

DARE UK Community Groups: Strengthening collaborations, driving impact

Monday, 24 February 2025









Agenda for today



Agenda Item	Details
Welcome and Introduction (5-10 mins) – DARE UK	Brief introduction to DARE UK Community Groups initiative and where we are now.
Lightning Talks from Interest Groups (~50 mins total)	3-4 mins for presentation + 2 mins for questions for each group. Presentation order:
	 Statistical Disclosure Control - Reducing Barriers to Outputs from TREs (SDC-REBOOT) UK TRE Community - Community Management and Engagement Public Engagement and Data Research Initiative (PEDRI) UK TRE Community Information Governance Interest Group Best Practice for Research Imaging (BESTIES) Standardised Architecture for Trusted Research Environments (SATRE) Secure Data Access Professionals (SDAP) International Secure Data Facility Professionals Network (ISDFPN)
Lightning Talks from Working Groups (~12 mins total)	3-4 mins for presentation + 2mins for questions for each group Presentation order:
	 Synthetic Data Working Group PEDRI PIE Good Practice Standards Evaluation and Adoption Working Group
Closing Remarks and Call to Action (5 mins) – DARE UK	









House rules

- Please remain muted at all times unless invited to ask a question during the Q&A session.
- To ask a question during the Q&A session: please use the 'raise hand' function if you wish to ask a question out loud; or you can ask your question via the chat box before and during the Q&A session.
- During the discussion, please remain respectful, do not interrupt others and ensure any questions or comments remain on-topic.
- Please keep your **video** off during presentations you are encouraged to turn your video on during the Q&A/discussion session.
- The webinar is being recorded.









Introduction to the DARE UK Community Groups initiative











Vision

For all research and innovation to benefit from seamless, secure use of diverse sensitive data, at a pace, efficiency and scale approaching that of the open data ecosystem, that revolutionises research productivity and accelerates research to deliver public good.

Mission

To put the UK at the forefront of sensitive data research and innovation by assembling the tools, technologies and standards needed to streamline secure data linkage and use.









DAREUK

Phase 1: Design & Dialogue

July '21 - March '24

£7.5m



Phase 2: Build, Test & Establish

August '24 - March '27

£18.2m







DARE UK Community Groups (The 'what')



Modelled on the <u>Research Data Alliance</u> community groups: open invitation for people to submit proposals that align with the DARE UK vision and objectives

Community Charter



Common, specific purpose

Code of conduct



Clear rules and expectations

Co-Chair leadership



Identifiable, motivated leadership for community coordination

Open publishing



Transparent







Different kinds of DARE UK community groups for different functions...



1) Interest Groups











Different kinds of DARE UK community groups for different functions...



1) Interest Groups

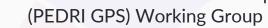














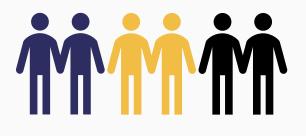




Different kinds of DARE UK community groups for different functions...



1) Interest Groups



















Current DARE UK community groups: 2023

















DARE UK community groups: 2025































PIE Standards Evaluation and Adoption (PEDRI GPS) Working Group



Community Group Lightning Talks









Statistical Disclosure Control – Reducing Barriers to Outputs from TREs (SDC-REBOOT)













developing and **sharing** and best practices in Output Statistical Disclosure Control ...

...to ensure **research outputs** (especially AI) **protect private data** and keep it secure



Building Capability and Capacity



Themes

Peer support for adoption

Theory and standards development and dissemination

Risk assessment of ML models

Activities

Monthly virtual meetings

Website development(s)

F2F meeting in September

Partners

DARE Communities of Interest

TREvolution







Data services – not just TREs

International partners



mailto:sdc-reboot@jiscmail.ac.uk



UK TRE Community – Community Management and Engagement Interest Group













The UK TRE community:

DARE UK Community Groups: Strengthening Collaborations,
Driving Impact

Pete Barnsley (UK TRE co-chair, Francis Crick Institute)

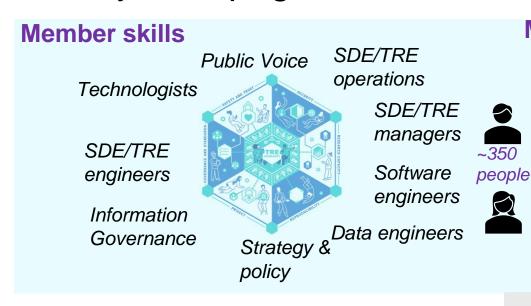
Current Co-chairs

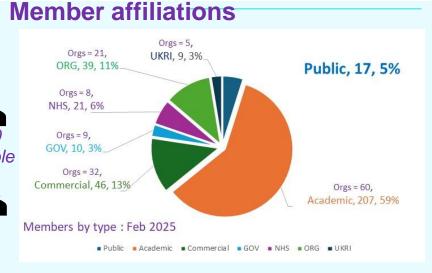
Pete Barnsley, Francis Crick Institute Madalyn Hardaker, Kings College London David Sarmiento, Alan Turing Institute Simon Li, Dundee University 24th Februrary 2025

The UK TRE Community on a page









~320 Sep people 2024

Sep

Conference: Accelerating secure digital science through community collaboration

~³⁵⁰ Jan people 2025 Dec: Engaging the Public in moving to a Federated research world

Jan: Funding & Sustainability – what we need

Mar 26th: Information Governance for federated UK research

Jun: How to fund the federated UK research eco-system Jun: Technical deep dive into federation options

Sep: TRE Community Conference: with other communities

Seven Working groups

Funding and Sustainability – how to fund

Citizen Agency –ensure public choice

SDE / TRE Definition – how they differ

Extending Control (Information Governance)

Glossary - what words mean

Cybersecurity Risks

SATRE – specification for TREs



Thank You!

