

Filled worksheet photos - Activity 3: Solution Design

Activity 3 - Solution Design

New Solution name:

Data Bridging

GROUP:

Challenge:

Challenge 1: Data access
Friction, slow process
Stats + time pressures

Early Risk mitigation -
Failed analysis
data protection policy

Who is it targeted at?
Researchers, Funders, Data Custodians

10% option to reduce risk of failure / delay downstream

Initial brainstorming of potential solutions:

- X Funded governance window
- X Development Awards
- X More open meta-data
- X Synthetic 'sample' data
- X Better data discovery tools
- X Data quality information

~~Info on data changes as part of governance / data prep.~~

- X Modes of the data provider

Details of solution:

What, where, how, resources needed, cost, timeframe (short-term or long-term project)

Generic bridging funded phase @ start of project for data discovery and bespoke curation

Upstream DPIA

Diagram illustrating the data flow and access process. A database feeds into a 'Digital Gateway', which then connects to 'Data Access Request' and 'Data Access' boxes. A person icon is shown interacting with the system. The process is labeled 'Generic bridging funded phase @ start of project for data discovery and bespoke curation'.

Diagram illustrating the data access process. A 'Data Access Request' box leads to 'Data Access' and 'Data Access' boxes. The process is labeled 'Research Institutions Investment'.

Activity 3 - Solution Design

New Solution name: The Big Stick

GROUP: 2

Challenge:

FRAGMENTED INFRASTRUCTURE
AND LACK OF INTEGRATED
SYSTEMS;

Initial brainstorming of potential solutions:

A UNIVERSAL METADATA
CATALOGUE - MANDATED AND
FUNDED ^{AND COORDINATED} CENTRALLY
TO ENABLE DISCOVERABILITY.
WITH A NATIONALLY ^{UK} AGREED
MINIMUM METADATA ^{STANDARD} incl DATA
DICTIONARY

IMPROVING INTEROPERABILITY BY
ENSURING AVAILABILITY OF DATA
SPECIFICATIONS / DATA DICTIONARIES
AND INFORMATION ABOUT DATA
COLLECTION & PROCESSING -
FUNDING DATA CURATION

Value proposition statement:

Standardising health data assets to make them AI ready
and easily ~~interoperable~~ findable.

People and organisations:

Who is it targeted at? Researchs trying to do analysis across multiple datasets.
(academic + commercial)

Who needs to be involved in developing this solution?

Black Leel by Government / HDRS
PPIE input at appropriate points, - Research community consultation
Repositories: SAIC, CRUK, Biobank, - Clinical community
NHS. - Data professionals.
- Specialist / experienced organisations: HDRUK etc.

Details of solution:

What, where, how, resources needed, cost, timeframe (short-term or long-term project)

Funded by ~~HDR~~ HDRS \Rightarrow EIS in if you build on HDRUK gateway

- Metadata
- Development of standards + agree it
 - How to regulate / mandate data input
 - Centralised organisation to lead this: HDRUK / HDRS, informed by consultation
 - Start small w. MVP \Rightarrow metadata dictionary
 - mandate DOI for publication: 710 no for dataset
- \rightarrow actions governed by this budget.
- data curation at scale (prioritised)
 - creation of metadata dictionary
 - staff costs for standardising
 - infrastructure maintenance
- Curation
- Prioritisation of key datasets
 - Funding of staff for curation
 - Development of automation tools

Funding pot for showing utility

- Research call to test out the value
- Impact assessment / review of success.

Activity 3 - Solution Design

Challenge:

Risk assessment and governance misalignment

Initial brainstorming of potential solutions:

Collaborative due diligence
risk assessment framework

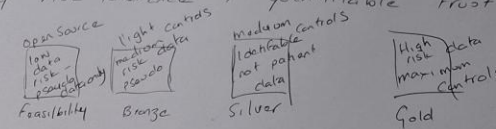
Putting more flesh onto the
5 Safes Framework

A Framework that sectors can agree to
on behalf of their user groups, which
can be cascaded for use and stops
extra local due diligence with
incongruent results

New Solution name: Low DARC - Low Data Access and ^{Risk} Calibration

GROUP: 3

Value proposition statement: Calibration tool to support access to data
A system that identifies data items by risk tolerance + quantifiable trust metrics.
An all sectors 'register'



People and organisations:

Who is it targeted at?

