DataKind UK

Data science and small charities

Giselle Cory, Executive Director

We’re a charity that builds **data science capacity** in social change organisations

NSPCC
Plan
Oxfam
SHARED ASSETS
ECOLOGICAL LAND CO-OPERATIVE
COFFEES OF CHARACTER CAFÉDIRECT MADE THE SMALL WAY
The Access Project
CAFÉDIRECT Producers' Foundation
Ark
STREET LEAGUE
WOODLAND TRUST
ONE THE CAMPAIGN TO MAKE POVERTY HISTORY
THEATRE FOR A CHANGE
Education · Awareness · Action
Prince's Trust
THE MIX
one world
global witness
NORTH EAST CHILD POVERTY COMMISSION
WeFarm
HelpAge International age helps
centre point give homeless young people a future
Shelter
The Welcome Centre
the key YOUR POTENTIAL UNLOCKED
Shooting Star Chase Children's Hospice Care
American Red Cross
the audience agency
Oasis Community Learning
CDI Responding To Needs, Driving Change
CONCERN worldwide
mind for better mental health
St Mungo's Broadway Rebuilding lives, day by day
WE ARE MACMILLAN. CANCER SUPPORT
Suas Educational Development
AMNESTY INTERNATIONAL
citizens advice
NCVO
Islington age UK
STREETS OF LONDON TACKLING HOMELESSNESS
the trussell trust Stop UK Hunger
BUTTLE uk FOR CHILDREN & YOUNG PEOPLE
ANTHONY NOLAN BE A MATCH, SAVE A LIFE
THE YOUNG FOUNDATION


DATA FOR GOOD
Meningitis now

Use of AI in social sector is growing

[Data Justice Lab]

Over 50 local councils using predictive models and 14 police forces

See Sky News [report]

[Guardian Research]

Automating poverty

Digital dystopia: how algorithms punish the poor

In an exclusive global series, the Guardian lays bare the tech revolution transforming the welfare system worldwide - while penalising the most vulnerable

by Ed Pilkington in New York

All around the world, from small-town Illinois in the US to Rochdale in England, from Perth, Australia, to Dumka in northern India, a revolution is under way in how governments treat the poor.

You can't see it happening, and may have heard nothing about it. It's being planned by engineers and coders behind closed doors, in secure government locations far from public view.

Only mathematicians and computer scientists fully understand the sea change, powered as it is by artificial intelligence (AI), predictive algorithms, risk modeling and biometrics. But if you are one of the millions of vulnerable people at the receiving end of the radical reshaping of welfare benefits, you know it is real and that its consequences can be serious - even deadly.

Using machine learning to identify food bank dependence early

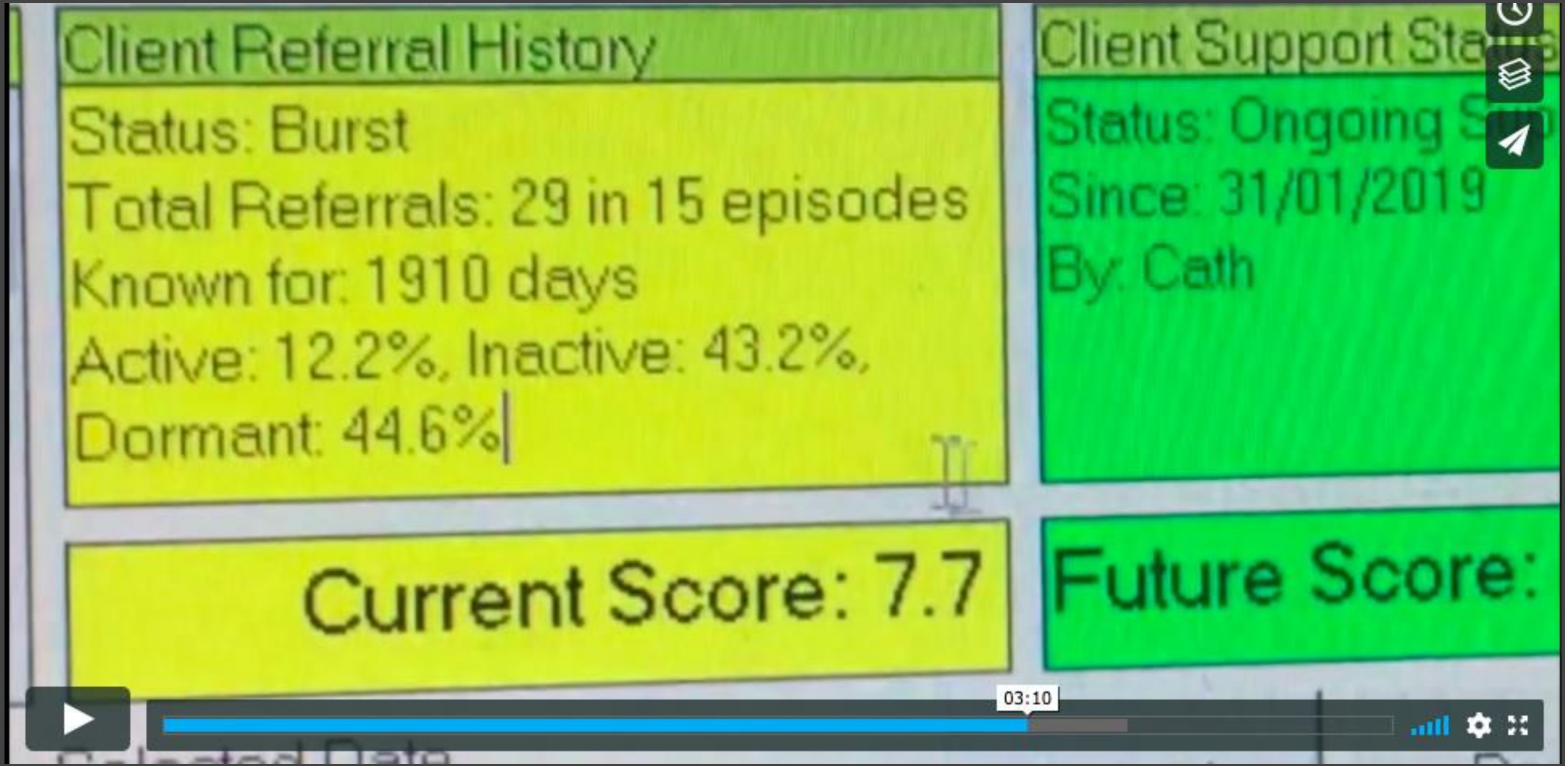
Project Funding, a real enabler organisation

Applying **data science** talent, tools and processes to achieve **social good**

MEETS

Providing packs of food, essentials and support to people **in crisis** in Huddersfield and the surrounding area

Both teams consisted of volunteers. All technical work completed ‘out of hours’



https://vimeo.com/343034018

Food bank use is on the rise...

The Welcome Centre
food bank and more...

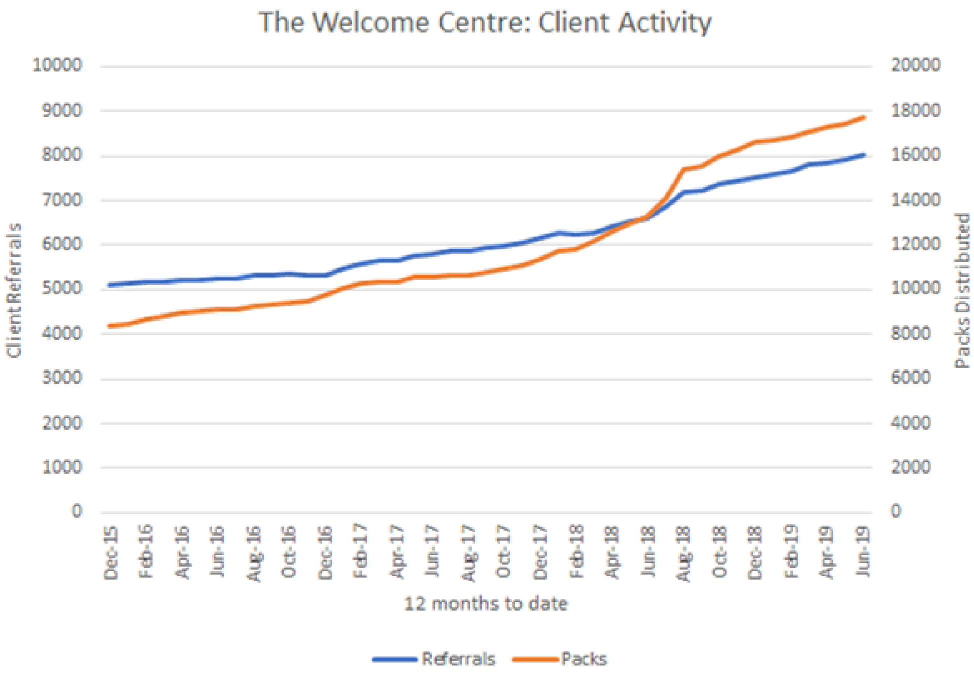
Hungry for answers

The real reasons why food-bank use is soaring

A shift in the nature of poverty, not its overall incidence, best explains their spread

Universal credit reform fails to satisfy food bank users

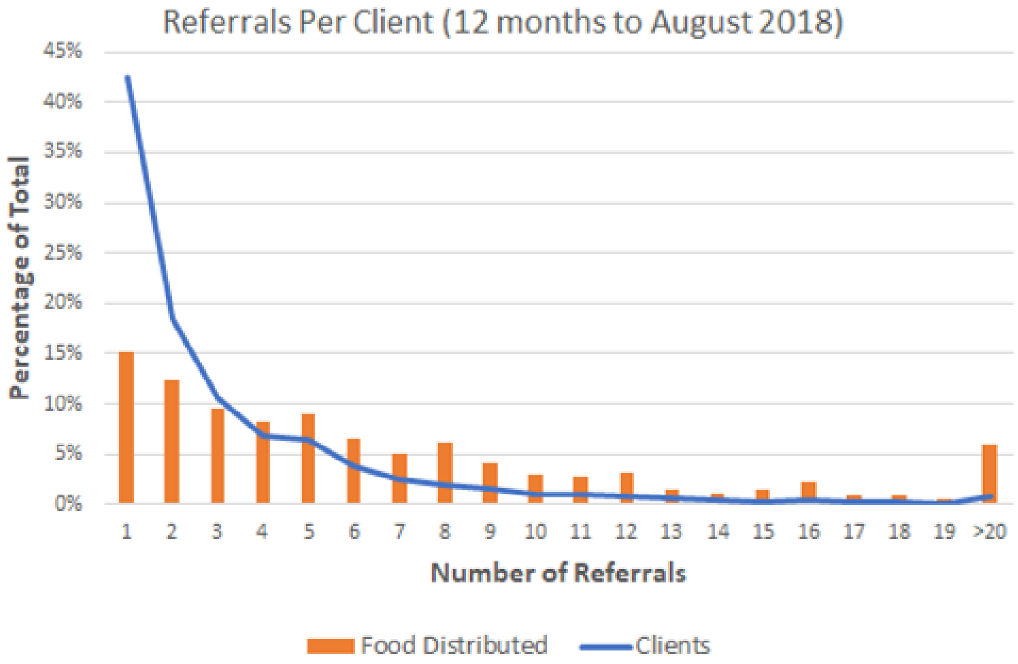
Slow payments and benefits freezes 'much bigger deal' than rise in work allowance



Data from the Welcome Centre

The way they are being used is also changing...