Please feel free to join (your Community)

Get involved



Website uktre.org



Repository github.com/uk-tre



Mailing list
UK-TRECOMM@JISCMAI
L.AC.UK



Slack channel



YouTube

https://www.youtube.com/ch annel/UCd7AZjyH33aCmloj GHMfqEg



Public Engagement in Data Research Initiative (PEDRI) Interest Group









Public Engagement in Data Research Initiative/ DARE Interest Group

Doreen Tembo (Strategic Lead)



PEDRI Strategic Objectives



Collaboration and Partnerships

Good PIE Practices

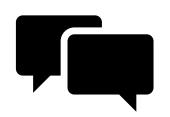
Two Way Engagement and Communication

Learning and development

Culture of PIE





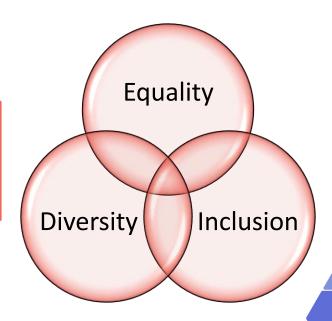








"It's not about getting people to trust you; it's about demonstrating trustworthiness" (Public Contributor)



Key Achievements

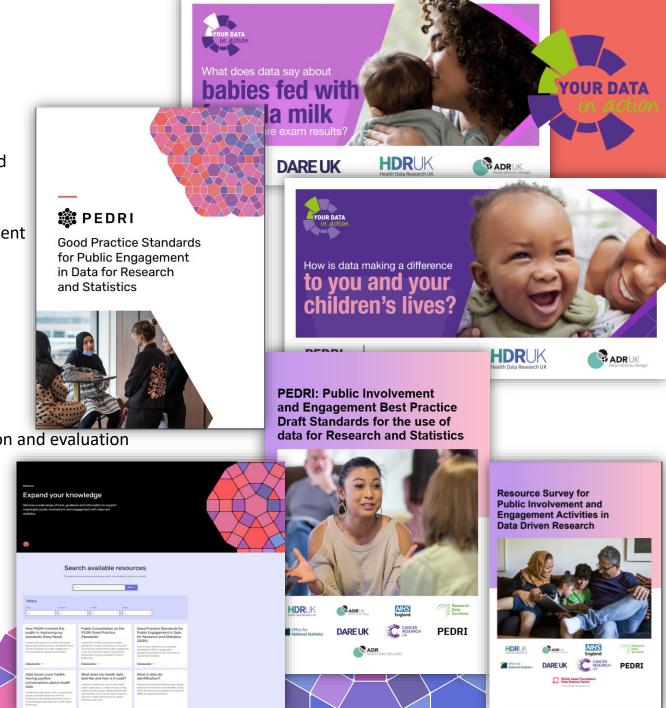
Phase I

- Building a partnership and precedence for collaborative working in the field
- Developing PEDRI Good Practice Standards
- Identifying community needs and requirements for learning and development
- Pilot communications campaign focused on families
- Delivering a programme of community activities

Phase II

- Strategic expansion and community development
- Good practice standards: public pledge, adoption, funding call, accreditation and evaluation
- Website and resource hub
- Thought leadership and community approach to challenges
- co-developing resources/learning and knowledge-sharing events
- Cross-data domain research engagement and communications









In the beginning....





- Convened by HDR UK in Feb 2019, the Alliance was established through funding from Innovate UK as part of the Digital Innovation Hub (DIH) programme, to develop tools, processes and capabilities to help organisations unlock data-driven research and innovation.
- We launched with 9 founding members:





















What are we trying to achieve?



Our mission is to accelerate improvements in biomedicine, health and care by encouraging widespread and responsible use of health-relevant data for research and development.

