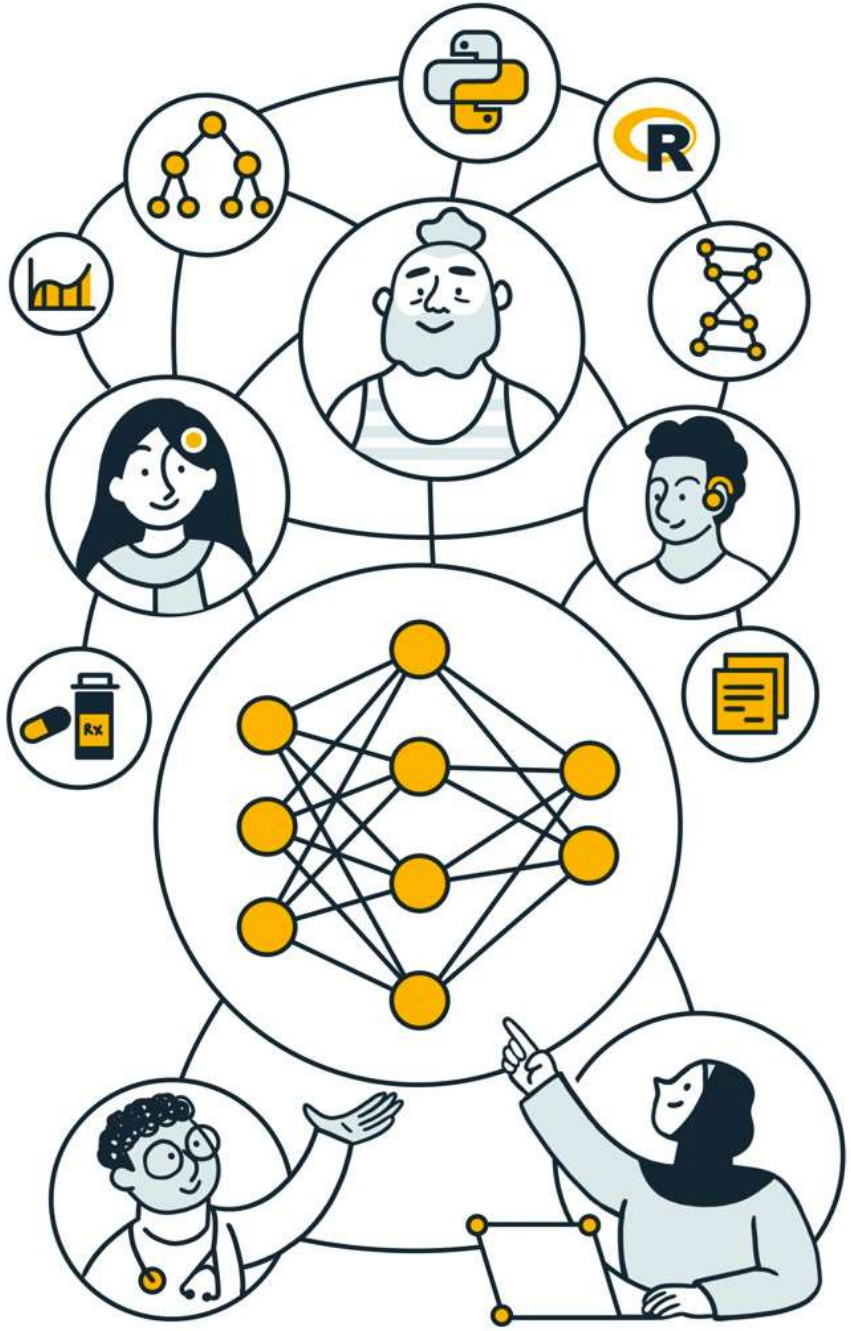
AI for Multiple Long-term Conditions
Research Support Facility

Building capacity in data science for MLTC research: the AIM RSF approach

Kirstie Whitaker

Pronouns: she/her



https://doi.org/10.5281/zenodo.10992603

Research Programme Director for Tools, Practices & Systems

The Alan Turing Institute

https://doi.org/10.5281/zenodo.10992603

AI for Multiple Long Term Conditions

- Purpose: Develop insights into the identification and subsequent prevention of MLTC
  - “Spearhead the use of advanced data science and AI methods - combined with existing methodology and expertise in clinical practice, applied health and care research and social science - to systematically identify new clusters of disease and the development of conditions over the life course.”
  - “Grow capacity and capability for multi-disciplinary working in MLTC-M for the benefit of patients, practitioners and the public.”
- In 2020-2021 the £23m call supported seven research collaborations and partnerships between leading academic institutions, health and care researchers, AI experts, and practitioners.
  - Additionally funded a Research Support Facility (RSF) to work with Research Collaborations

FUNDED BY
NIHR | National Institute for Health and Care Research

https://www.nihr.ac.uk/explore-nihr/funding-programmes/ai-award.htm

https://doi.org/10.5281/zenodo.10992603

AI for Multiple Long Term Conditions

- AIM-CISC: Analyse and understand which patterns of multimorbidity are most common, which most affect people's lives, and help improve the quality and safety of care.
  - University of Edinburgh, The Roslin Institute, NHS Lothian, and University College London.
- AI-MULTIPLY: Characterise the dynamic relationships of MLTC and polypharmacy and inform healthcare pathways.
  - University of Newcastle and Queen Mary University.
- Cluster-AIM: Develop and validate population clusters to integrate health and social care using mixed methods.
  - Southampton, Oxford, Kent, Nottingham and Leicester.
- DECODE: Mapping the challenges and requirements for Data-driven, machinE learning-aided stratification and management of multiple long-term COnditions in adults with intellectual DisabilitiEs (ID).
  - Leicestershire NHS Trust, Loughborough University, University of Leicester, and De Montfort University.
- DynAIRx: Develop new, easy-to-use AI and data visualisation tools that help GPs and pharmacists treat patients with MLTC.
  - Liverpool, Leeds, Manchester and Glasgow.
- MELD-B: Identify life-course time points and targets for the prevention of early-onset, burdensome Multiple Long-Term Conditions.
  - University of Southampton, University of Glasgow, Swansea University, Southampton City Council, University of Aberdeen, and King's College London.
- OPTIMAL: OPTIMising therapies, discovering therapeutic targets and AI-assisted clinical management for patients living with complex multiple long-term conditions.
  - University of Birmingham, University of Manchester, University Hospitals Birmingham NHS Trust, NHS Greater Glasgow & Clyde, and University of St Andrews.