Food bank users are coming back more regularly and becoming **dependent** on the support

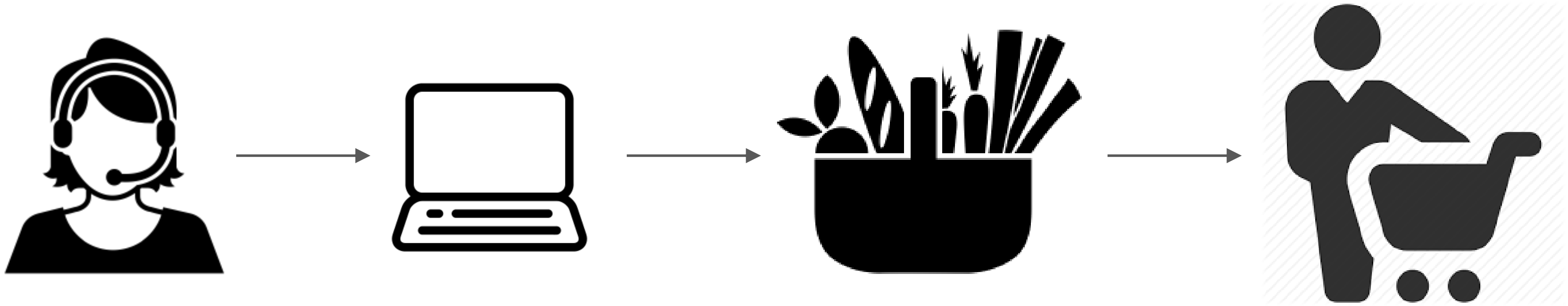


- Most clients visit relatively few times
- But the few clients who visit regularly consume a high proportion of resources and are at risk of becoming dependent
- Hence need for support

Data from the Welcome Centre

Typical Journey

The Welcome Centre
food bank and more...



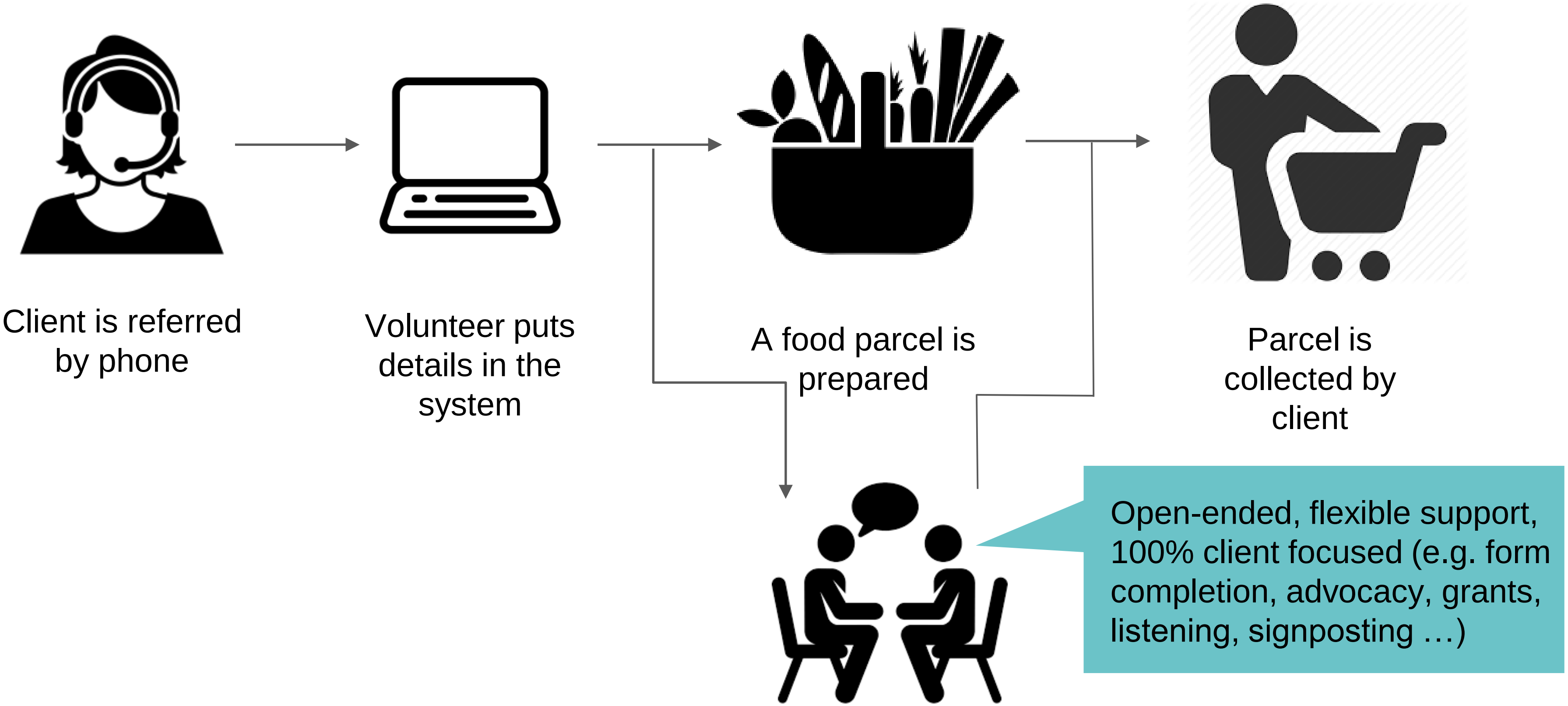
Client is referred by phone

Volunteer puts details in the system

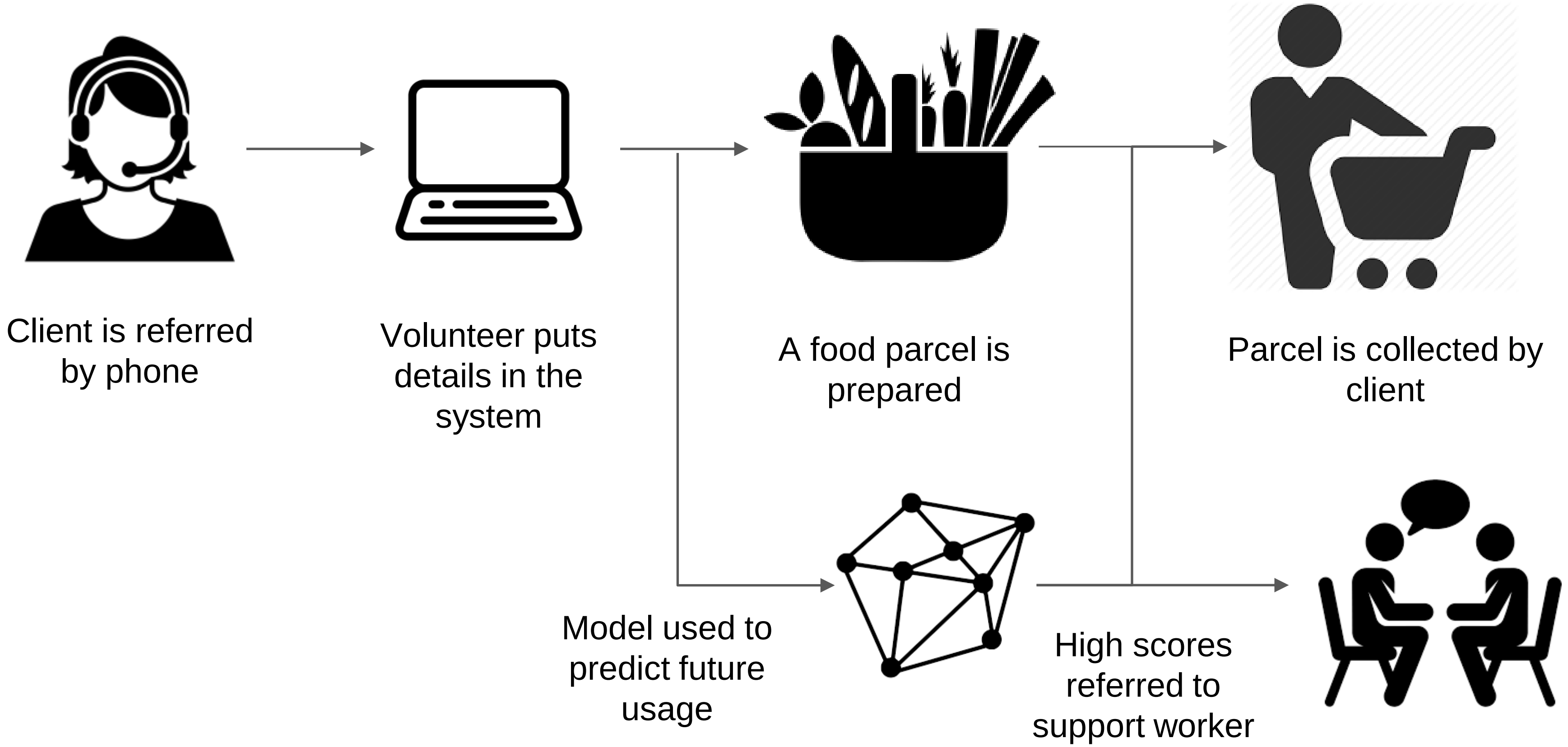
A food parcel is prepared

Parcel is collected by client

Supported Journey



Using Machine Learning to Triage



The Challenge

DKUK

Integrate it into a system written in **`BASIC`** on Windows servers

Have it run with zero maintenance and low cost

Provide a way to retrain the model

Build a production grade model that predicts future visits

Avoid any socio-demographic bias

Ensure it’s GDPR compliant

Make sure the staff understand it

The Model

Client Features:
These are features that describe the client, They are typically static throughout a client history.

- Address Length
- Dependent Details
- Living with Partner
- Client Issues

Referral History Features:
Features derived from the pattern of previous referrals.

