

Advancing Collaboration and Data Sharing Agreements in Biomedical AI Research Workshop

Filled worksheet photos - Activity 1: Imagining Utopia

Activity 1 - Imagining Utopia

GROUP: PALADINS 1

A perfect **data sharing** agreement is one that...

is STANDARDISED ^{HAS} CLEAR DEFINITION ^{ACCESSIBLE} LANGUAGE

HAS CLARITY ON USE, PARAMETERS OF USE

INCLUDES TERMS OF SECURE ENVIRONMENT

This enables: A POSITION STATEMENT ON WHAT THE DATA

QUICKER RESOLUTION, MORE STREAMLINED IS

FLEXIBILITY

A perfect **collaboration** agreement is one that...

This enables:

Something concrete that makes it work

ACCREDITED TRAINING FOR ALL

STANDARD TEXT ROLES & RESPONSIBILITIES

Something concrete that makes it work

WHO makes this work?

CONTRACT MANAGERS RESEARCH GOVTS

DATA ACCESS COMM, TEE REGULATORS

... who benefits the most?

ALL PARTIES, DATA HOLDERS

DATA USERS, RESEARCHERS, PATIENTS

WHO makes this work?

... who benefits the most?

in Turing
ute

EMBL-EBI

Advancing collaboration and data sharing agreements in biomedical AI research works

Activity I - Imagining Utopia

Paladin
GROUP: II

A perfect data sharing agreement is one that...

- short, clear agreement & process
- clear precedent (e.g. suppression) relevant to AI
- clear roles and responsibilities
- doesn't use legal/technical jargon to exclusion
- practical and realistic risk assessment / approach
- modular → activate clause if relevant (suppl. components)
- PPIE involvement
- should only cover DS considerations (not CA & contract)

This enables:

→ Culture change

also applies to

A perfect collaboration agreement is one that...

- clear distinction DS, CA & contract
- built to work together (DS/CA/C)

This enables:

→ focussed discussion on purpose of the document



Something concrete that makes it work

- Understanding what groups sign up to
- Relevant training & awareness for both researchers & legal staff
- Visual representation of the ecosystem (e.g. governance, data)
- Breakdown of who does what, when
- Core agreement (IP, ...) with suppl. components that needs discussion

Something concrete that makes it work

- Starting point like "Brunswick" but with industry considerations.

WHO makes this work?

- "Someone" who has the authority to set precedent.
↳ community-driven
- Someone in the organisation needs the overview of DS/CA/
↳ is this the decision maker? (upskilling required) contact
- Authority figures (e.g. Ewan - EMBL) writing up opinion piece with precedences

... who benefits the most?

- patients / public + society in general
- data holders (organisations) → also reduces chance of reputational damage

WHO makes this work?

- Funder
- Host organisation ("data controller")

... who benefits the most?



Advancing Utopia

GROUP: 3

A perfect **data sharing** agreement is one that...

main part should
not exceed 2 pages
(=summary)

This enables:

use easy language
(e.g. govUK standard/
a nine year old can
understand)

*Generate Revenue/Payment
- benefits for different
stakeholders, internal & external
→ good or bad?
- difficult to tease who gets
access

Allow scalability
across projects

TREs allow
Scalability

WH (data protection is important
but so is innovation -
balance)

... who benefits the most

For health-medical
data the patients/
public would benefit
the most

Gov-fundes
programmes to
enable this in
organisations

Experienced
Senior management
in both public
and private sectors

Researchers

Foundation of
engaging with the
original data owner
- e.g. patents

No CONFLICTS
WITH EXISTING
POLICIES
(all parties involved)

SCALABLE
and FLEXIBLE,
easy to adapt

ensures datasets
are well labelled
and high quality
to ensure
rich output.

A perfect **collaboration** agreement is one that

- Shared language,
Jargon free
- "5 year old can understand
as a standard"

has clearly
defined roles &
tasks for each
collaborating
party.

Template could
be helpful.

OR
transparency/
accountability
to the contribution
of each partner

Clarity around
IP
for any outputs
coming from the
collab

Clarity
around benefits
and
responsibilities
of each
partner

Some

makes it work

Better
Builds on a clear,
existing
template

Underpinned
by infrastructures
that enable
cross-border collab

Share definition of terms.
> established
standards

A team which
brings together
different professionals
including PPIE

Legal team
with understanding
of the
research environment

All stakeholders
involved!

Advancing

and data sharing agreements in biomedical

workshop

Activity 1 - Imagine

GROUP: 5

- Standard ethics + data sharing approval

→ accepted by one provider + grant approval at others.

A perfect data sharing agreement is one that...

- minimise complication in system
- could happen at programme level
- all use same agreement.

Research leader - All data is all over place.
→ Need data in one place → one system.

↓
not for everyone because of trust.

This enables:

→ speed + reduce cost.

Something concrete that makes it work

- build more public trust.

→ Better planning of data access / what data.

→ use cases of data use / reuse.

→ Data professional as a co-applicant. - LG.

WHO makes this work?

- funders + data providers need to work closer together - coordinate.

→ Need data stewards / professionals to look at grant applications / DMP.