TRC Providers, Data owners / DPO / IGA / Legal / Caldicott Guardians / Ethics
Researchers, Risk managers / owners

Who needs to be involved in developing this solution?

All of the above

Details of solution:

What, where, how, resources needed, cost, timeframe (short-term or long-term project)

1. All sectors sharing
~~Discovery / Feasibility System that includes low risk data~~
Better Catalogues, Synthetic data, metadata

2. Discovery / Feasibility system that includes low risk data



Activity 3 - Solution Design

Challenge:

CULTURAL AND MULTIDISCIPLINARY
BARRIERS

Training:
 - already @ uni
 - for PIs who apply for funding
 - have so sort of legal
 - Scientific knowledge training
 for legal team PIs etc
 PIs - multidisciplinary activities
 → improve communication and

- Knowledge reps across
crps and disciplines
- Shared dictionary
of terms and
definitions
- "Facilitators" "Translators"
figures in each team
(role)
→ professionalised
and recognised!

Dictionary
of
Data
Definition
Language
Explanation
It's a doddle
with
"DODDLE"
training for
data leadership
- upskilling data scientists
who are willing to be the
future leaders within the
organisations

Culture of
explaining concepts
asking
questions.
training in the job
collectively

Build an organisational
culture that allows/
welcomes people to clarify
terms during meetings

UMBRELLA LEADERSHIP
 (PI, data, domain experts etc)
 at a levels of variants
 (1) Each member also belongs to
the category of experts (exp.
data scientist to a data group
division to a clinical group etc)
 (2) Formalised protocol for PIs
legal sustainability
 data career and/or roles

New Solution name: UMBRELLA LEADERSHIP

(United Management for Better Research Excellence with Long Lasting Activity) 4

GROUP:

Value proposition statement:

A change of culture that enables interdisciplinarity,
better science, talent retention and impact. With EDI and
collaboration at its core, it delivers change in a equitable and
effective manner by promoting shared language and understandings

People and organisations:

Who is it targeted at? Funders and applicants → Pilot new approach: consult, implement, give feedback
potentially deploy widely.

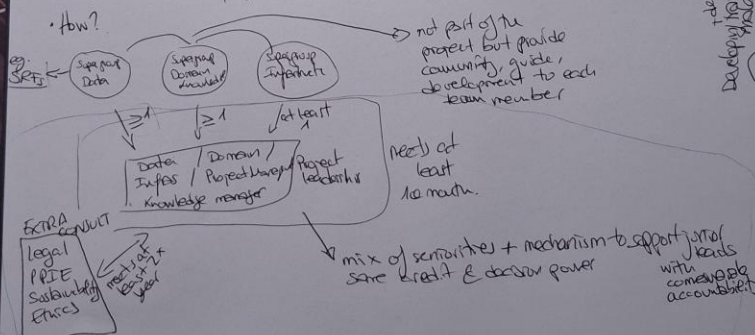
Who needs to be involved in developing this solution?

- Reducing short term
contracts & career
precarity
- Recognises diverse
roles & contributions
- Research Quality
- Career
- Accelerating science
- Efficiency
- Higher inclusivity
- Collaboration across roles
with shared language

Details of solution:

What, where, how, resources needed, cost, timeframe (short-term or long-term project)

- What: Pilot
new project leadership approach
- How?



- Change on leadership needs
- Allocation for knowledge manager
- budget for training / development on teamwork + cross-disciplinary transition

Tools
 DODDLE
 Dictionary of Data
 Definition language and
 Explanation
 (LW) facilitated
 knowledge sharing
 platform

- HR costs
- change cost from funders

financing collaboration and data sharing agreements in biomedical AI research workshop

Activity 3 - Solution Design

New Solution name: **E-CLAIRe**
(open source)

GROUP: 5

Challenge:

AI Contract

~~current agreements do not take into consideration~~

Value proposition statement:

E-CLAIRe

~~Ethical Contract Library for AI Risk Evaluation~~
Ethical Contract Library for AI Records Risk Evaluation

People and organisations:

Who is it targeted at?

Anyone who is developing a contract.

* Legal

* Programme Management

* PPIE

* Data Protection

* AI regulators

* Data providers

* Researchers

* Funders

* Private Organisations

Who needs to be involved in developing this solution?