- # of prev. referrals
- # of referrals in given window lengths
- Time since last referral

Extra Trees Regressor

Model

Score

Current Referral Features:
These are features that describe the current referral

- Agency
- Referral Issues
- Referral Benefit
- Referral Domestic Circumstances

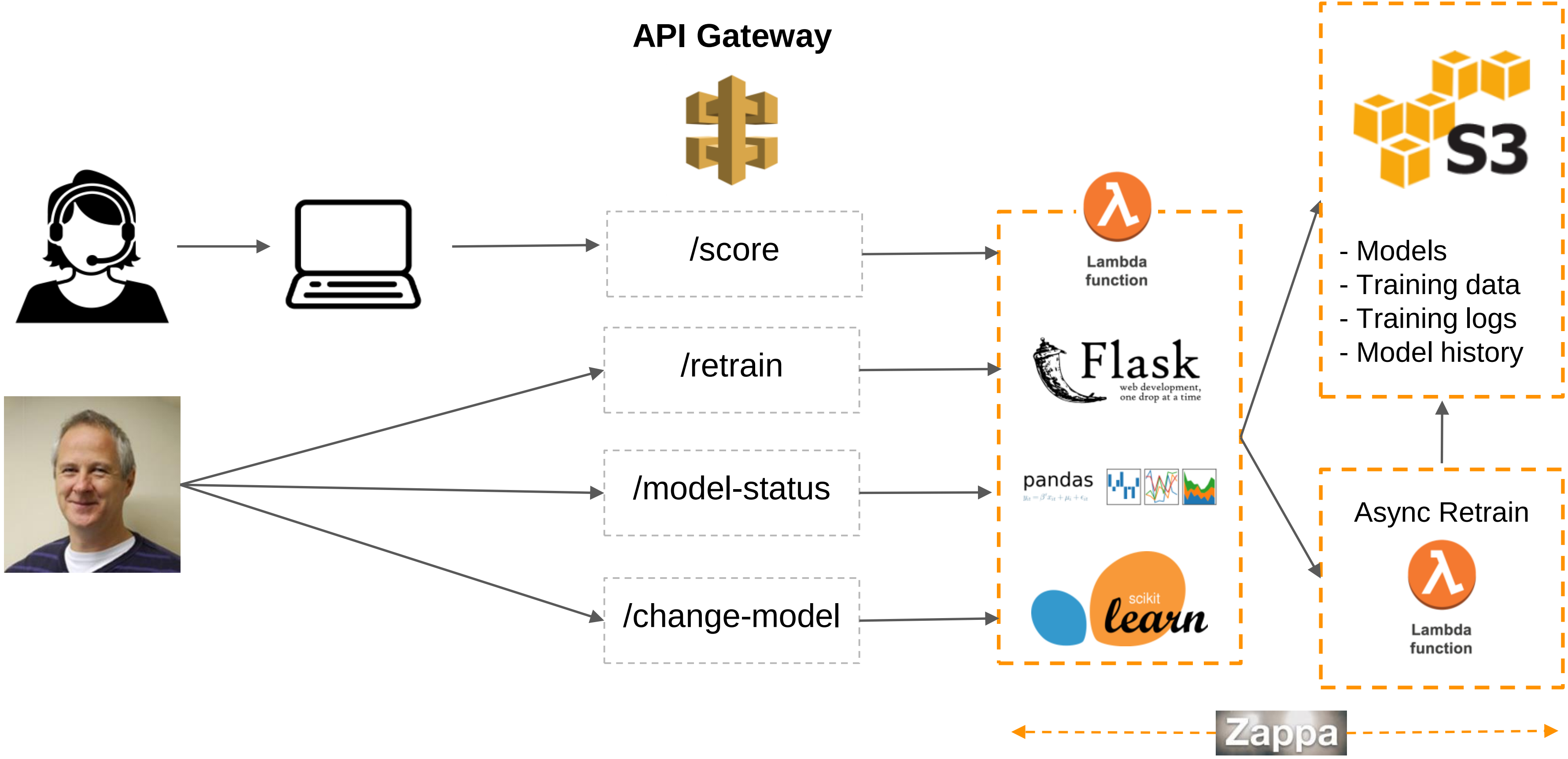
Previous Referral Features:
Features from previous referrals

- Previous Agencies
- Previous Issues
- Previous Benefits
- Previous Domestic Circumstances

** we will not be using free text fields as these are very sensitive to the agent inputting the data

Advanced technology

DKUK



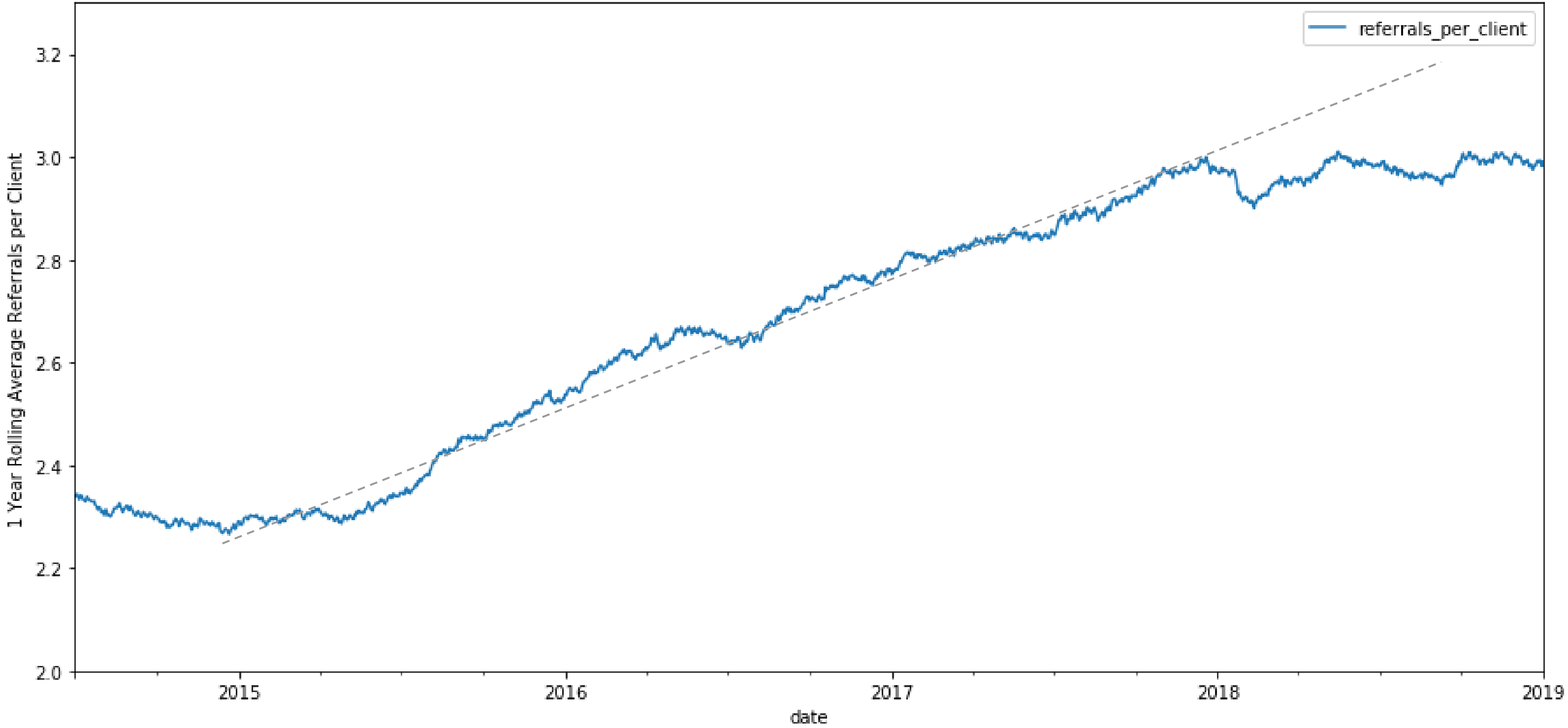
...at a low and sustainable cost

AWS server

£1.20 a month!

(and no local support overhead for The Welcome Centre)

Support and the model is reducing dependency



How well is the model working? (Anecdotal)

The Welcome Centre
food bank and more...

- 40 referrals per day, around 50% from new clients.
- ~10 model-based support recommendations made per day, depending upon overall demand and the specific clients being referred
- Very few false positives

Anecdotal evidence

- The support worker ‘trusts’ the recommendations and no longer spends time on manual checking
- After 9 months the system is an integral part of TWC’s business processes (not sitting on a shelf gathering dust)

Direct impacts of the Welcome Centre

Early intervention

Allows the Welcome Centre to realise the goal of meeting with clients at the beginning of their journey

Time Saving

The support worker has more time to spend in direct contact with clients and spends less time reviewing new client referrals

More effective client engagement

Provision of food packs becomes part of support management plan rather than being based on ad hoc referrals

New Ways of Working

Facilitates the move from a purely reactive to a more proactive way of engaging with clients

Wider impacts of the Welcome Centre

Funding bid

Funding bid for expanding the support service submitted. This will provide a second support worker, backed by volunteers. Use of the recommendations made by the model is central to this bid

(model threshold can be adjusted to match support capacity)

Enhanced strategy

Enabled more effective proactive approach of support service to become a ‘core’ function within TWC

Early intervention is vital

[Mary’s story]

Data-driven organisation

Good data and data analysis is now regarded as a key tool for informing strategic decision making at board level

The Welcome Centre
food bank and more...

Wider community impacts

DKUK

TIC TeC

dataiku

Strata Data Conference

nesta

superhighways
harnessing technology for community benefit

Impact Aloud 2018 - Presentation

"Really inspiring and eye-opening about what could be achieved by a small organisation with this sort of help."

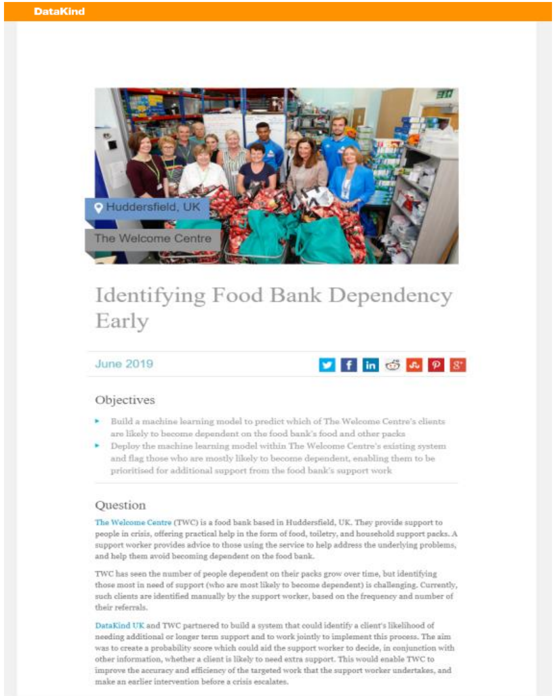
DatSci AWARDS 2019

Finalist for *Best Use of Data to Achieve Social Impact*

Babbage: Data to the rescue
Babbage from Economist Radio

THE BIG ISSUE

Wider community impacts - further projects


DataKind

Huddersfield, UK

The Welcome Centre

Identifying Food Bank Dependency Early

June 2019

Objectives

- Build a machine learning model to predict which of The Welcome Centre's clients are likely to become dependent on the food bank's food and other packs
- Deploy the machine learning model within The Welcome Centre's existing system and flag those who are mostly likely to become dependent, enabling them to be prioritised for additional support from the food bank's support work

Question

The Welcome Centre (TWC) is a food bank based in Huddersfield, UK. They provide support to people in crisis, offering practical help in the form of food, toiletry, and household support packs. A support worker provides advice to those using the service to help address the underlying problems, and help them avoid becoming dependent on the food bank.