- CONVENE
- SHARE KNOWLEDGE
- REDUCE DUPLICATION
- ALIGN



- AGREE ON BEST PRACTICE
- CO-DEVELOP STANDARDS
- PUBLISH RECOMMENDATIONS
- SHARE RESOURCES

- SUPPORT ORGANISATIONS
- FACILITATE IMPLEMENTATION
- INCENTIVISE CHANGE

Current Alliance workstreams and working groups HDRUK



Trust & Transparency

- Pan UK Data Governance Steering Group and Action Forces
- Public Engagement in Data Research Initiative (PEDRI)
- Researcher Registry Information Governance & Policy Working Group

Data Standards

- Data Officers Group
- OMOP Special Interest Group
- Diversity in Data Special Interest Group
- Clinical Trials Key Stakeholder Prioritisation Forum

Technology Services

- Co-development of the 'Enhanced Gateway'
- Contribute to various communities convened by DARE UK

Capacity Building

- Black Internship Programme
- Support to development of new training and resources on the HDR UK Futures Platform

International communities with Alliance representation

QUANTUM – EU Horizon Programme

- **Background:** Data holders are expected to make their datasets available for secondary use, providing a notion of their quality and utility and the maturity of their data quality procedures. Article 56 proposes this would take the form of a label
- Goal: Developing and implementing a label mechanism that could be ideally adopted in the future HealthData@EU

What impact has the Alliance made?



Trust & Transparency

- Responding to the COVID-19 pandemic by enabling UK wide research through harmonsied data access.
- Improving transparency in data use through the <u>Data Use Register Standard</u> and more recently access processes through the <u>Pan-UK Data Governance Steering Group Transparency Standards</u>.

Data Standards

- Supporting interoperability and useability of data from standards on paper to standardised analytics:
- Publication of <u>Data Standards White Paper</u> in 2021, supporting the <u>EHDEN UK Data Partner call</u> in 2022, establishing <u>OHDSI UK</u> and the <u>Alliance OMOP SIG</u> in 2023 and more recently establishing the <u>Real World Evidence</u> (RWE) Network in 2024.

Technology Services

- Making data **Findable, Accessible, Interoperable and Reusable (FAIR)** through the Gateway by developing a **Metadata Specification** and encouraging alliance members to publish their datasets.
- Shaping policy and best practice for Trusted Research Environments (TREs) through the White Paper on Principles and Best Practice for TREs.

Capacity Building

 Addressing the data science skills gap by encouraging Alliance members to become host organisations in the <u>Black Internship Programme</u>.



UK TRE Community – Information Governance Interest Group









HELLO FROM THE CO-CHAIRS











KATHERINE
O'SULLIVAN
UNIVERSITY
OF SHEFFIELD

AMY TILBROOK
UNIVERSITY OF

NIVERSITY OF EDINBURGH

FELIX RITCHIE
UNIVERSITY OF
THE WEST OF
ENGLAND

PETE BARNSLEY
THE FRANCIS
CRICK
INSTITUTE

YVONNE ADEBOLA PUBLIC MEMBER

WHY DO WE NEED AN INFORMATION GOVERNANCE INTEREST GROUP?



AIMS AND DELIVERABLES









PLEASE JOIN US!

TRE-IG@JISCMAIL.AC.UK

OUR FIRST EVENT: 26TH MARCH ONLINE HTTPS://LU.MA/NM27F056

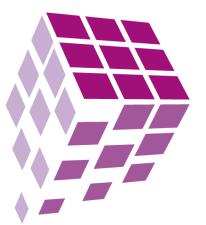


Best Practice for Research Images (BESTIES)









BESTIES

Best Practice for Research Images

Funded by DARE UK Community Interest Groups 24-25

Group Info Page: https://dareuk.org.uk/community-groups-listing/best-practice-for-research-imaging/

Sign up here: https://forms.office.com/e/acmtWfhFW8

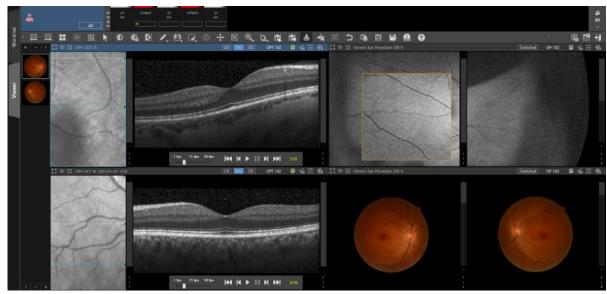


Image taken from Image Technology News: https://www.itnonline.com/article/infinitt-showcase-new-products-enhancements-himss24



Image provided by Claire MacDonald

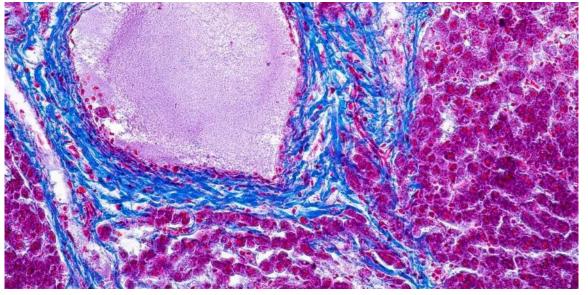


Image taken from Lab Manager: https://www.labmanager.com/integrating-artificial-intelligence-with-digital-pathology-22148



Digital Pathology

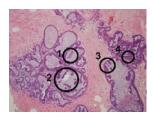


Image taken from University of Washington: https://homes.cs.washington.edu/~shapiro/digipath.html

Ophthalmology PACS



Image taken from University of Edinburgh: https://clinicalsciences.ed.ac.uk/edinburgh-imaging/research/themesand-topics/body-systems/eyes-retinal

Radiology PACS



Image taken from Radiology Café: https://www.radiologycafe.com/interviewapplication/career-in-radiology/









Short-term outcomes:

- Generate ideas from participant discussions for areas of focus
- Create at least one working group
- Develop at least one BPR
- Produce a report on the analysis of work package outputs and recommendations for areas to develop best practice recommendations and test in future workstreams

Long term aims:

- To develop WGs which will explore further areas for focus. The WGs will conduct deeper explorations of the issues and how they should be addressed.
- Use the outputs from the WGs to generate best practice recommendations.