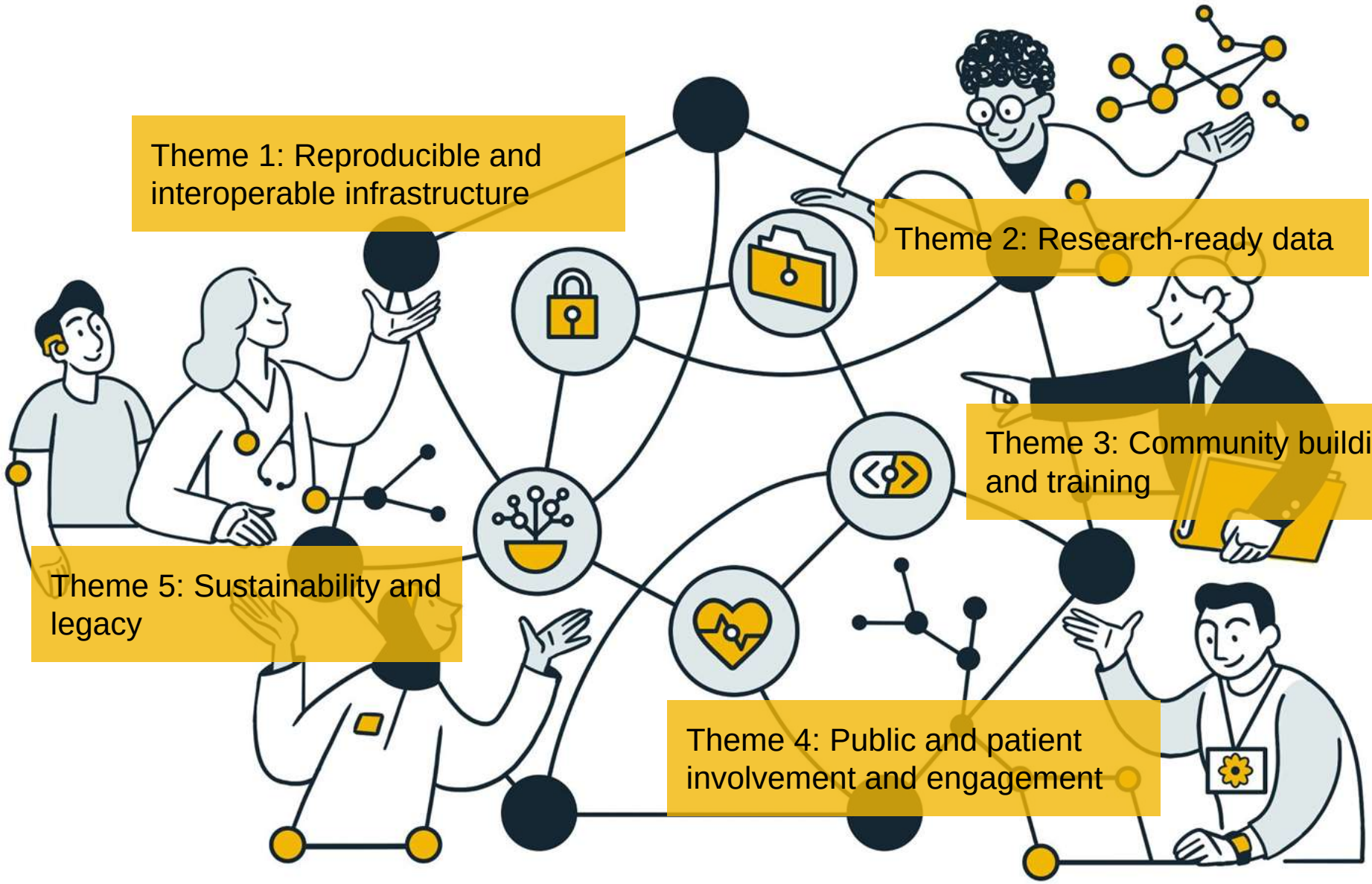
AI for Multiple Long Term Conditions: Research Support Facility (AIM RSF), AIM-CISC, AI-MULTIPLY, Cluster-AIM, DECODE, DynAIRx, MELD-B, & OPTIMAL. (2024). AI for Multiple Long-Term Conditions Handbook. Zenodo. https://doi.org/10.5281/zenodo.10867842

https://doi.org/10.5281/zenodo.10992603

Swansea University
Prifysgol Abertawe

The Alan Turing Institute

THE UNIVERSITY of EDINBURGH



https://doi.org/10.5281/zenodo.10992603

Swansea University
Prifysgol Abertawe

The Alan Turing Institute

THE UNIVERSITY of EDINBURGH

Theme 1: Reproducible and interoperable infrastructure

Theme 2: Research-ready data







Theme 3: Community building and training

Theme 4: Public and patient involvement and engagement

Programme management







Theme 5: Sustainability and legacy

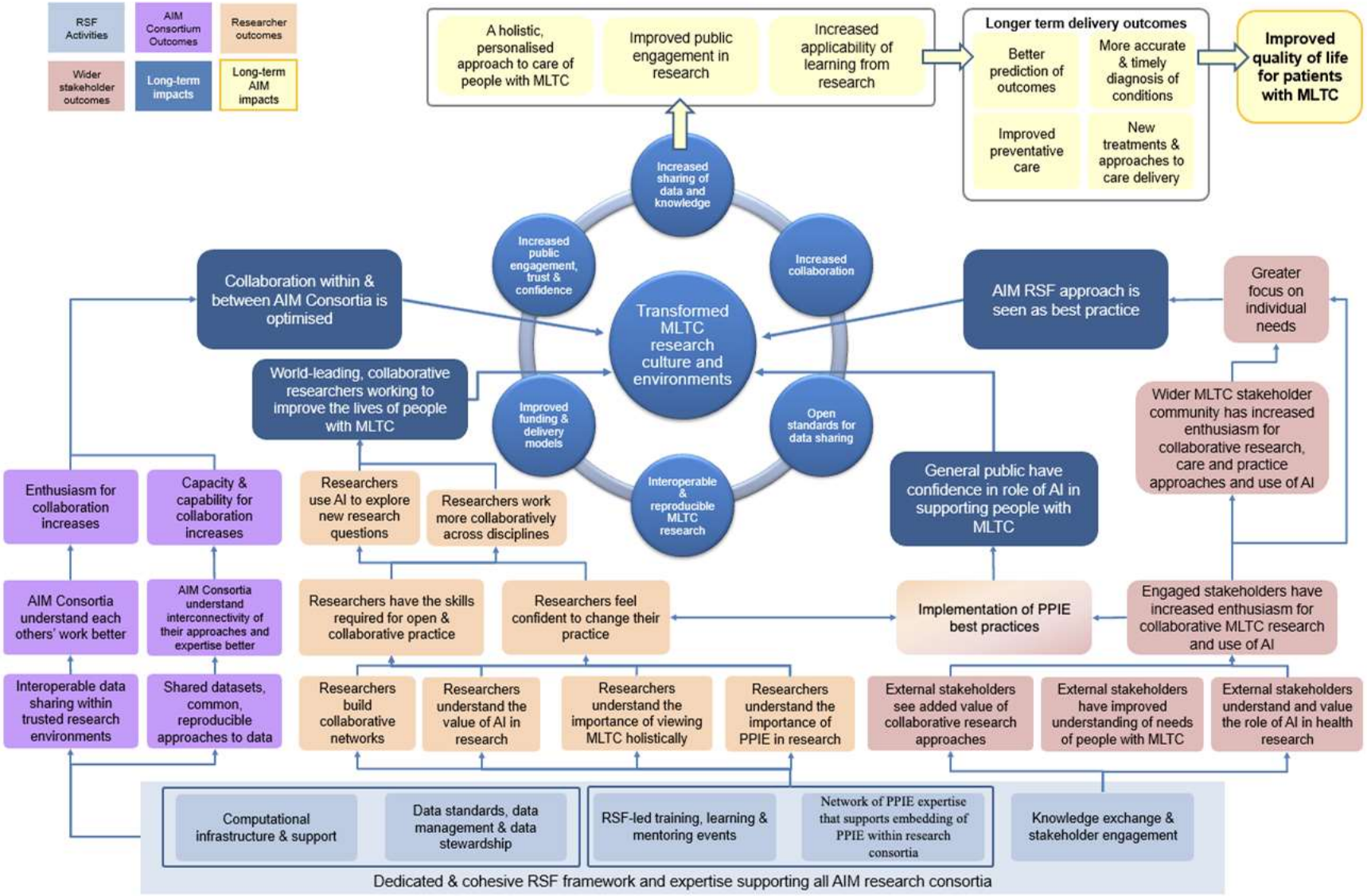


The AIM RSF Team

- David Ford
- Ronan Lyons
- Chris Orton
- Reece Labrom
- Ann-Marie Mallon
- Rachael Stickland
- Mahwish Mohammad
- Evelina Gabasova
- Emma Karoune
- Bastian Greshake Tzovaras
- Sophia Batchelor
- Aziz Sheikh
- Monica Fletcher
- Beth Collop
- Kirstie Whitaker
- Sydney Ambrose
- Batool Almarzouq