TWC has seen the number of people dependent on their packs grow over time, but identifying those most in need of support (who are most likely to become dependent) is challenging. Currently, such clients are identified manually by the support worker, based on the frequency and number of their referrals.

DataKind UK and TWC partnered to build a system that could identify a client's likelihood of needing additional or longer term support and to work jointly to implement this process. The aim was to create a probability score which could aid the support worker to decide, in conjunction with other information, whether a client is likely to need extra support. This would enable TWC to improve the accuracy and efficiency of the targeted work that the support worker undertakes, and make an earlier intervention before a crisis escalates.

DKUK Workshop

GREATER LONDON AUTHORITY

superhighways harnessing technology for community benefit

DKUK

hear humanity. equality. rights.

london plus

2-year initiative to help small charities in London to improve their digital skills and analysis of data

TWC Client Journey - Mary

Refereed by Local Welfare and recommended for support by the DK model 47 year old Mary had never claimed benefits, she had worked all her life, brought up three children and was a part-time carer for her granddaughter. Sadly she had a breakdown and became very unwell. She was signed off work for three months, but soon realised that she needed to give up work to concentrate on her health and well being.

Mary made a Universal Credit claim on line with the help of her sister, but there were issues with her bank account and therefore payment was delayed. Cath saw Mary on a one 2 one basis and over the next few months, she helped Mary to do the following:

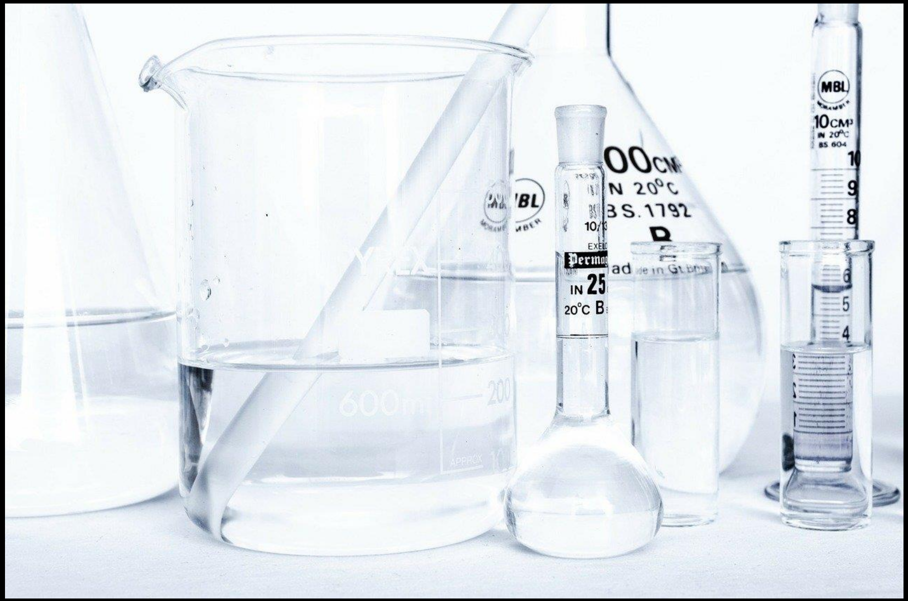
- Phone DWP to sort out the issue with her bank account.
- Complete a UC50 form – medical health form.
- Prepare for the medical assessment that she had to attend.
- Apply for PIP – personal independence payment.
- Apply for council tax reduction.
- Speak to the carers allowance team regarding a change in circumstance.
- Refer Mary for money advice via KNH money advice unit.

Mary now has her benefits in place, she has been found unfit for work and is entitled to higher rate benefit for 12 months, giving her the time she needs to attend counselling, so that she can get better. She was given seven food parcels during this time, five whilst her Universal Credit was processed (standard for all UC claimants) and two extra food parcels whilst the issue with her bank account was resolved.

What is AI
and Machine Learning?

- →What is **AI?**
- →What is **machine learning?**
- →How could this be applied to **your organisation?**

➔But first....


MBL
IN 20°C


SLOWLY
PLEASE

What is AI?

→What is **machine learning**?

→How could this be applied to **your organisation**?


A.I. is a fundamental risk to the existence
of human civilization.

Artificial intelligence

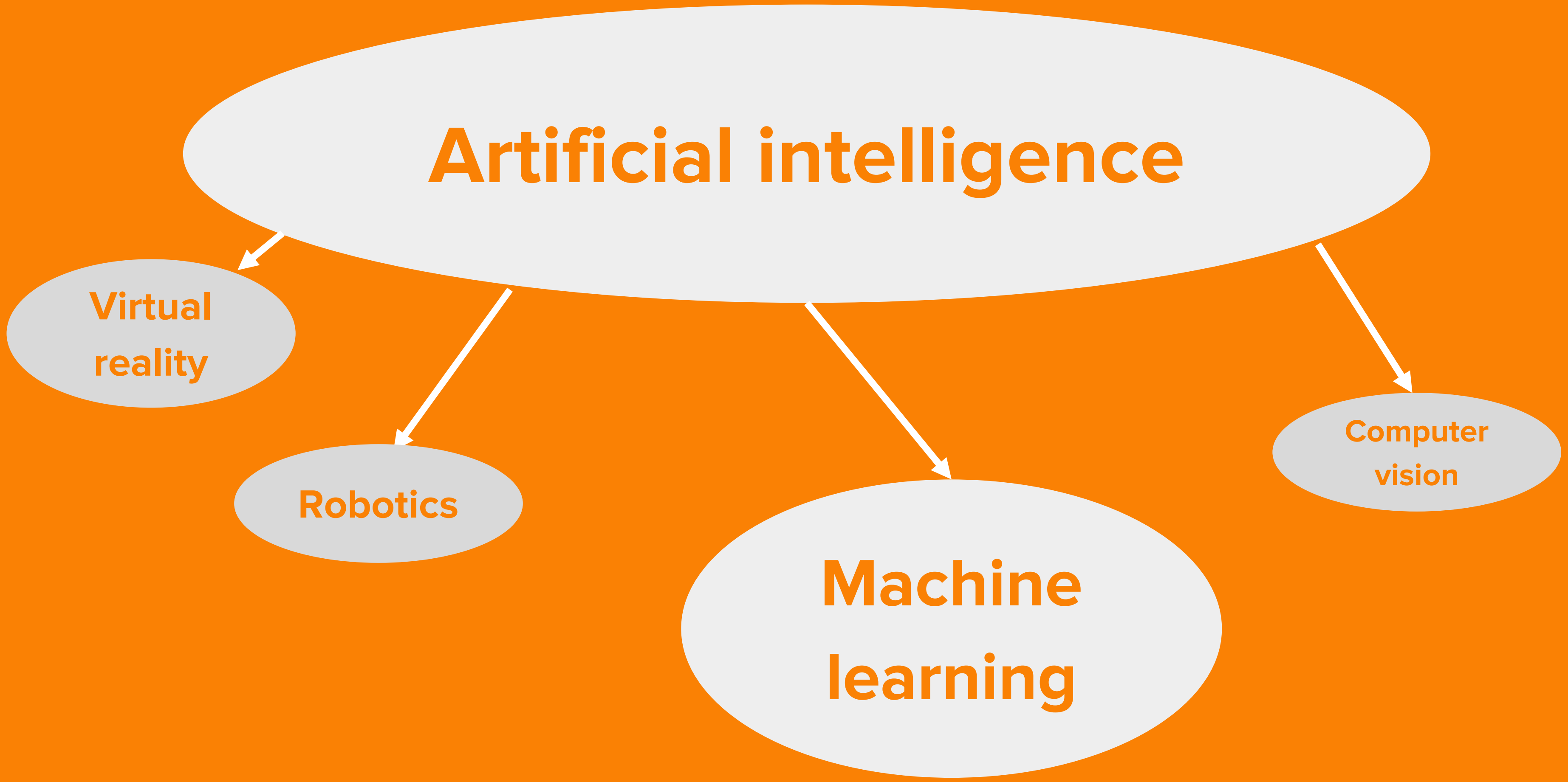
Artificial intelligence

General intelligence

A machine capable of true learning

Narrow intelligence

A machines that is good at one task


Artificial intelligence
Virtual reality
Robotics
Machine learning
Computer vision

- → What is **AI**?
- What is **machine learning?**
- → How could this be applied to **your organisation?**

Machine learning

Supervised: We tell the model what we want to look out for

Unsupervised Learning: The model does its own thing

→What is AI?

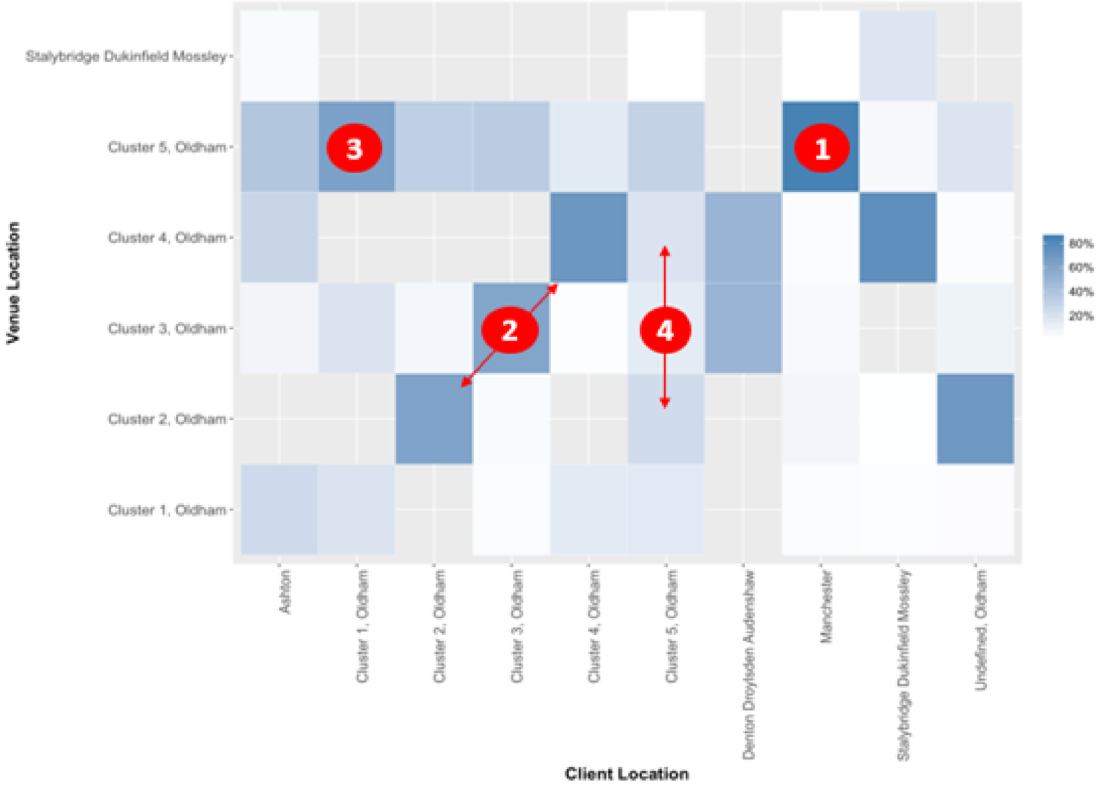
→**What is machine learning?**

How could this be applied to your organisation?