... who benefits the most?

→ Researchers - faster access.

- Make sure patients are acknowledged for their data.

- Early, long term collab agreement.

→ Shared understanding of how to collab → known IP, open collab.

→ data sharing agreement
→ standard process for access.

on agreement is one that...

R/leader -

R/leader - collab approach.

→ longer term contracts for researchers.

→ data sci prof in facilities that specialise in specific areas / data.

This enables:

Something concrete that makes it work

WHO makes this work?

... who benefits the most?



Enables:

- ↳ Common "good practice" guidance / legal cover for data access decision makers & public panels
- ↳ Data curated to "composable" data models / common "wrapper" data
- ↳ Pathways for "discover" access to learn if an approach is useful / what data is needed
- ↳ Common terms for accepted additional uses (e.g. reproducibility)

A perfect data sharing agreement is one that...

- Flexible enough to allow data to be on any (suitable) system (e.g. diff systems & diff partners or federation).
- ~~Fast~~ Fast, easy agreement on sensitivity of data & suitable systems that can host it. (ideally common sensitivity tiers & system TIR/SDE standards)
- Clarity on how data can be used.
- Clarity on review points
- Transparency on how data shared & allowed onward share
- Data provenance (incl. ownership) & expenditure.

This enables: Expectations on contributions back

Activity 1 - Imagining Utopia

GROUP: 6 - Wizard

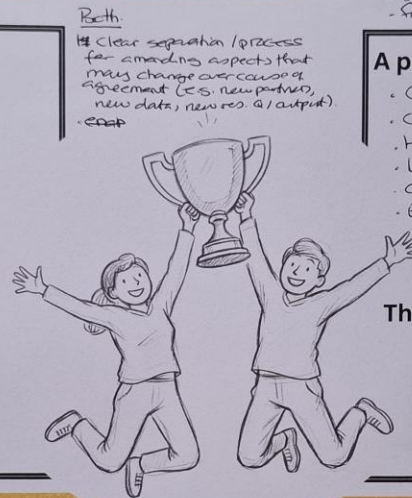
Enables:

- Sufficient compute & access
- Fair access to compute

A perfect collaboration agreement is one that...

- Clear roles, responsibility & roadmap
- Clarity on impact & value (along way, not just at end) benefits, inhibitors & patients
- How data & generated data & models can be used
- What contributes new data / IP & who can use it
- Clarity on review points
- Ensuring fair & clear & mutually beneficial outcomes & contributions from partners - makes fair access to resources (e.g. compute)

This enables:



Something concrete that makes it work

Something concrete that makes it work

WHO makes this work?

... who benefits the most?

WHO makes this work?

... who benefits the most?

Activity 1 - Imagining Utopia

GROUP: 7

A perfect **data sharing** agreement is one that...

- * CLEAR COMMON LANGUAGE
- * ALIGNMENT OF SENSITIVE DATA ACROSS JURISDICTIONS
- * PRE-AWARD PERIOD FOR NEGOTIATION

This enables:

- * TIMELY START OF RESEARCH / PROJECT
- * PROTECTION ACROSS ALL GROUPS (incl. organisation, researchers etc.)



A perfect **collaboration** agreement is one that...

- * CLEARLY SET EXPECTATIONS
- * CLEAR, COMMON LANGUAGE
- * ALLOWS ADAPTABILITY AND FLEXIBILITY
- * PRE-AWARD PERIOD FOR NEGOTIATION

This enables:

- * TIMELY START OF RESEARCH / PROJECT
- * PROTECTION ACROSS ALL GROUPS

Something concrete that makes it work

- * DATA SHARING LITERACY TRAINING ACROSS ALL JOB LEVELS
- * ACCOUNTABILITY
- * CLEAR CONTACTS

Something concrete that makes it work

- * ACCOUNTABILITY
- * UNDERSTANDING OF WHAT COLLABORATION INVOLVES (e.g. culture)
- * CLEAR CONTACTS

WHO makes this work?

- * DATA PROTECTION TEAM
- * LEGAL
- * PROGRAMME MANAGEMENT
- * RESEARCH TEAMS (SCIENTISTS)
- * DA

... who benefits the most?

- * DATA SUBJECTS
- * PUBLIC

WHO makes this work?

- * LEGAL
- * OPERATIONAL TEAM
- * SCIENTIFIC TEAMS

... who benefits the most?

- * THE PROJECT TEAM



List of groups for activity 1:

- Group 1 – Paladins 1 - Data and AI Governance
- Group 2 – Paladins 2 - Data and AI Governance
- Group 3 – Ambassadors (Stakeholder, Public and Patient Engagement, Partnerships) combined with Group 4 – Dwarfs & Seers (Data Infrastructure, Data Analysis & Wrangling)
- Group 5 – Leprechauns & Navigators (Funding & Policy, Programme Management)
- Group 6 – Wizards (Data Science and AI research)
- Group 7 – Navigators (Programme Management)

Group 4 was merged with group 3 for this task due to some late arrivals at the start of the day, so no separate - Group 4 – Dwarfs & Seers (Data Infrastructure, Data Analysis & Wrangling).