Group Co-Chairs



Claire MacDonald
Head of Clinical Data
Science
Manchester University
NHS Foundation Trust



Dr Anthony Wilson Group Chief Medical Informatics Officer (Research & Innovation) Manchester University NHS Foundation Trust



Steve Sweeney (Public Representative)



Dr Steve Harris Chief Research Information Officer University College London



Dr Nicholas Payne Medical Physicist and Research Associate University of Cambridge



Dr Katie Wilde Head of Digital Research University of Aberdeen



Sign up here:

https://forms.office.com/e/acmtWfhFW8



BESTIES (BEST practice for research ImagES)

An Interest Group dedicated to those developing imaging pipelines for health and social care research.

Start now



Standardised Architecture for Trusted Research Environments (SATRE)







Standard Architecture for Trusted Research Environments



https://satre.uktre.org

satre-contact@dundee.ac.uk

Who benefits?

Public: Better understanding of what's happening with your data

TRE operators & developers: more sharing of knowledge and technology

Users: More consistency amongst TREs

What are we doing?

Maintaining our best-practice guide to building and operating Trusted Research Environments (TREs)

Finding a sustainable model for SATRE:

- Governance
- Formal accreditation
- Extending SATRE

Get involved



https://uktre.org



https://securedatagroup.org

Join a collaboration café



Secure Data Access Professionals (SDAP)









An Introduction to SDAP

Leigh Tate – UK Data Service







Co-chairs

Dr James Scott
UK Data Service





Leigh Tate
UK Data Service





Our purpose

To facilitate the sharing of expertise, good practice and knowledge between those working In Trusted Research Environments

To support collaboration across the Secure Data landscape, intended to drive innovation and service development



Our members

Membership from over 40 organisations, representing the day-to-day management or administration of a Safe Data setting

Working in collaboration with other Data Community networks and Interest Groups to broaden our reach and impact



Our events and resources

Quarterly online events, with presentations from guest speakers on a wide range of community-led topics

A suite of online resources (such as a Statistical Disclosure Control handbook and competency framework for staff working in a Safe Setting) available at securedatagroup.org

Register to join SDAP

securedatagroup.org/membership





Thank you





International Secure Data Facility Professionals Network (ISDFPN)













A new endorsed DARE UK Interest Group: ISDFPN - The International Secure Data Facility Professionals Network

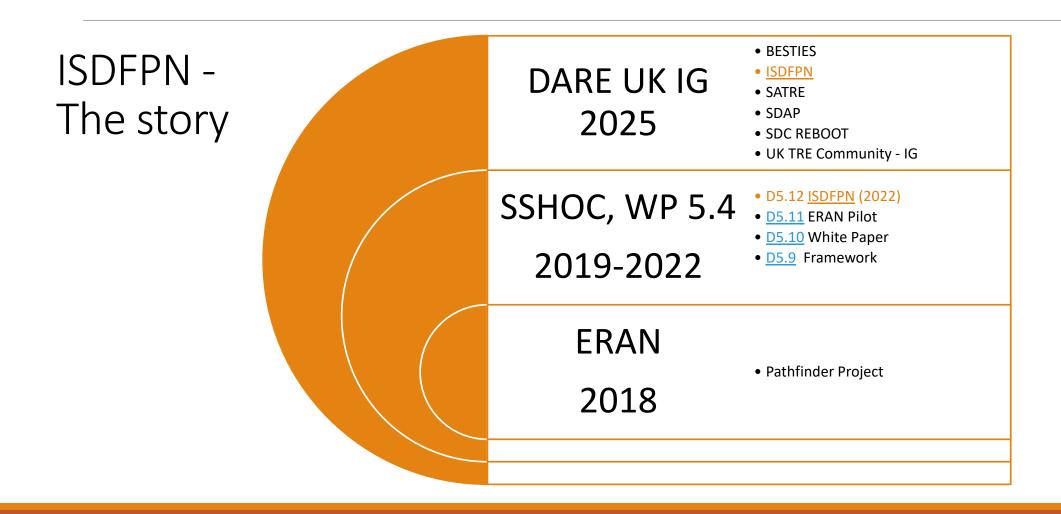
Beate Lichtwardt (UKDS), Sharon Bolton (UKDS), Kristi Winters (GESIS) Co- and Deputy Chairs ISDFPN

DARE UK Community Groups: Strengthening Collaborations, Driving Impact.
Online, 24 February 2025

