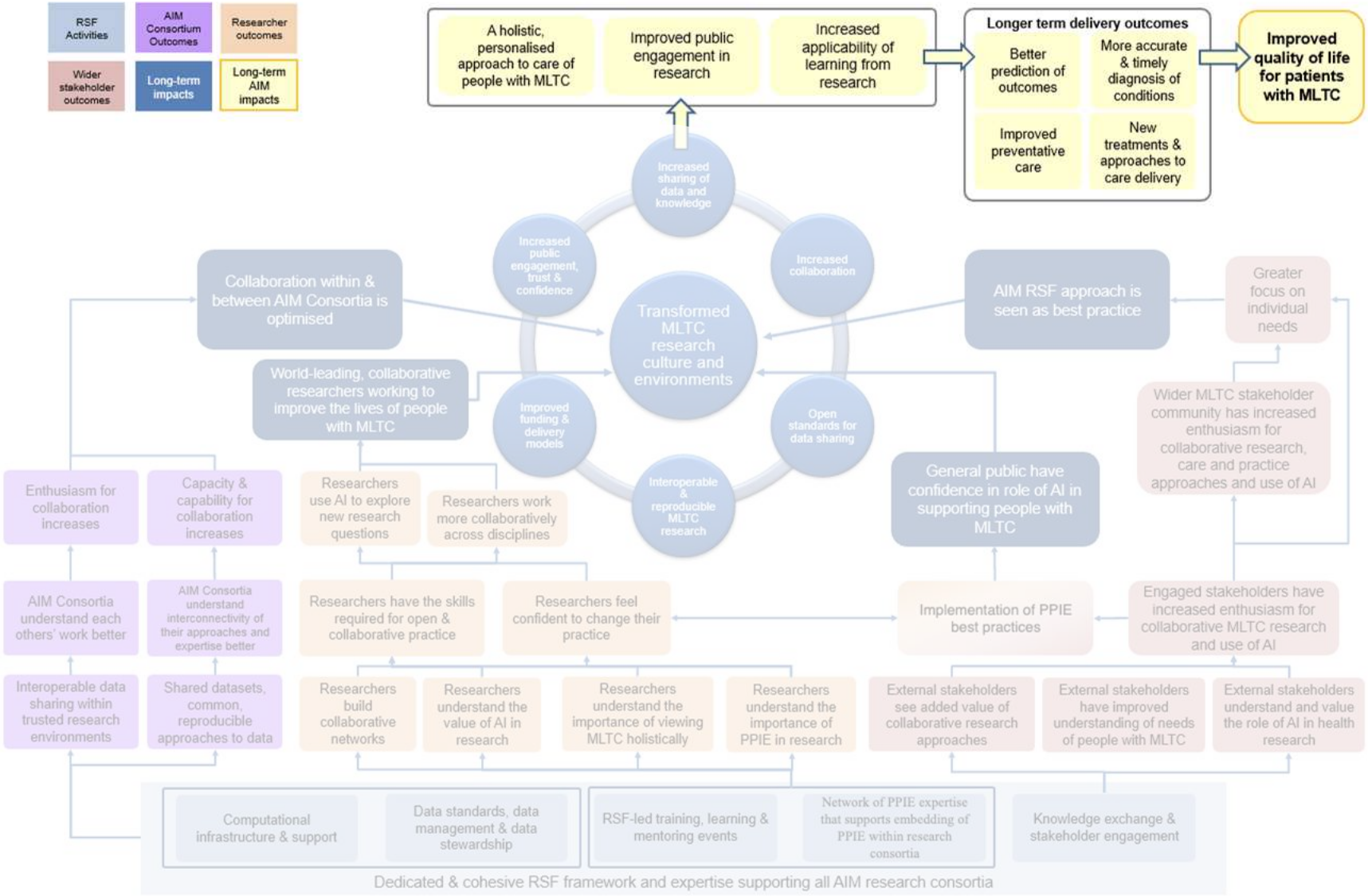
https://doi.org/10.5281/zenodo.10992603

AIM RSF Theory of Change



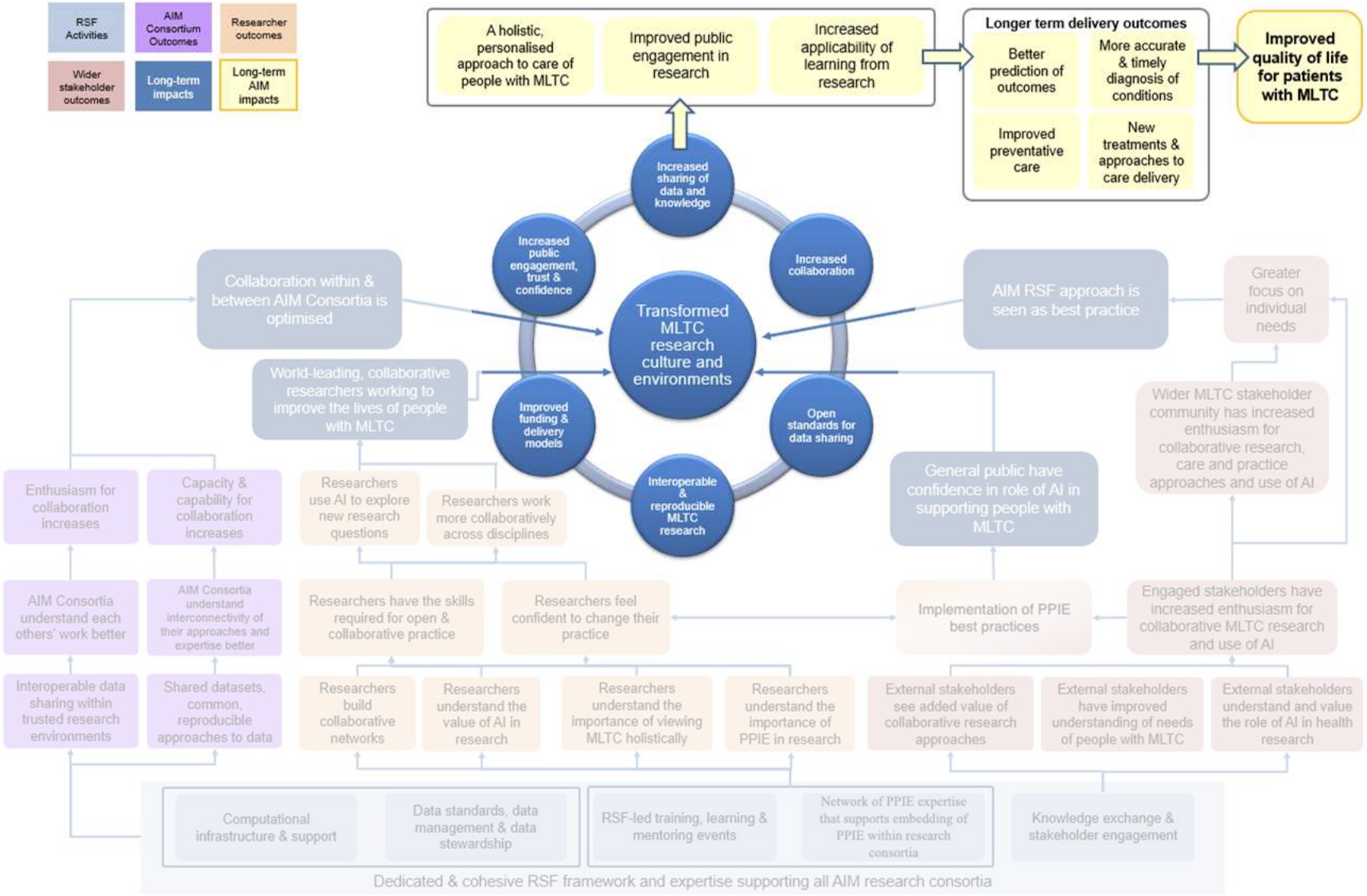
https://doi.org/10.5281/zenodo.10992603

AIM RSF Theory of Change



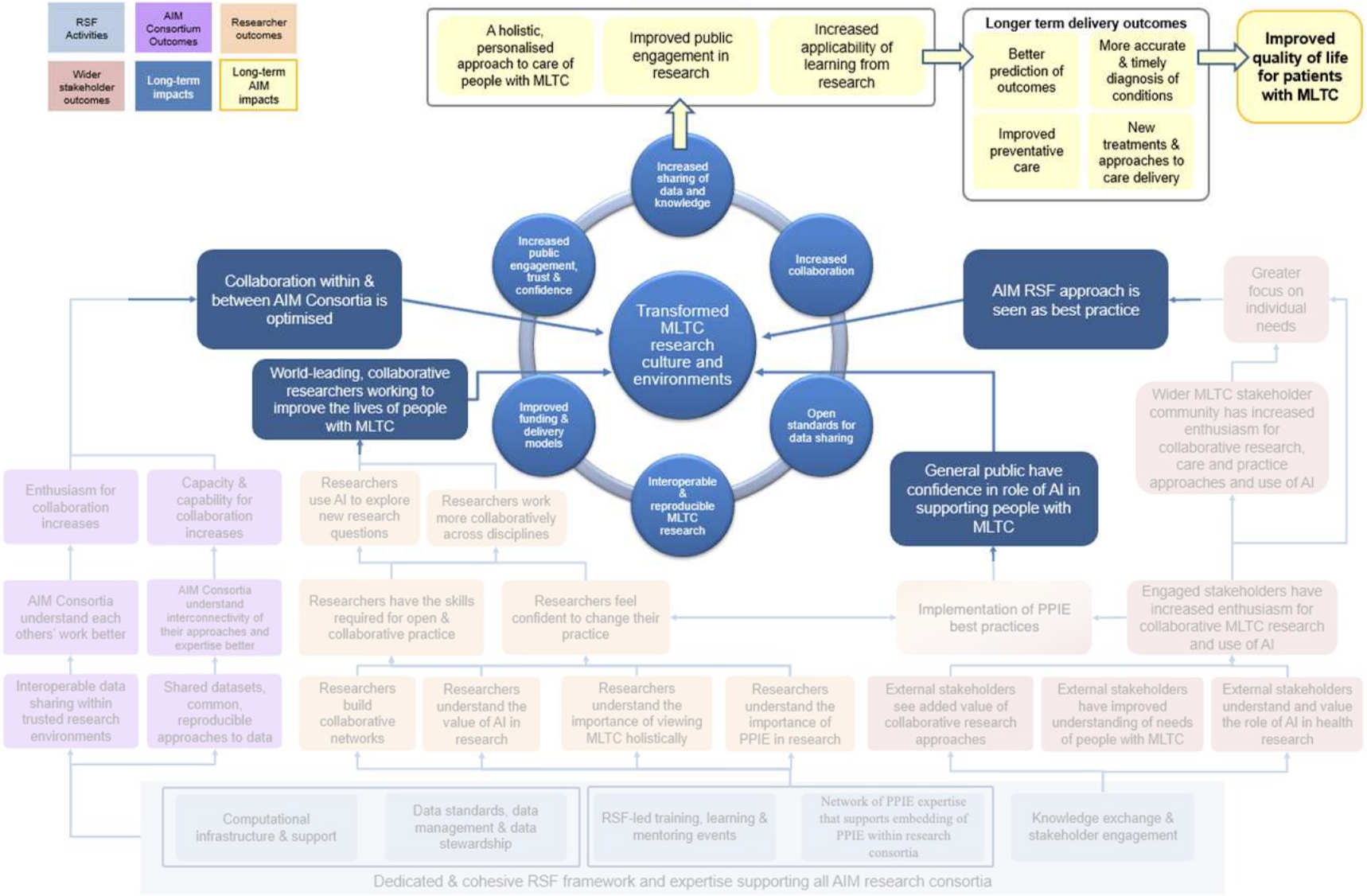
https://doi.org/10.5281/zenodo.10992603

AIM RSF Theory of Change



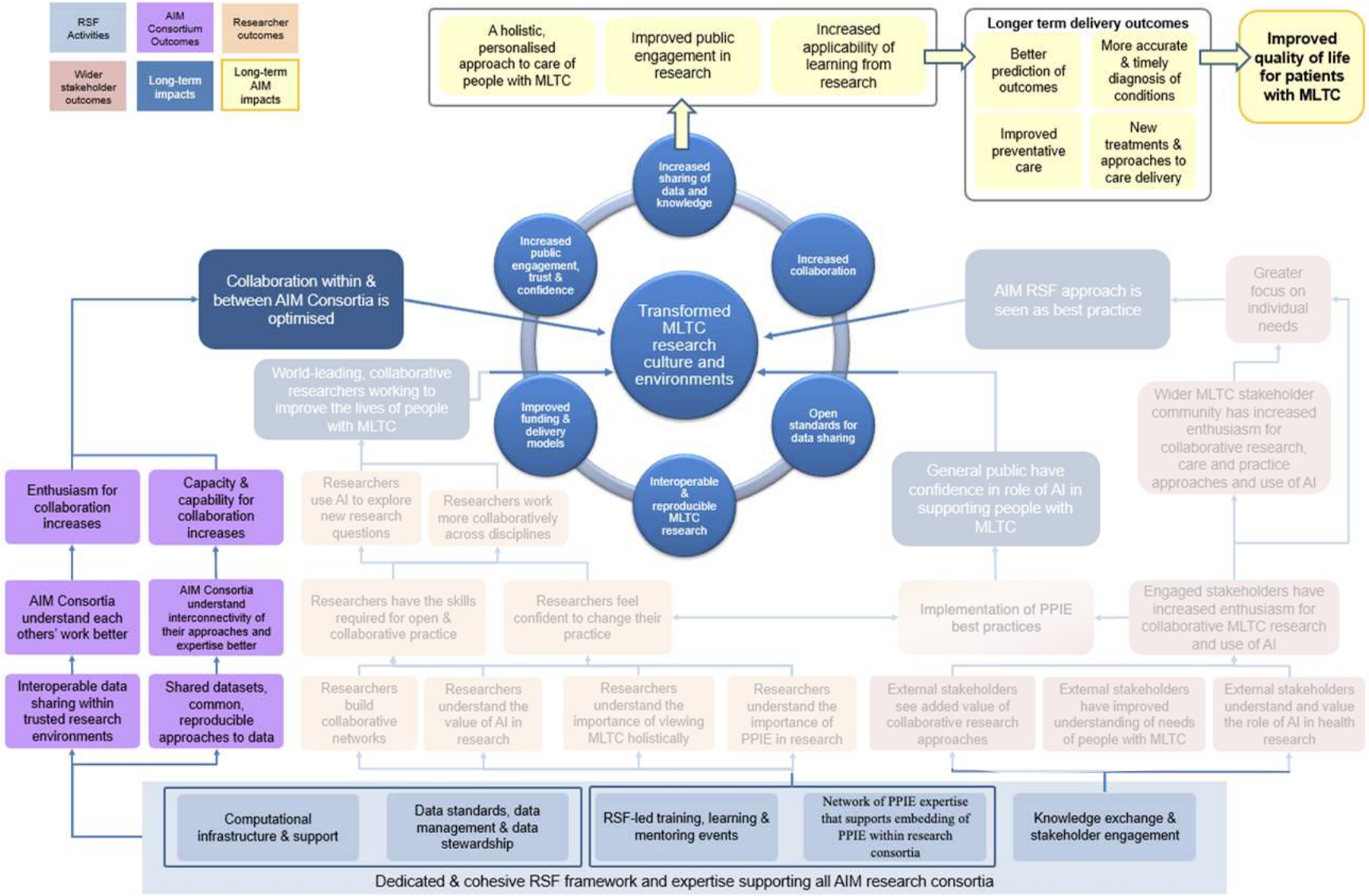
https://doi.org/10.5281/zenodo.10992603

AIM RSF Theory of Change



https://doi.org/10.5281/zenodo.10992603

AIM RSF Theory of Change



https://doi.org/10.5281/zenodo.10992603

Collaboration between and within AIM Consortia is optimised

- Driven by researcher need, specifically the DECODE and MELD-B projects, AIM RSF team have facilitated access and use of GPUs at the SAIL Databank at Swansea University.
- Currently working on a horizon scan of Trusted Research Environments appropriate for health data analysis across the UK.
- Working closely with the Clinical Practice Research Datalink team:
  - Gathered and compiled feedback on application procedures, access waiting times, data license agreements and infrastructure.
  - Accessed synthetic data to develop analytical workflows in open source programming languages curated to the CPRD datasets.

https://www.cprd.com
https://saildatabank.com

https://doi.org/10.5281/zenodo.10992603

Collaboration between and within AIM Consortia is optimised

"RSF has created a community of collaboration" – Simon Fraser, MELD-B PI

"The co-ordinating and communicating pieces wouldn't happen without dedicated resource [from the RSF]" – Thomas Jun, DECODE PI

- Regular cross-consortia meetings: PIs, project managers and ECRs
  - Appropriate level of detail and actions rather than percolating all information through senior leadership
- Hybrid conference
  - May 2023, internal to AIM Consortia and RSF in Birmingham
  - Diverse attendance: 10% Experts by lived experience, 16% clinical staff, 24% postdoctoral researchers
  - Online facilitators, meal reimbursements and accessibility fund
- Save the date for our open conference on 9 September in Manchester!