* Research Council / funders

* Researchers / Data Scientists for the development of the AI solution

* Regulator: AISI / Alan Turing Institute / Innovate UK / HDBRI

* Communication / Marketing teams

* Legal teams

* Web developers

* UX developers

* Illustrators

Details of solution:

What, where, how, resources needed, cost, timeframe (short-term or long-term project)

3-year plan (National)

Design and Development

* What:

- Data Collection

* Where:

National body (or International)

* How:

- Research Council (initial funding)

- National Institute with Scientific expertise

- Regulator sandbox

* Resources needed:

- Human

- Funding

- Writing parties to upload historic contracts

-

* Costs:

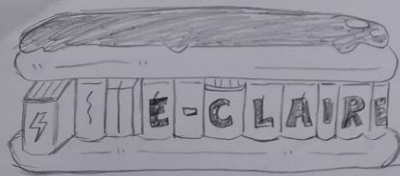
- Data Protection x1

Deployment

- governmental server to deploy the website + LLM use.

- training platform within the website.

-



"Illuminating AI Contracts"

- Costs**
- * Data Protection
 - * Data Scientists x2
 - * Legal Advisor
 - * Project Management team
 - * web developer - UX
 - * Comms team
 - * Government server
 - * Illustrators

Activity 3 - Solution Design

Challenge: 6

- Fear
- Misinformation
- Mistrust
- Responsible use of Administrative data

Initial brainstorming of potential solutions:

- Engagement events
- Scoping review of fears, feelings & the future
- Community level, faith organisations
- Schools, GPs, charities around victims/peeps.
- Trusted body (CW).
- Advocates / spokesperson
- Patient stories / real use cases - positive public good.
- AI developers <=> public engagement [interface]
- Government mandated principles - data ethics framework
- Behavioural scientists
- Framework for public engagement
 - ↳ themes, evidence, transparency, claims.

New Solution name: CLEAR ^{citizen} Community Learning + Engagement in AI Research

GROUP: 6.

Value proposition statement:

People and organisations:

Who is it targeted at?

Anyone developing AI solutions from unconsented data.

Who needs to be involved in developing this solution?

- Co-creation; representatives from UK nations, diverse communities, sectors, ^{UPD} Ada Lovelace
- Commissioners; governmental representatives, DSIT, UKRI; charities / advocacy groups ^{DAREUK / DORUK / FEDAI}
- Build on existing research & work ^{Finches / NHR}

Details of solution:

What, where, how, resources needed, cost, timeframe (short-term or long-term project)

- ↳ Co-created framework for building public trust in AI methods, research & development.
- ↳ Independently driven, research focused.
- ↳ Government mandated

↳ Existing governance bodies & structures.

↳ Highlighting patient stories - children → adults.

Budget - 3+ years [LONG TERM]

- Honoraria
- Salary costs ^{→ social scientists}
^{→ grant researchers}
^{→ regional ambassadors}
- Travel ^{→ PM}
- Event costs
- Comms & marketing
- Transcription
- Merch
- Project evaluation
- Dissemination

Activity 3 - Solution Design

Challenge:

Skills, training + workforce development

Initial brainstorming of potential solutions:

- Co-create. *bring together relevant stakeholders*
- Setting up competencies profiles
- Adding AI specific elements
- Training + knowledge exchange programme.
- Toolkit development.
- Interventions that fit in with existing UKRI training programmes.
- Training policy makers / higher up than research leads.
- *6 months work*
Reverse mentoring.

New Solution name: Bridging AI understanding. (BAU)

• New business as usual.

GROUP: 7.

Value proposition statement:

→ tailored value proposition for each target group

People and organisations:

Who is it targeted at?

- Grant Managers, Programme Managers
- Data Protection / Legal / Decision makers

Paired with

AI researchers → trained to communicate effectively.

Who needs to be involved in developing this solution?

Competencies

- Key Biomedical AI research orgs.
- Different roles.
- + focus groups
- legal
- data protection

Training / mentoring scheme.

Same people.

→ learning + ~~loss~~ skills specialist.

Details of solution:

What, where, how, resources needed, cost, timeframe (short-term or long-term project)

- Competency profile development
- Sprint.
- starting from existing profiles.
- Persona development
- needs analysis
- ⇒ from workshop by relevant stakeholders.

⇒ Funding call for development of training → from needs analysis

Description of target group

AI researchers:

- indicators what is required
- a couple of AI projects rather than job title
- = competence
- community nomination / organisation suggestion

⇒ Reverse mentoring.

Data Protection / Legal / Decision makers:

- tangible outcomes
- target places / activities where they are
-