ISDFPN – Topical Network Events

- > 3 Events/year
- 2 Talks/event
- Online
- Next event:05 March 2025

Presentation 1	Presentation 2
International Secure Data Facility Professionals Network (ISDFPN)	Mind the Skills Gap: Creating Capacity for Data Access. SDAP: Competency Framework
Certifying reproducibility with confidential data: first results from the French cascad/Casd cooperation	Data Sharing with RDC Qualiservice
Enabling Access to Confidential Qualitative Data through Data Enclaves	Building a Safe Researcher Accreditation Scheme
Researcher Passport: A digital user credential for accessing restricted data	Introducing the SafePoints
SANE: an off-the-shelve, data holder-agnostic TRE	Output Disclosure Control for Qualitative Data in Trusted Research Environments: Current State and Next Steps
Setting up Remote Access to Sensitive Data (when things don't quite go to plan)	Making secure qualitative data available via TREs



ISDFPN – Some facts

K Data Service

Webviews by country

- ➤ 30 Individual members
- ➤ 20 Organisations
- > 9 Countries



- Zambia
- USA
- UK
- Turkey
- Tunisia
- Thailand
- Syria
- Switzerland
- Sweden
- Spain
- South Africa
- Russia
- Poland
- Netherlands
- Jordan
- Japan
- Italy
- Israel
- Ireland
- Indonesia
- India
- Hungary
- Germany
- France
- Dominican Republic
- Croatia
- Colombia
- China
- Cameroon
- Brazil
- Bolivia
- 5011110
- Belgium
- Argentina

ISDFPN Zenodo Community

Space

- > Event presentations
- > Other resources

Presentations and Publications

about ISDFPN

- > WCQR
- > IASSIST
- > ESRA
- UKDS Impact Blog
- > IASSIST Quarterly
- > SSHOC Zenodo







Thank you!

If you would like to join the ISDFPN, offer a talk or request more information regarding the Network, please email isdfpn@ukdataservice.ac.uk.

https://ukdataservice.ac.uk/about/research-and-development/international-data-access-network-idan/

https://dareuk.org.uk/community-groups-listing/international-secure-data-facility-professionals-network/



Synthetic Data Working Group







DARE UK Community Group

Plans for the UK Synthetic Data Community Group

Find out more: www.syntheticdata.uk



Meet the Co-Chairs



Lewis HotchkissDementias Platform UK



Emma Squires
Dementias Platform UK



Simon Thompson
Dementias Platform UK,
SAIL Databank, SeRP



John Gallacher Dementias Platform UK



Tim Rittman University of Cambridge



Emily Oliver Administrative Data Research UK



Cristina Magder
UK Data Service



Fiona Lugg-Widger Centre for Trials Research



Robert Trubey
Centre for Trials Research



Sophie McCall Research Data Scotland



Anmol Arora University College London



Steve MoorePublic Representative



Current Landscape



Lewis Hotchkiss



Emily Oliver



Sophie McCall



Cristina Magder



Fiona Lugg-Widger



Robert Trubey



DPUK/SeRP Synthetic Data Working Group



ADR UK Synthetic Data Working Group



Scottish Synthetic Data Working Group



Data providers' perspectives of low-fidelity synthetic data Project



Public attitudes to the use of low-fidelity synthetic data Project



DPUK/SeRP Synthetic Data Work







Dementias Platform UK has developed several tools and processes for the generation and evaluation of synthetic data and are now working towards the practical implementation of these in various scenarios.

Using **Synthetic Data** in Practice

Dementias Platform UK is working on a number of projects for the practical implementation of synthetic data within and outside of our Trusted Research Environment.



Creating "Synthetic Data TREs"

DPUK along with the **SAIL Databank** are developing "Synthetic TRE" versions of themselves to enable greater access to data with lower governance controls and restrictions.



Data Federation with ADDI

We are working with **ADDI** to deploy data federation to enable federated queries across TREs. Synthetic data helps in this scenario where the actual data cannot be seen.



Enabling Researcher/Student Training

We are working on releasing synthetic datasets for researcher and student training in health data research.



The Need for Synthetic Data across DARE UK Projects



Al Risk Evaluation Group



DARE UK Federation TRE-FX



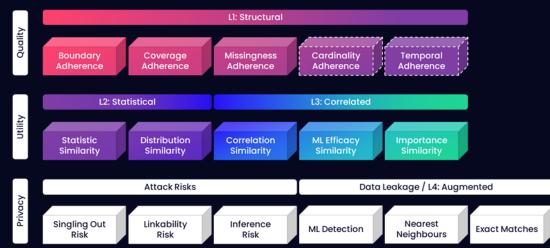
SDC-REBOOT Group (Tools for Al Disclosure Control)





DPUK/SeRP Synthetic Data Work







Aims & Objectives

- Host a series of workshops to discover solutions for the deployment of synthetic data in Trusted Research Environments.
- Develop standards for the evaluation of quality, utility and privacy in synthetic data
- Develop and refine "SynthOpt", a community tool for the generation and evaluation of synthetic data at various levels
- Create learning resources and guides to help researchers and data owners understand synthetic data and its uses
- Put together a governance framework for Trusted Research Environments to be able to use synthetic data in practice





Workshops

Researchers

Understanding user requirements and use cases for different levels of fidelity of synthetic data.