Collaboration between and within AIM Consortia is optimised

ECR Network Day in November 2023

“Brilliant opportunity to learn from and connect with your peers in the community, who you might otherwise take years to meet in person with"

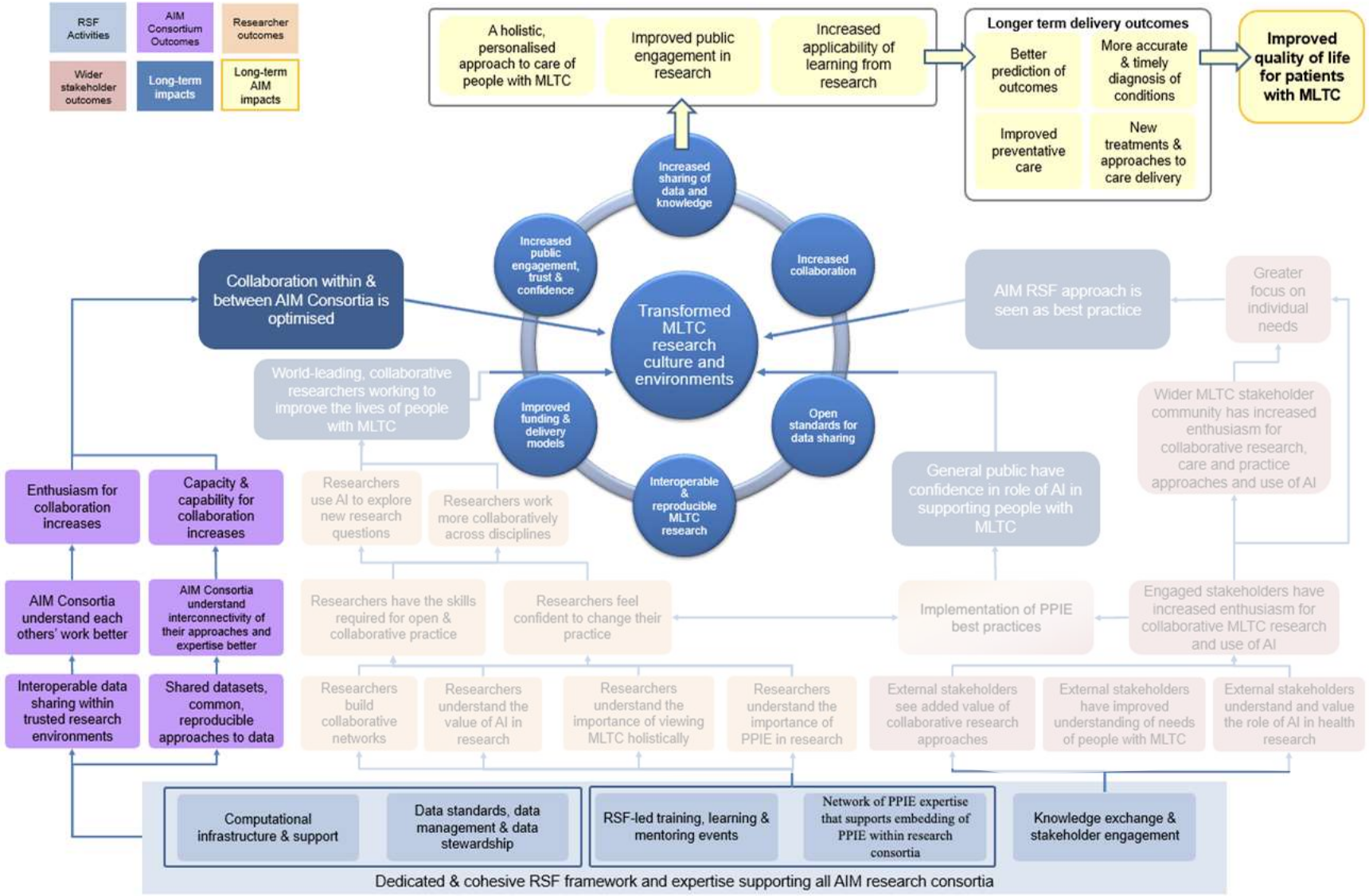
“This was a fantastic event and a great opportunity for collaboration"

“We need to do this more often!”

https://github.com/aim-rsf/AIM-ECR-network
https://github.com/aim-rsf/AIM-ECR-network/blob/main/ECR-days/London-23-Nov/planning-docs/ReproHack.md