To provide additional insight or challenge/confirm existing assumptions to enable humans to deliver better services

- Understand need and demand
- Effective campaigning
- Understand the people we work with
- Evaluate services
- Improve operational efficiency

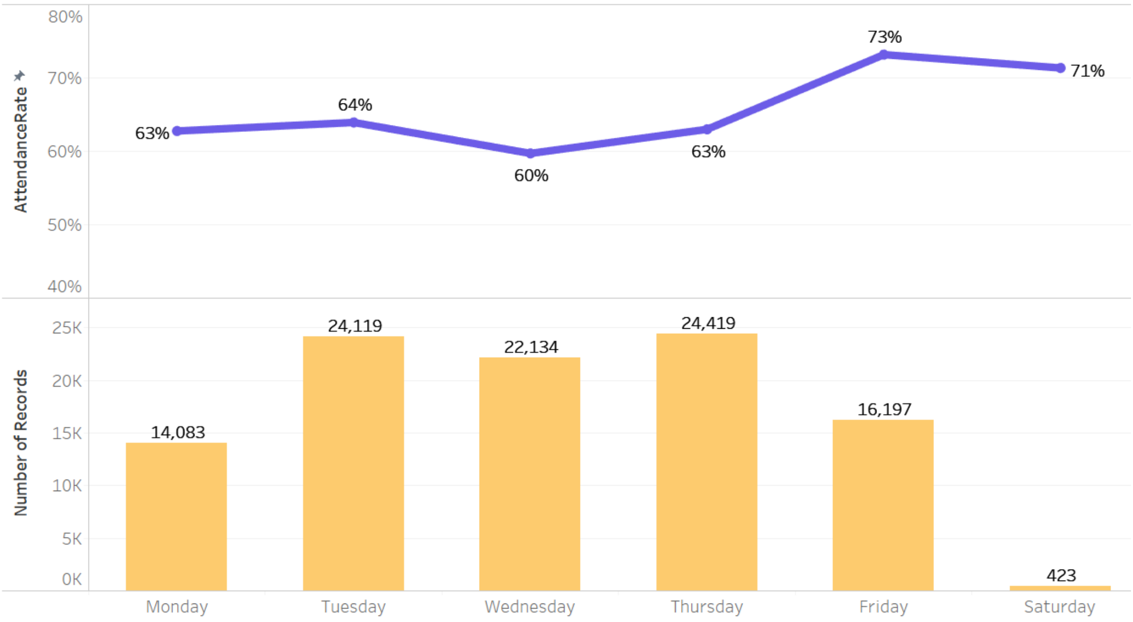
Q. Who accesses our services and where?



1. Clients who live in Manchester are much more likely to visit venues in Cluster 5 – this is the closest cluster to central Manchester
2. We would expect to see a pattern of clients visiting venues where they live – this trend is not evident for clients who live in cluster 1 or 5.
3. Clients who live in cluster 1 are more likely to visit cluster 5 venues
4. Clients who live in Cluster 5 are visiting venues in all clusters, with no pronounced preference

mind for better mental health | Tameside, Oldham and Glossop

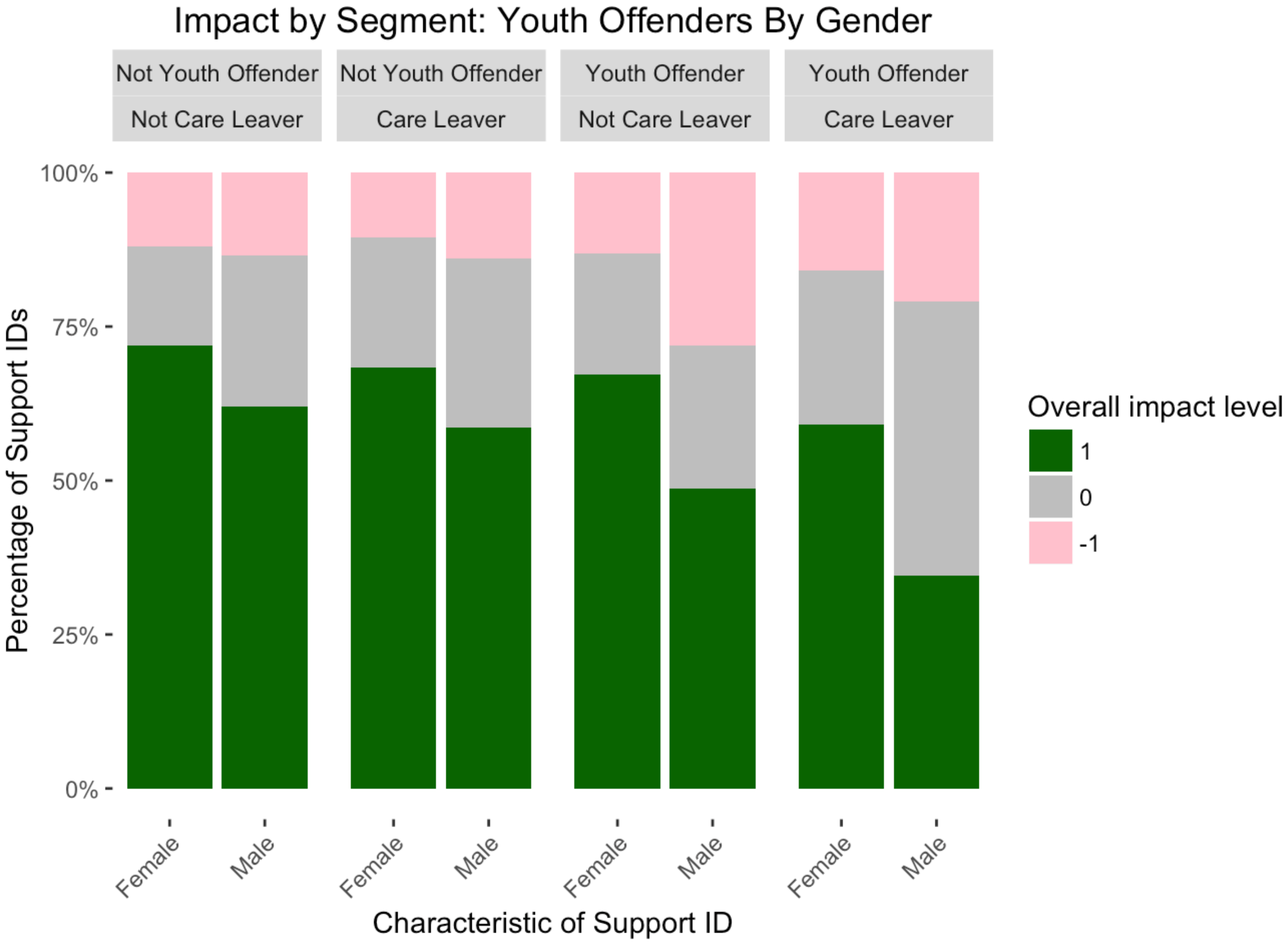
Q. What are the best times to offer services?



