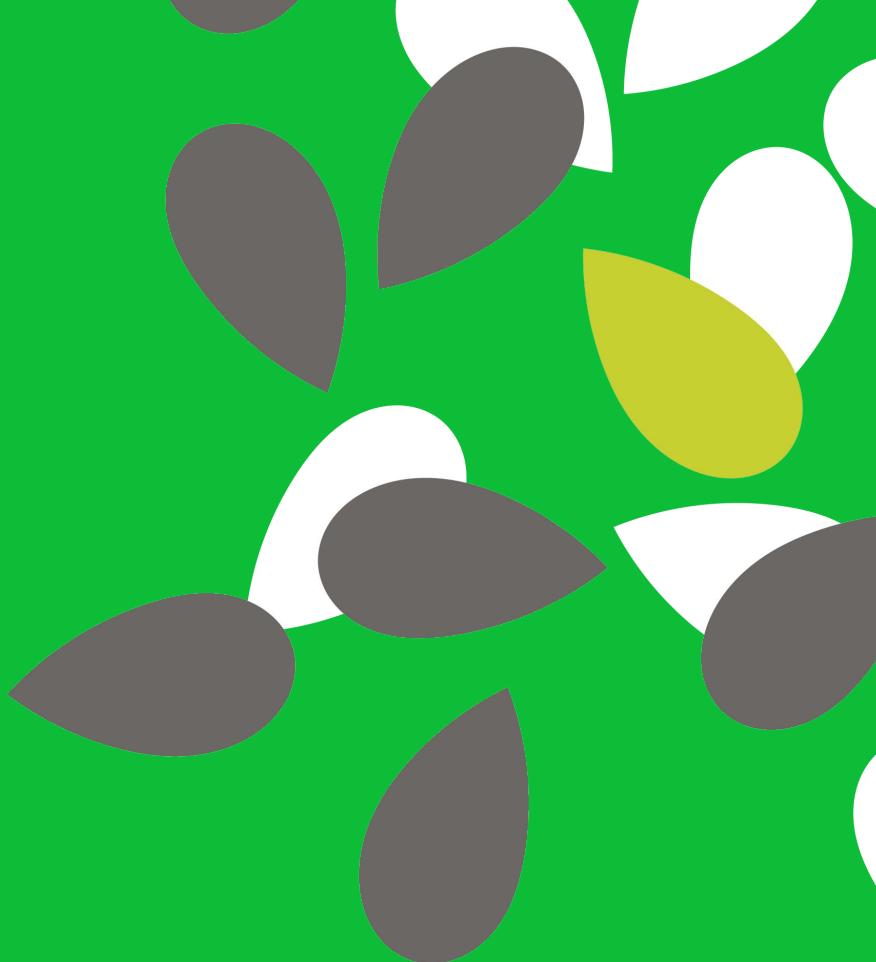
Any questions?

If you would like to talk to us about collaborating, partnering, supporting our work, or anything else, we'd love you to get in touch.

info@theodi.org

+44 (0)20 3598 9395

@ODIHQ



Discuss!

- What approaches to data access can be most beneficial?
- What are the difficulties you face in adopting these approaches?
- What would you need to more quickly and effectively adopt new approaches to data access?

Thank you!