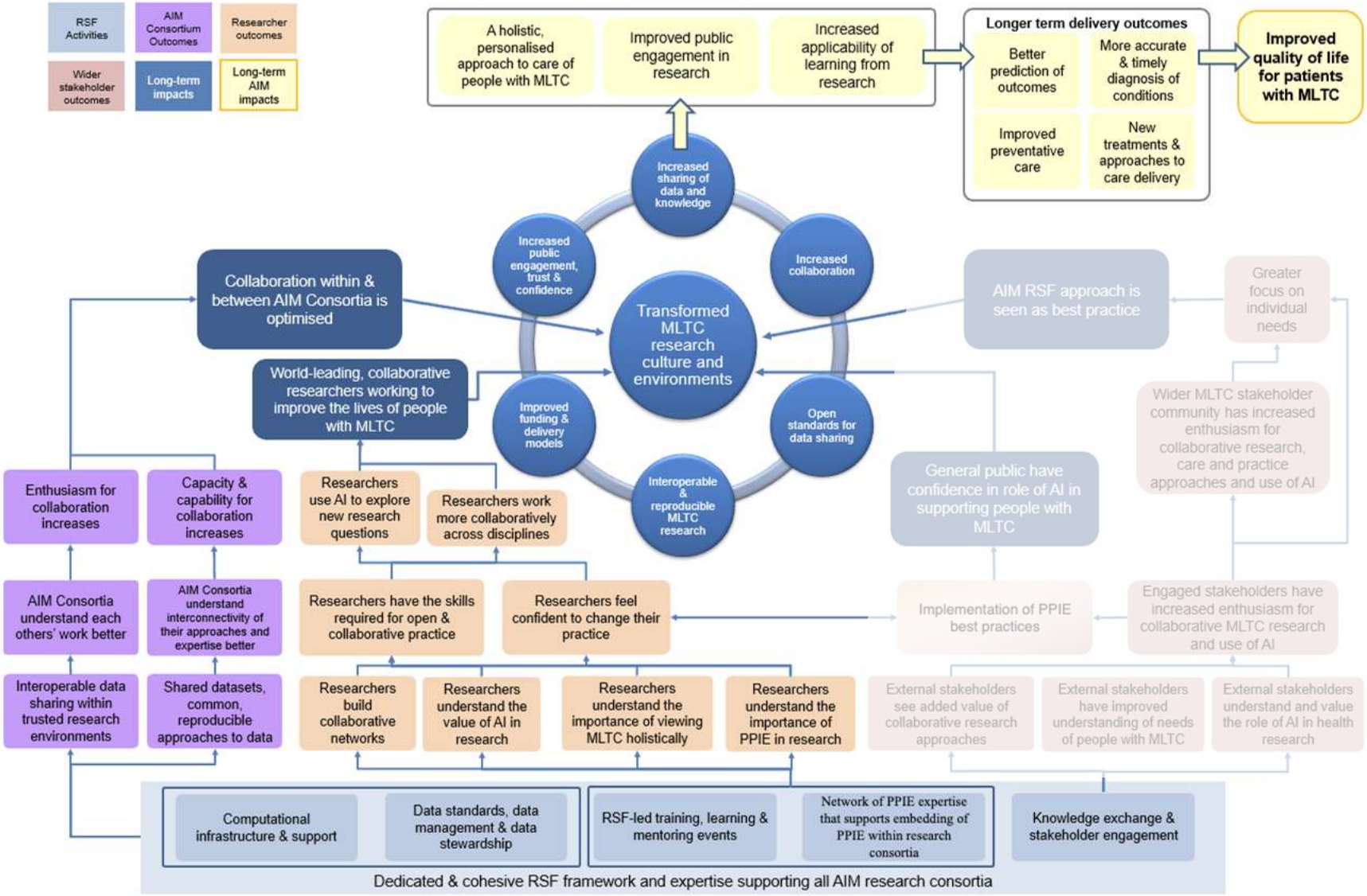
https://doi.org/10.5281/zenodo.10992603

AIM RSF Theory of Change



https://doi.org/10.5281/zenodo.10992603

AIM RSF Theory of Change



https://doi.org/10.5281/zenodo.10992603

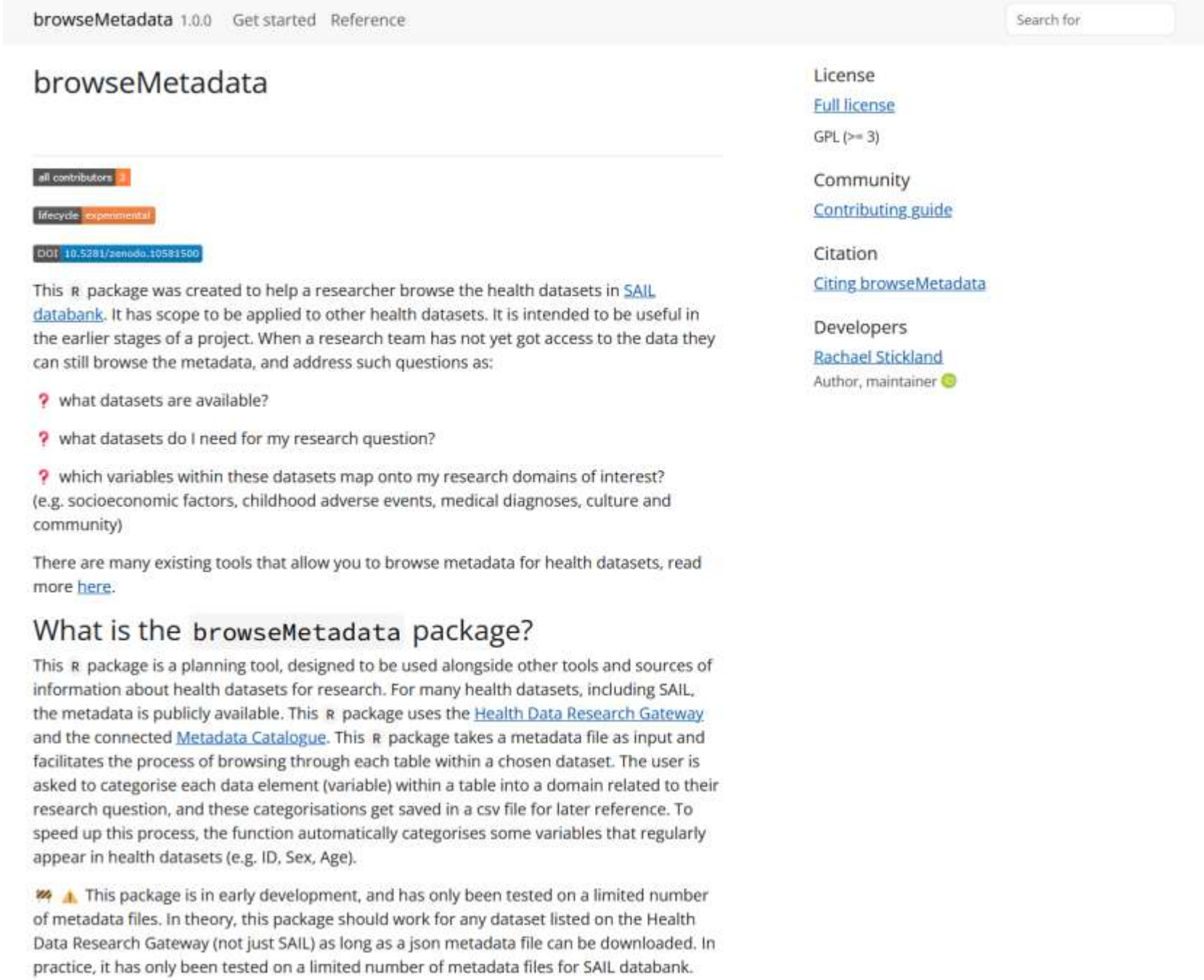
World-leading, collaborative researchers working to improve the lives of people with MLTC

- Committed to sustainability that transcends the lifetime of the AIM funding.
- Encouraging and demonstrating a shared and reproducible approach to the use and development of resources, tools and methods.
- Open source trainings available for anyone to use and reuse on:
  - Version control with git and GitHub
  - Introduction to GitHub for Collaboration
  - Publishing FAIR research outputs
  - Interactive dashboards with R Shiny
  - Literate programming with R Markdown
  - Data wrangling with dplyr and tidyr
  - Data visualisation with ggplot

https://github.com/aim-rsf/training
https://doi.org/10.5281/zenodo.7739071

https://doi.org/10.5281/zenodo.10992603

- A tool to map research domains of interest to corresponding variables within the SAIL databank.
- Inspired by the work carried out by Data Wranglers with MELD-B.
- Open source R package.
- Designed to be compatible with any dataset listed on the Health Data Research Gateway, provided a JSON metadata file is available.



https://github.com/aim-rsf/browseMetadata
https://aim-rsf.github.io/browseMetadata

https://doi.org/10.5281/zenodo.10992603

README CC-BY-4.0 license

AI-MLTC Drug Lists

all contributors 3

DOI 10.5281/zenodo.10816326

License CC BY 4.0

This repository has been setup to share drug lists across NIHR's AIM projects. Currently, the drug lists come from the AIM Consortium OPTIMAL, based in Birmingham. You can find all the drug lists created by OPTIMAL in this repository.

About the lists

There is variation in how different systems approach the classification and mapping of various drugs, and it has been noted that there is a lack of single data consistency.

The drug lists have been created using different methods to find a more efficient process to integrate as part of the Code Builder tool:

- **Method 1**: Involved manually curating a list of generic drug names on the BNF, and searching for respective brand names on the dm+d browser
- **Method 2**: Involves the use of both the `snomedizer` `R` package and a dm+d tool implementation (downloading latest dm+d distribution from NHS Technology Reference Update Distribution TRUD webpage). It's in the pipeline to create a self-contained wrapper for this process.
  - this method is a substance-based approach and involves running through a list of IDs for medicinal products.