Challenges assumption that weekend classes are less attended

mind for better mental health | in the City, Hackney and Waltham Forest

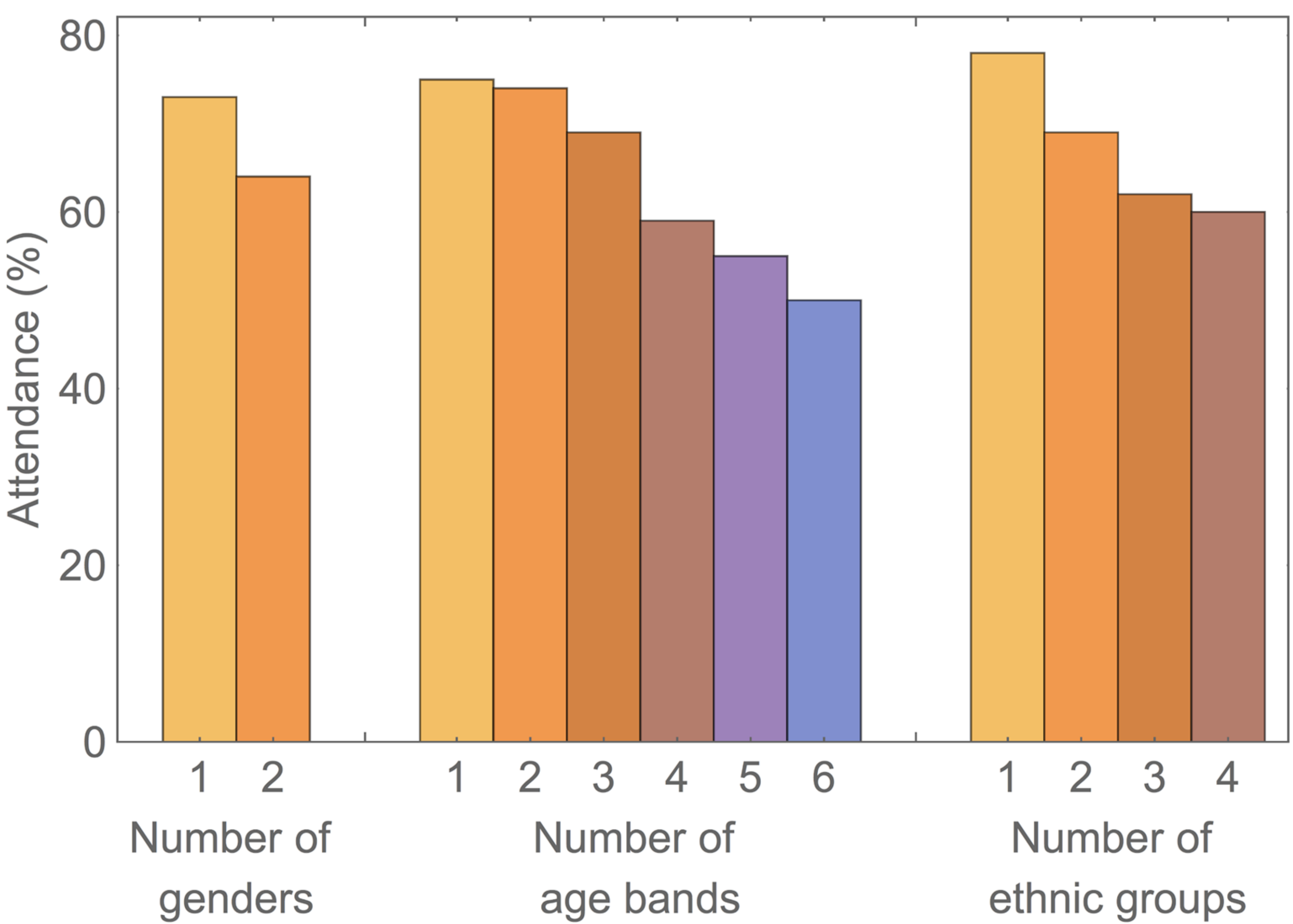
BUTTLE uk
FOR CHILDREN & YOUNG PEOPLE

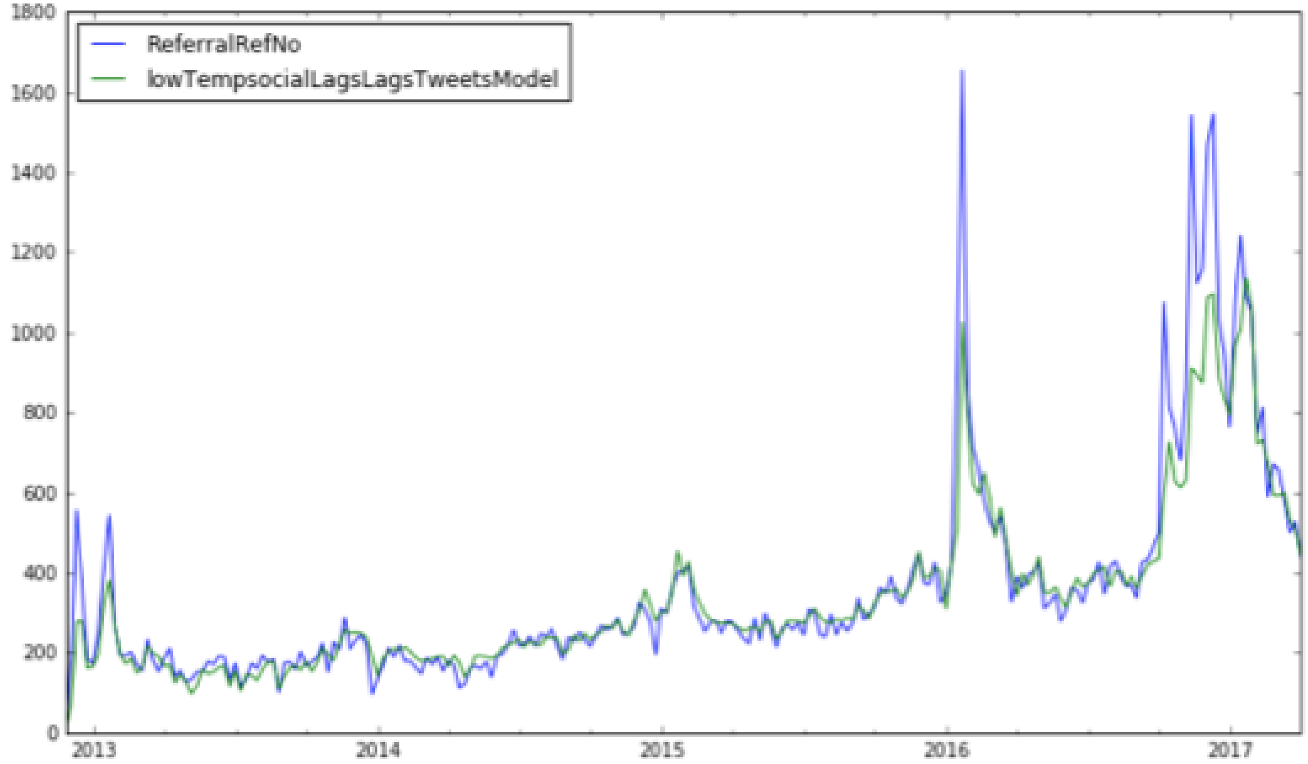


Impact by Segment: Youth Offenders By Gender
Not Youth Offender
Not Care Leaver
Not Youth Offender
Care Leaver
Youth Offender
Not Care Leaver
Youth Offender
Care Leaver
100%
75%
50%
25%
0%
Percentage of Support IDs
Female
Male
Female
Male
Female
Male
Female
Male
Characteristic of Support ID
Overall impact level
1
0
-1

Mind for better mental health

in the City, Hackney and Waltham Forest





ReferralRefNo
lowTempsocialLagsLagsTweetsModel
1800
1600
1400
1200
1000
800
600
400
200
0
2013
2014
2015
2016
2017

Data maturity

the key

The Themes of Data Maturity

CULTURE
Team approach, self-questioning, openness and sharing, governance.

SKILLS
Internal capacity, roles and skill levels, access to external knowledge and expertise.

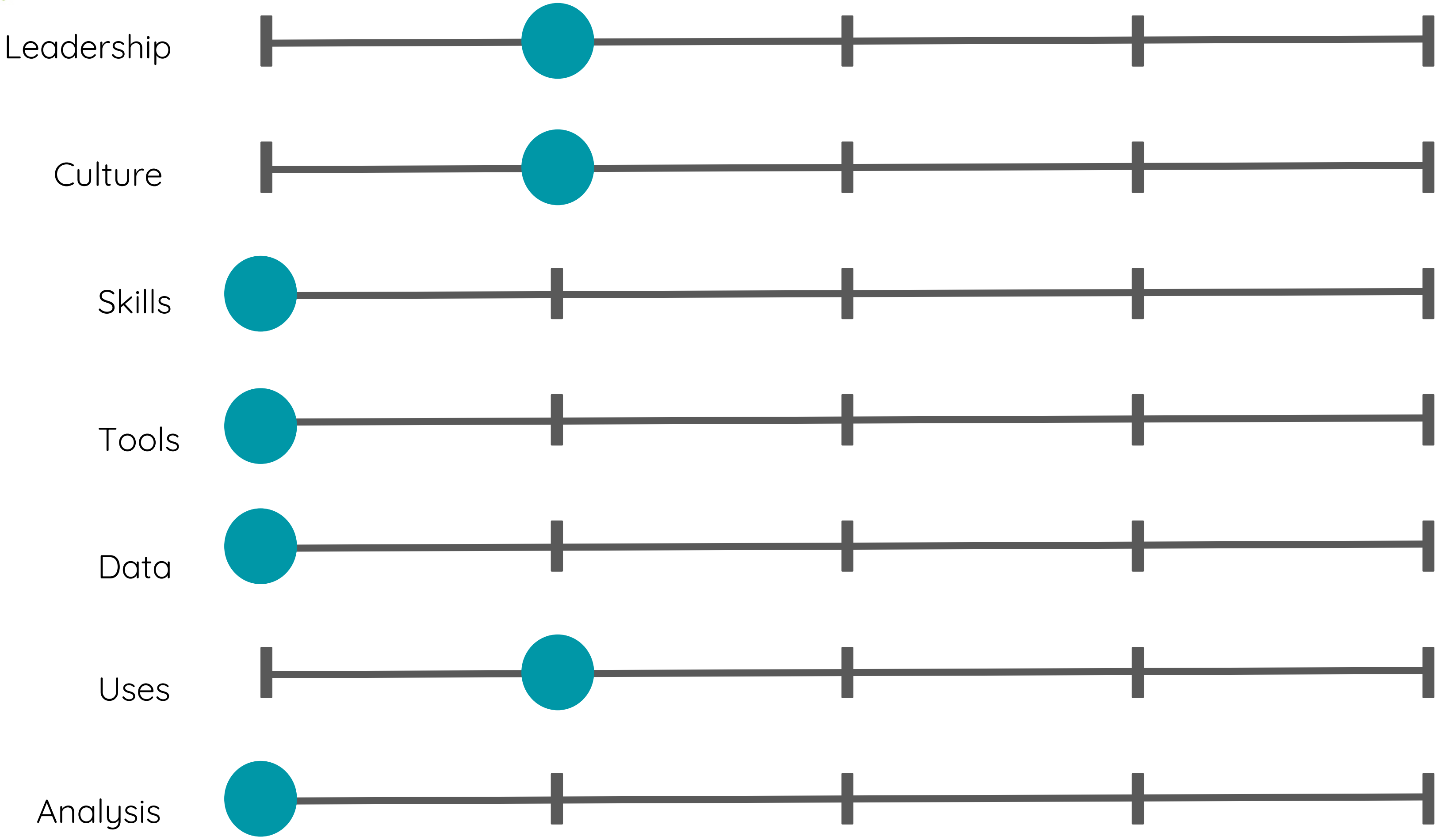
LEADERSHIP
Attitude, investment, plans for data development, alignment to business plans, capability.