April



Experts

Developing tools and setting standards for the generation and evaluation of synthetic data.

June



Data Owners

Exploring issues such as data ownership, transparency and communication.

August



Public

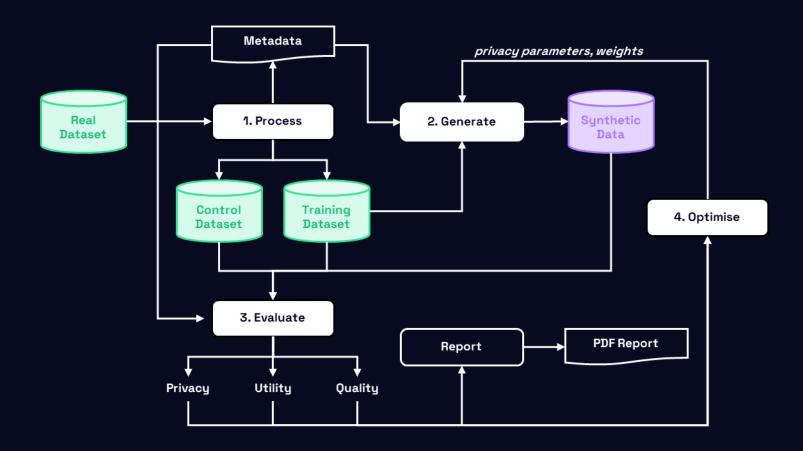
Ensuring synthetic data is used for the public good and protects privacy sufficiently.

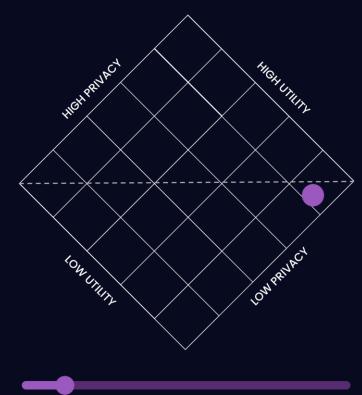
October





Community Tools







Community Tools

Synthetic Data Evaluation Report This report details the quality, privacy and utility evaluation metrics galened from the synthetic data, and visualisations to help interpret them.

Metrics Summary

		Wettics Sair	-	Utility Metrics	300.0
	Score	Quality Metrics	Buo.	Statistic Similarity	0.97
Privacy Metrics	100	Boundary Adherence	1.0	Correlation	0.91
Exact Matches	1.0		0.87		0.0
	0.24	Coverage	0.89	ML Efficacy	nan
Detection	0.04	Complement	nan	nan	nan
Inference Protection	0.16	nan	nan	nan	nan
Singling Risk	0.04	nan	nan	nan	Esser
Linkability Risk	0.05	nan	- Esseri		
Inference Risk	0.00				

Synthetic Data Categorisation Level: Correlated Synthetic Data Correlated Synthetic Data is categorised as the highest risk due to capturing information about relationships and patterns between variables. Therefore, the privacy metrics should be evaluated carefully to ensure individuals arrent at risk of being identifiable.

What are the Synthetic Data Categorisation Levels? Synthetic data exists on a spectrum of low to high fidelity depending on how similar it is to the real data. We can categorise these levels of fidelity as below: Level 3 (Statistical Synthetic Data) Roughly keeps the same summary statistics

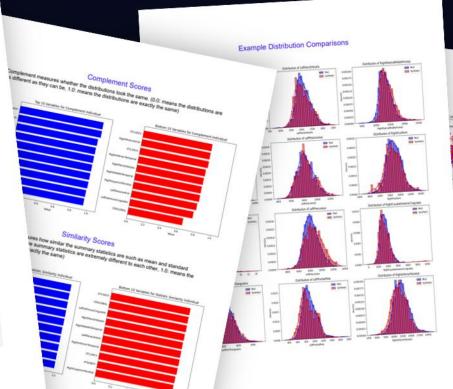
Level 0 (Random Synthetic Data) Values which are randomly generated, without using the real data or metadata.

Level 1 (Metadata Synthetic Data) Random data which is generated based on real mandom data sinch is generated based on real metadata such as data types, value ranges etc.

Generated by a synthesiser to mimic the structure of the real data, but without capturing the relationships between attributes.

such as mean, median, standard deviation etc. variables are captured.





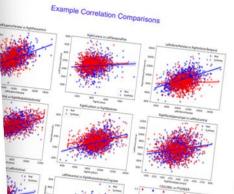


This metric measures whether each row in the synthetic data is novel, or whether it exactly matches

(best) 1.0: The rows in the synthetic data are all new. There are no matches with the real data, (worst) 0.0: All the rows in the synthetic data are copies of rows in the real data.