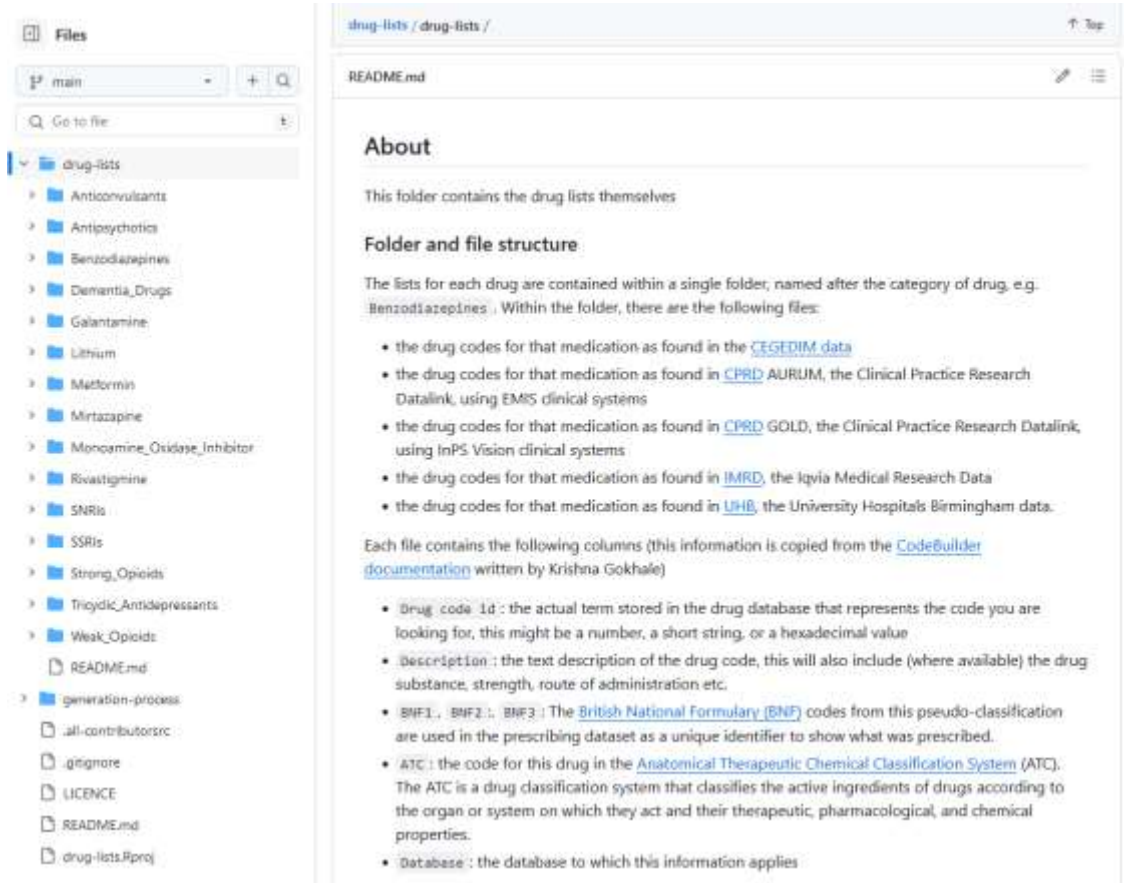
The process for the creation of the lists can be found in the generation-process folder and the drug lists themselves in the drug-lists folder.

Contributors ✨

Thanks goes to these wonderful people (emoji key):

This project follows the all-contributors specification. Specifically, `content` refers to creation of drug lists, `documentation` refers to the documentation created in this repository, and `Feedback` refers to clinical auditing of the lists.

- In collaboration with the OPTIMAL team
- To date, 15 code lists generated and filtered by the RSF data wranglers, and quality checked by OPTIMAL's clinical team.
- Openly available under CC-BY license



https://github.com/aim-rsf/drug-lists

https://doi.org/10.5281/zenodo.10992603

World-leading, collaborative researchers working to improve the lives of people with MLTC

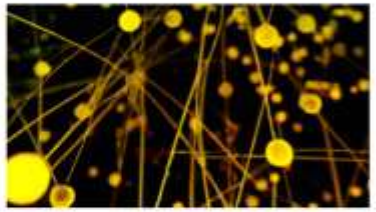
The Good, the Bad and the Public
Tuesday 13 Aug 2024
Time: 13:30 - 14:30
Anica Alvarez Nishio

Data-Centric AI
Tuesday 14 May 2024
Time: 13:30 - 14:30
Prof. Paolo Missier

AI-based modelling of EHRs of patients with learning disabilities and multiple long-term health conditions
Tuesday 09 Apr 2024
Time: 13:30 - 14:30
Professor Georgina Cosma | Dr. Rania Kousovista | Dr. Emeka Abakasanga

Multimorbidity and place
Thursday 14 Mar 2024
Time: 11:00 - 12:00
Alan Marshall

Dynamic Artificial Intelligence (DynAIRx)
Tuesday 13 Feb 2024
Time: 13:30 - 14:30
Dr. Aseel Abuzour



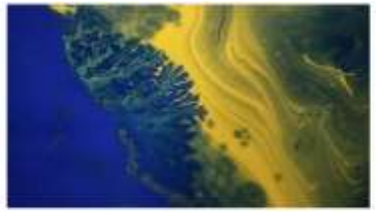
Translating 'Practice into Evidence'
Tuesday 09 Jan 2024
Time: 13:30 - 14:30
Deborah Swinglehurst

Drawing straight lines along blurred boundaries
Tuesday 14 Nov 2023
Time: 13:30 - 14:30
Louise Locock

Fairness: The cascading effects in medical devices
Tuesday 10 Oct 2023
Time: 13:30 - 14:30
Claudia Fischer | Smera Jayadeva

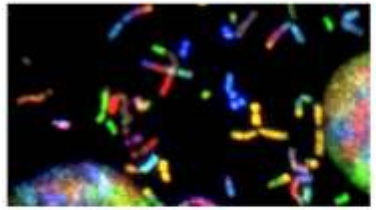
The SAIL Databank research project journey
Tuesday 26 Sep 2023
Time: 13:30 - 14:30
Ian Farr | Sarah Rees

How DataLoch Can Be Utilised for Researchers Studying Multi-Long-Term Conditions
Tuesday 08 Aug 2023
Time: 13:30 - 14:30
Kathy Harrison

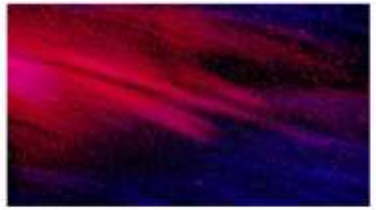


How to Deal with Privacy, Bias & Drift in Synthetic Primary Care Data
Tuesday 11 Jul 2023
Time: 13:30 - 14:30
Allan Tucker

The potential for AI to help with better prescribing and medicines optimisation
Tuesday 13 Jun 2023
Time: 13:30 - 14:30
Dr. Anthony J Avery OBE



Public engagement: Data for science and health
Tuesday 09 May 2023
Time: 13:30 - 14:30
Emily Jesper-Mir



Structured missing data: Grand challenges in learning from multi-modal data at scale
Tuesday 14 Mar 2023
Time: 13:30 - 14:30
Robin Mitra | Sarah McGough

RSF seminar series: Using routinely collected electronic healthcare record data to study respiratory disease
Tuesday 14 Feb 2023
Time: 13:30 - 14:30
Jennifer Quint

RSF seminar series: Developing and publishing code for trusted research environments
Tuesday 08 Nov 2022
Time: 13:30 - 14:30
Ed Chalstrey