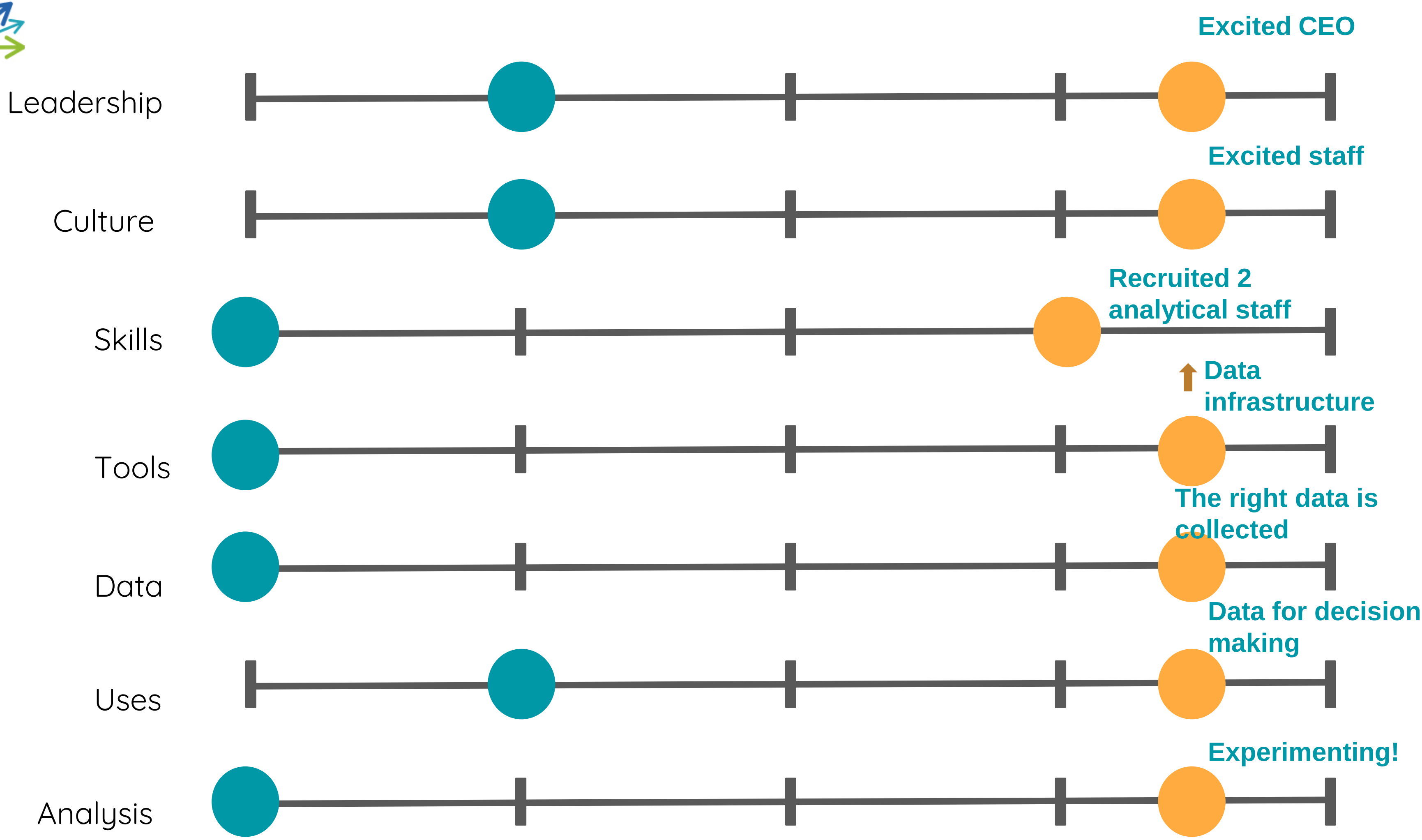
DATA
Assets, Collection, sources, quality.

TOOLS
Storage, type and quality of tools and infrastructure.

USES
Range and extent of reasons for collecting and analysing data, and benefits and rewards reaped.

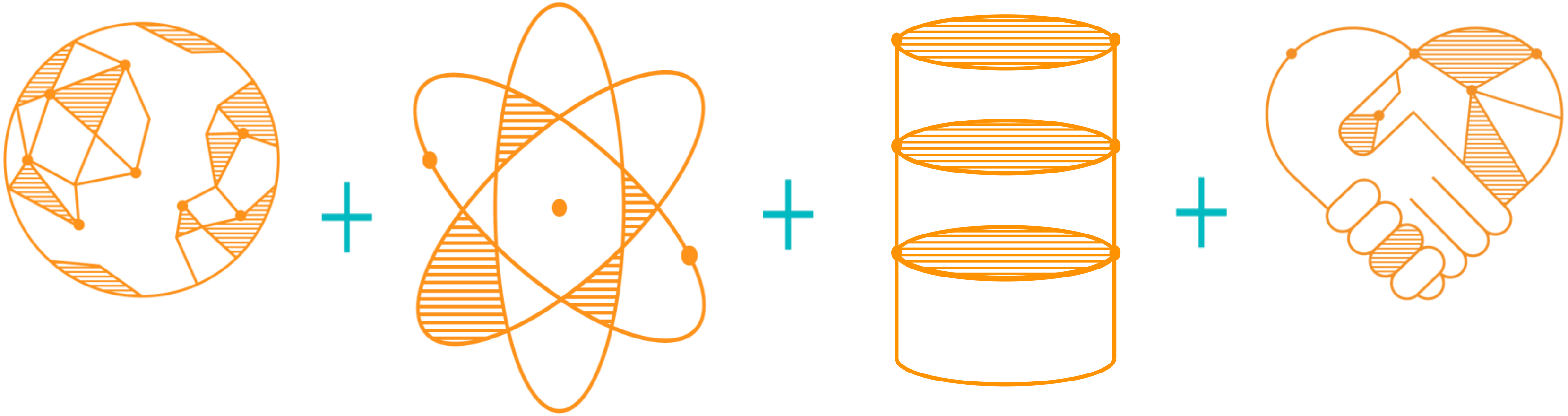
ANALYSIS
Type of data analysed, techniques, presenting and communicating.


Leadership
Culture
Skills
Tools
Data
Uses
Analysis


Excited CEO
Leadership
Excited staff
Culture
Recruited 2 analytical staff
Skills
Data infrastructure
Tools
The right data is collected
Data
Data for decision making
Uses
Experimenting!
Analysis

Machine learning **problem framing**

Ideal data for good projects have …



Well-framed
Problem

Simple
(enough)
Solution

Relevant &
Responsible
Data

Social actor
Partner

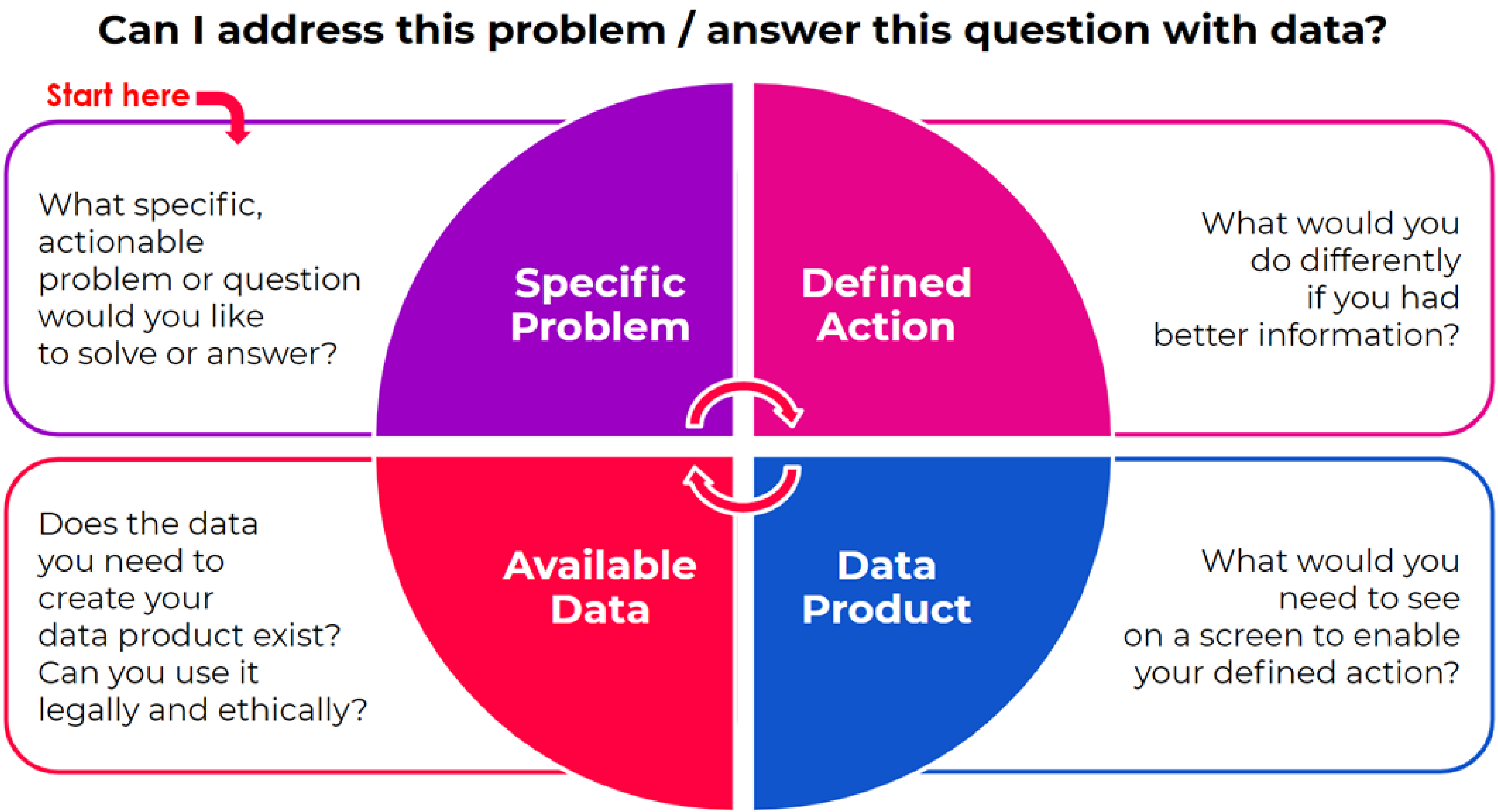
What is the problem?

- What are some of the barriers you are facing?
- Does it align with a strategic priority?

Two helpful resources:

1. Blog from Eddie Copeland A four step step approach to collaborative data projects
2. Google: Introduction to Machine Learning Problem Framing part of Google AI for Social Good Guide

Eddie Copeland: [A four step step approach to collaborative data projects]



Google's Introduction to Machine Learning Problem Framing Exercise

1. What would you like the machine learning model to do?
2. What is the ideal outcome?
3. How will you know if it is successful?
4. What is the output of the model?
5. When do you need the output and how will it be used?
6. If you didn't use ML, what would you use?

1. Is there a clear problem that is suitable for ML/AI techniques?

> *Look at applied cases e.g [DataKind UK](), [Google]()*

Adapted from Google: Introduction to Machine Learning Problem Framing- Identifying Good Problems for ML

1. Is there a clear problem that is suitable for ML/AI techniques?
2. Is there sufficient data?

   > ***Size*** *- The more data you have, more precise the estimate. More complicated ML models require hundreds of thousands of 'rows'*

   > ***Quality*** *- dodgy data in = dodgy results out*

Adapted from Google: Introduction to Machine Learning Problem Framing- Identifying Good Problems for ML

1. Is there a clear problem that is suitable for ML/AI techniques?
2. Is there sufficient data?
3. Are there assumptions that can be tested?