(Privacy) Detection

was how difficult it is to tell apart the real data from the synthetic data using is it is so one apairs are reas outs from the symmetric data using a second of 1 may indicate high quality but it could also be a clue that the symbolic data is copying the rows in the



at can be singled out by a unique combination of their

lata can help an attacker finding a combination of data. This attack evaluates the robustness of the attribute which single out an individual. gal can be singled out by a unique combination of

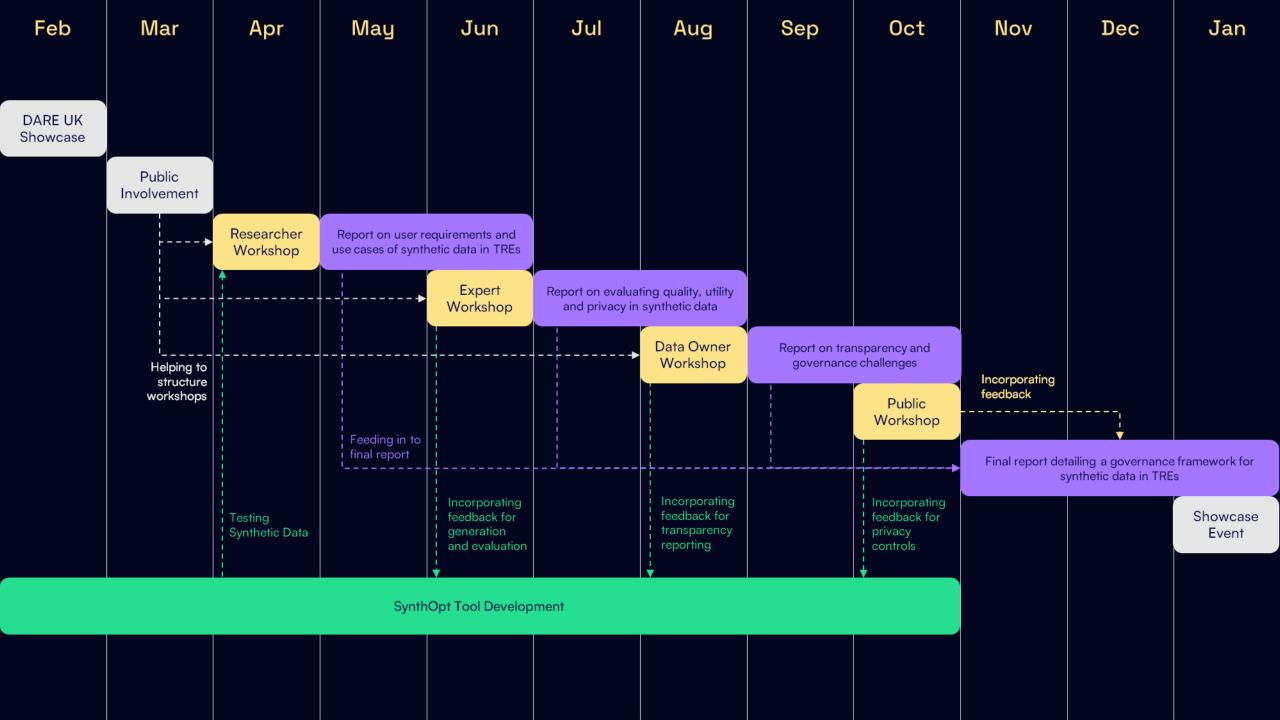
he synthetic data apart from the real data.

eatly synthetic data apart from the real

le to infer real, sensitive values. It is assumed al data; they will combine it with the synthetic . The attacker is not able to correctly guess attack. The attacker is able to correctly guess

setic data will help an adversary who tries to link two other sec data we neep an adversary who tries to tink two other. For example, suppose that the adversary finds dataset A na about the profession and education of people, and and health related information. Can the attacker use the eth? (It is assumed the attacker knows the key fields of the title.)





Upcoming Events



DEMON Network & Dementias Platform UK Synthetic Data for Dementia Research ECR Workshop

Attend expert talks and workshops on the applications of synthetic data in dementia research — from mitigating bias and privacy issues in data, to enabling federation. Plus, gain hands on experience with generating and evaluating synthetic data on datasets within the DPUK Data Portal.

Register Now!