Adapted from Google: Introduction to Machine Learning Problem Framing- Identifying Good Problems for ML

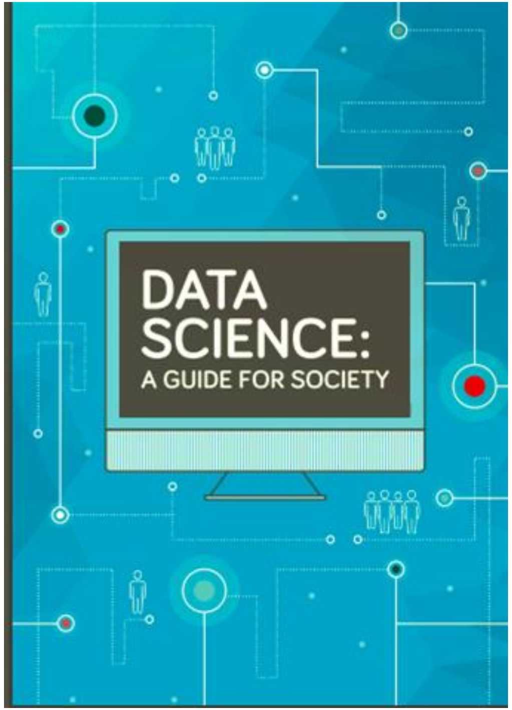
1. Is there a clear problem that is suitable for ML/AI techniques?
2. Is there sufficient data?
3. Are there assumptions that can be tested?
4. Will the resulting ML/AI be used for an action?

Adapted from Google: Introduction to Machine Learning Problem Framing- [Identifying Good Problems for ML]

Doing it right

1. Where does the data come from?
2. What assumptions are being made?
3. Can it bear the weight being put on it?

From [Sense About Science]



Data Science

Step	Example
1. Set the research goal.	I want to predict how heavy traffic will be on a given day.
2. Make a hypothesis.	I think the weather forecast is an informative signal.
3. Collect the data.	Collect historical traffic data and weather on each day.
4. Test your hypothesis.	Train a model using this data.
5. Analyze your results.	Is this model better than existing systems?
6. Reach a conclusion.	I should (not) use this model to make predictions, because of X, Y, and Z.
7. Refine hypothesis and repeat.	Time of year could be a helpful signal.

Reference: Google: Introduction to Machine Learning Problem Framing https://developers.google.com/machine-learning/problem-framing/big-questions

Tips for low resource settings

People

- Invest in people before platforms

- Invest in people before platforms

- Upskill staff

- Invest in people before platforms

- Upskill staff

- Invest in people before platforms

- Upskill staff

- Add an apprentice

- Invest in people before platforms

- Upskill staff

- Add an apprentice

- Invest in people before platforms

- Upskill staff
- Add an apprentice
- Use support networks

- Invest in people before platforms

- Upskill staff
- Add an apprentice
- Use support networks

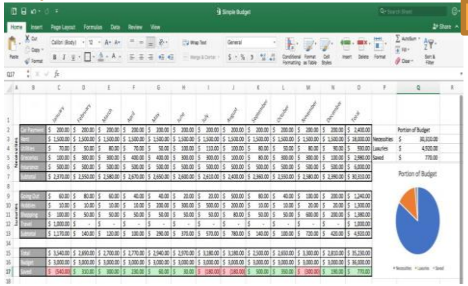
Social Data Society

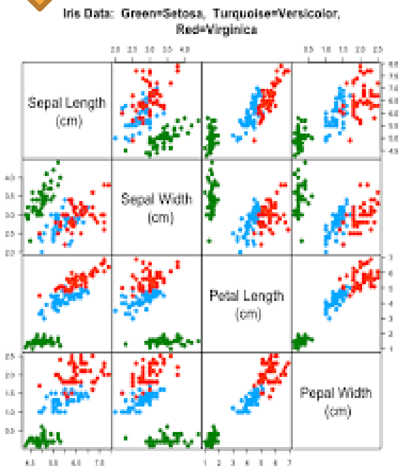
Responsible Data

Leading the way from best intentions to best practice.

Tools

Everyone is on a digital/data journey!





Get there without code

Get there without code

FRI, 31 JAN 13:30

Data essentials

Waterloo Action Centre, London

FREE

#CharitiesCauses #Course

WED, 26 FEB 14:00

Data science for social impact: small charity workshop

TBC - London

FREE

#CharitiesCauses #Course

TUE, 24 MAR 10:00

Excel for reporting and visualising your non-profit's data

The Foundry (A Place for Chnage), London

£48 - £96

#CharitiesCauses #Course

WED, 25 MAR 10:00

Putting your data on the map

Canada Water Theatre (Library), London

£30 - £54

#CharitiesCauses #Course

Make maps without code

Make maps without code

batchgeo

MapIt UK

Microsoft | Power BI

tableau

CARTO

LONDON DATASTORE

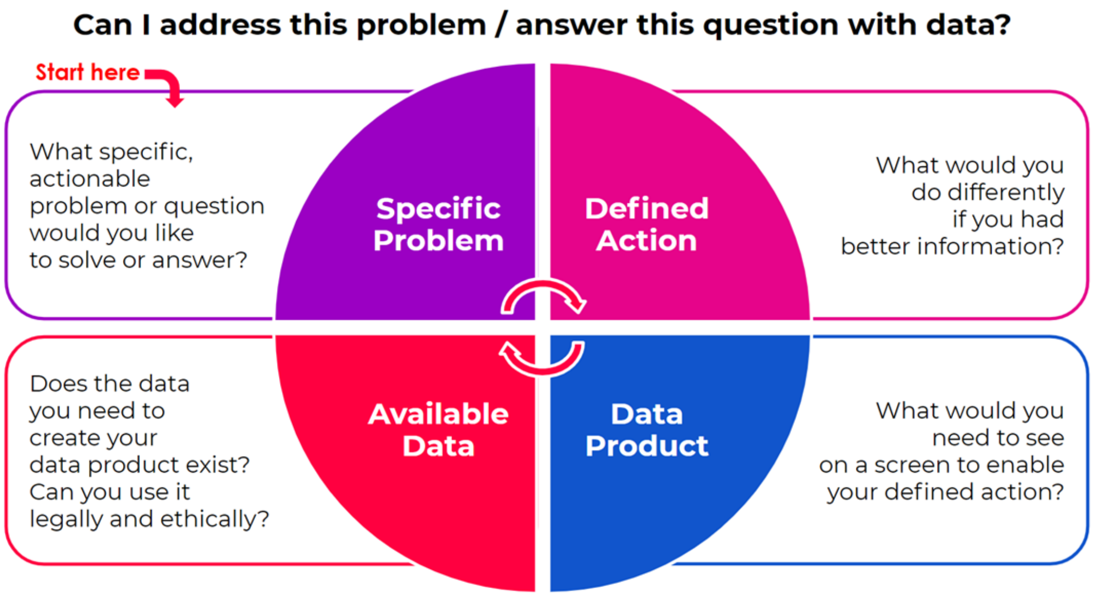
Get coding!

Get coding!

Approach

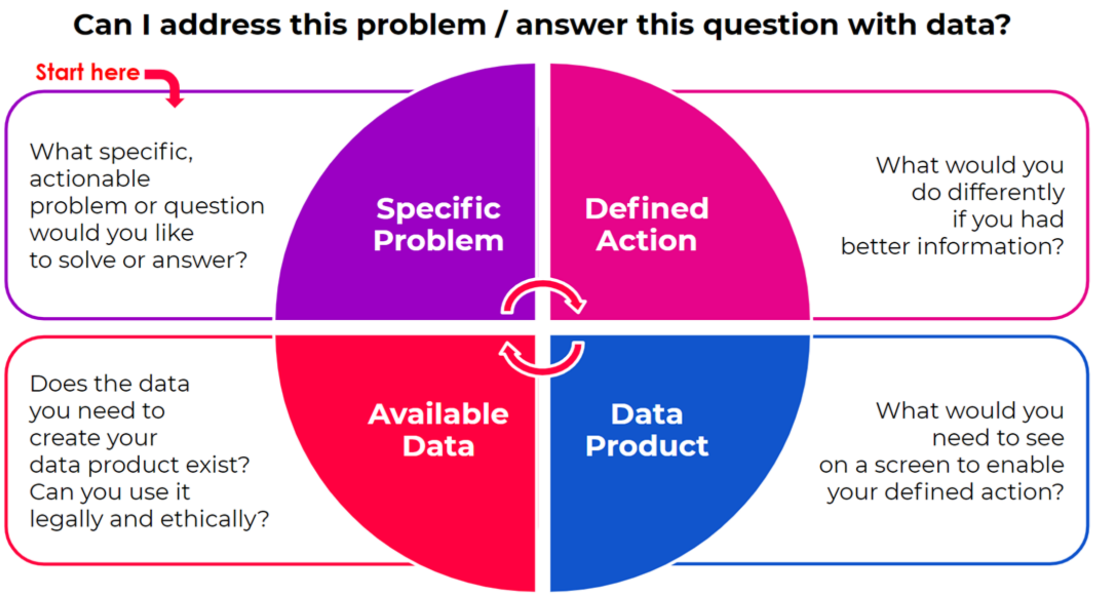
What is the problem?

What's the problem?



Eddie Copeland: [A four step step approach to collaborative data projects]

What's the problem?



INSPIRING IMPACT

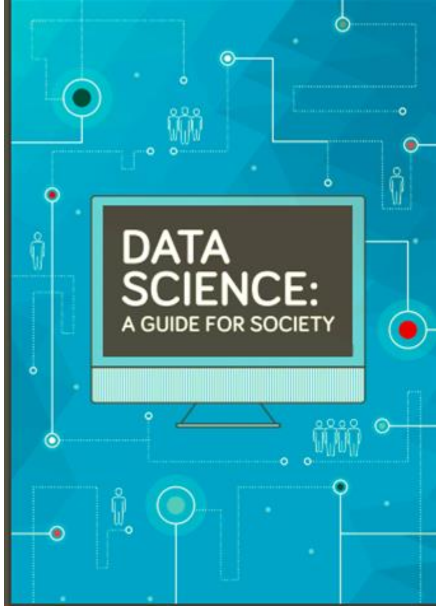
COALITION — FOR EFFICIENCY —

Eddie Copeland: A four step step approach to collaborative data projects

Ethics and law

From Sense About Science

1. Where does the data come from?
2. What assumptions are being made?
3. Can it bear the weight being put on it?



Make the case

- Show, don’t tell
    - Start with one thing
    - Don’t worry about the tech
    - Educate your funders!

- Show, don’t tell

- Use pro-bono support

- Show, don't tell

- Use pro-bono support

About us

What does DataKind UK do?

- Data therapy
- DataDives
- DataCorps

DataKind UK

@DataKindUK
giselle@datakind.org.uk