2nd & 3rd of April 2025 Zoom, Online



DARE UK Community Group

Plans for the UK Synthetic Data Community Group

Find out more: www.syntheticdata.uk



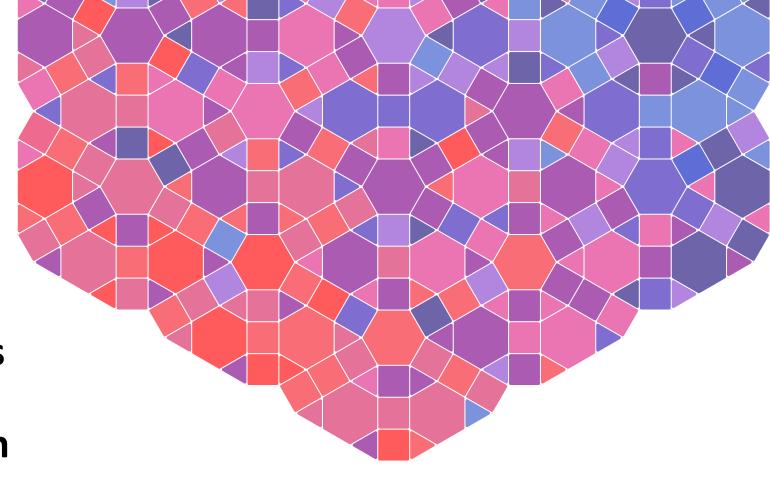


PEDRI PIE Good Practice Standards Evaluation and Adoption Working Group









Good Practice Standards

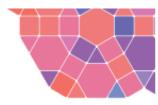
Adoption and Evaluation Working Group

Samaira Khan

Senior Partnerships and Operational Manager



Our seven Good Practice Standards for public involvement and engagement in data research





Equity, diversity and inclusion

Ensure representation of people from different backgrounds.



Data literacy and training

Empower the public to understand and contribute.



Two-way communication

Enable everyone to have open, honest, and clear conversations.



Transparency

Project information is freely accessible for discussions with the public.



Mutual benefit

There is benefit to everyone involved.



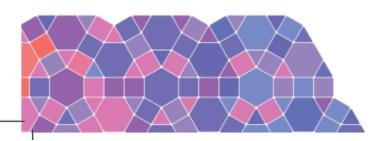
Effective involvement and engagement

There are clear tasks, purpose and impacts that are not tokenistic.



Creating a culture of involvement and engagement

Public involvement and engagement is embedded into ways of working.





amrc











HDRUK











LE Langitudinal Linkage Collaboration

NCMH







Upcoming Plans

Impact Specialist Project Officer



Manage, develop and support implementation of an evaluation tool for PEDRI GPS



Stakeholder consultations and workshops



Launch GPS Funding Call



Frontier meeting for PEDRI GPS, engaging executives and leaders in the field



Support exploration and creation of accreditation system for GPS.

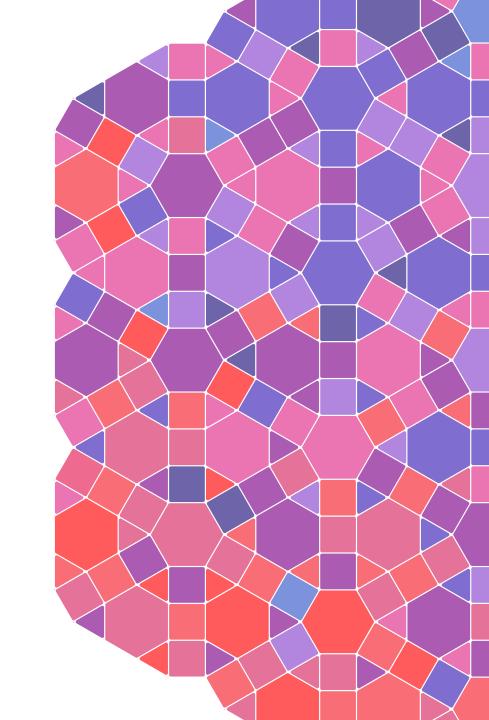


Stay Connected & Get Involved

contact@pedri.org.uk
www.pedri.org.uk
Sign up to our bulletin

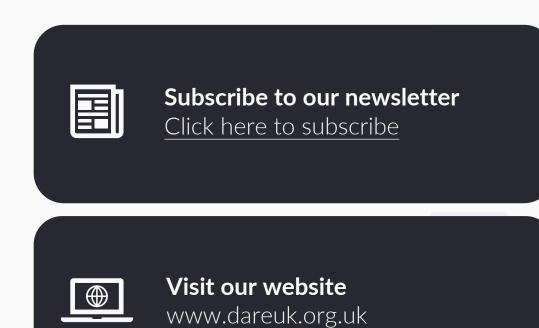


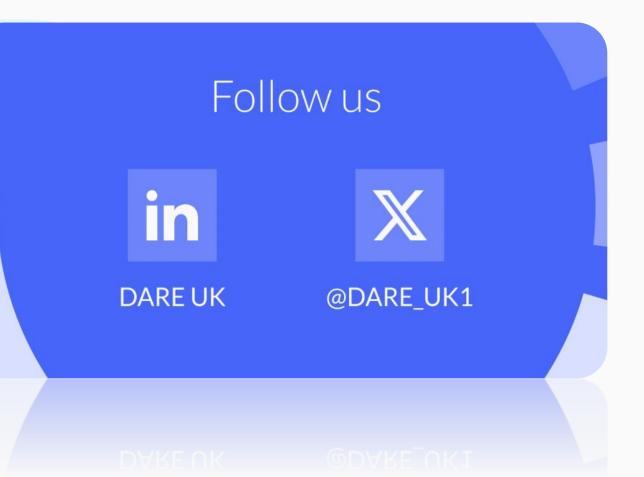






Stay in touch!













Stay in touch! - QR codes

Visit our website

SCAN ME

Subscribe to our newsletter

Subscribe to our newsletter

Subscribe to our newsletter

Follow us on LinkedIn

SCAN ME