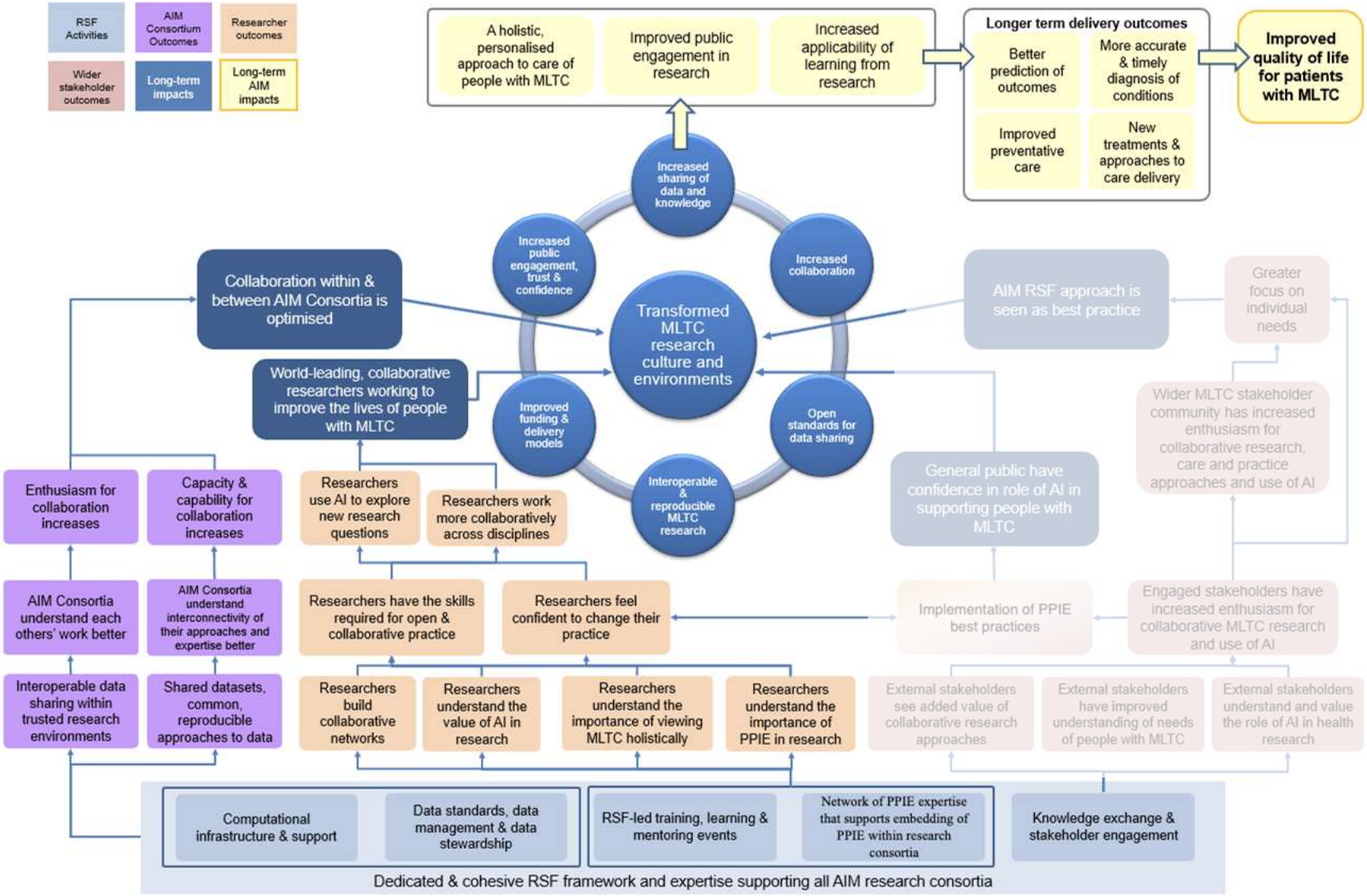
RSF seminar series: DExtER, a semi-automated epidemiology platform for electronic health record research
Tuesday 11 Oct 2022
Time: 13:30 - 14:30

RSF seminar series: Everyone has a say
Tuesday 13 Sep 2022
Time: 13:30 - 14:30
Meag Doherty

https://www.turing.ac.uk/events/rsf-monthly-seminars
https://zenodo.org/communities/ai-mltc-m/

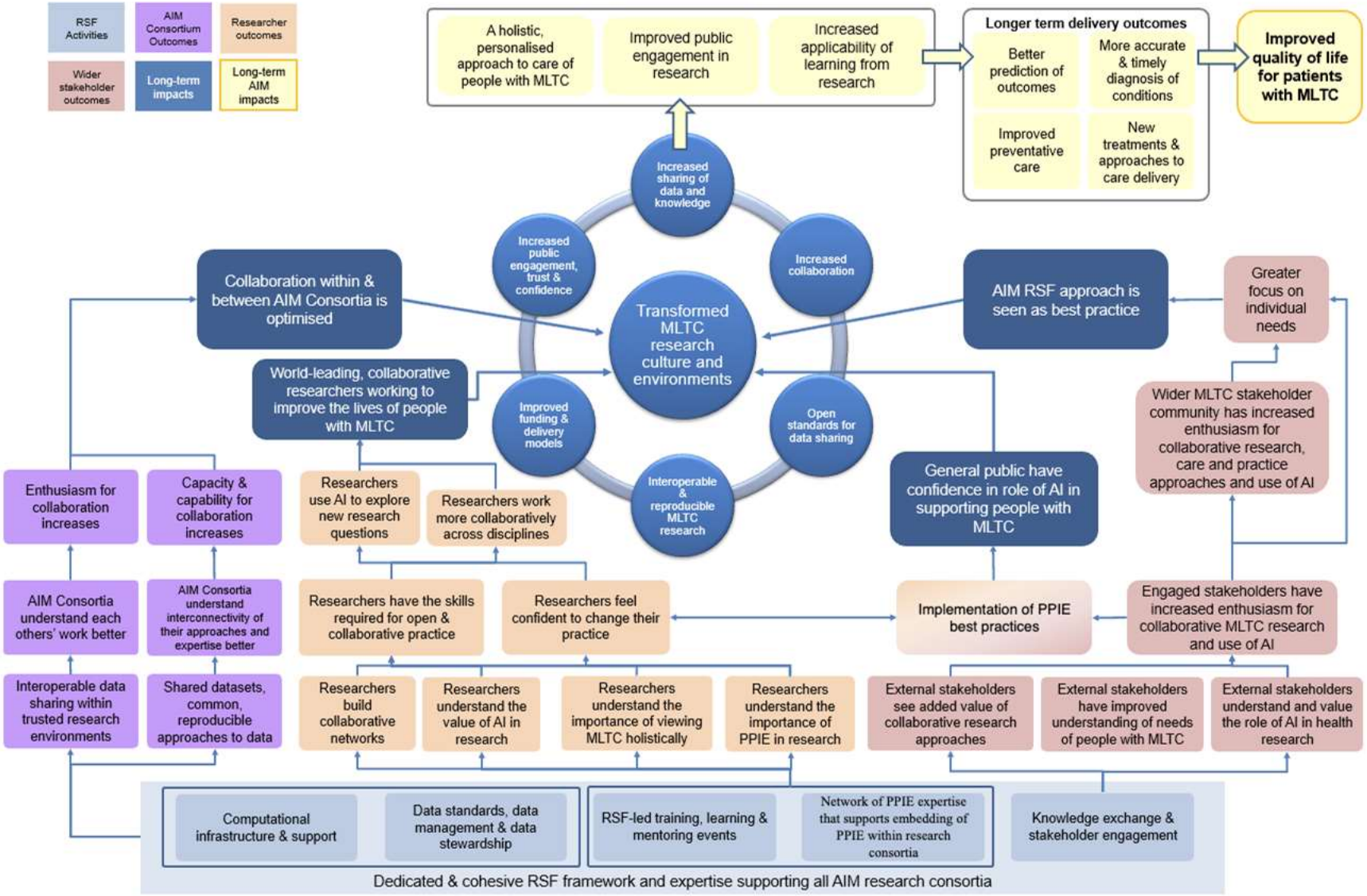
https://doi.org/10.5281/zenodo.10992603

AIM RSF Theory of Change



https://doi.org/10.5281/zenodo.10992603

AIM RSF Theory of Change



https://doi.org/10.5281/zenodo.10992603

General public have confidence in role of AI in supporting people with MLTC

- Co-hosting and facilitating meetings of the MLTC Community of Practice led by Stephanie Hanley
- Workshops on “What is Data?” and “What is AI?” alongside broader discussions, including recent input to the design of our annual conference.

“I have thoroughly enjoyed being part [of CoP], I am patient voice rather than early career researcher. So, I see my role as more to reassure early career researchers we are not scary and can help them do things better” - Amanda, expert by lived experience.

“The advantage of this community is that it is a big PPI group and meetings are well attended including several members of the research team. It is now a well-developed group which bodes well for the future” - Margaret, expert by lived experience.

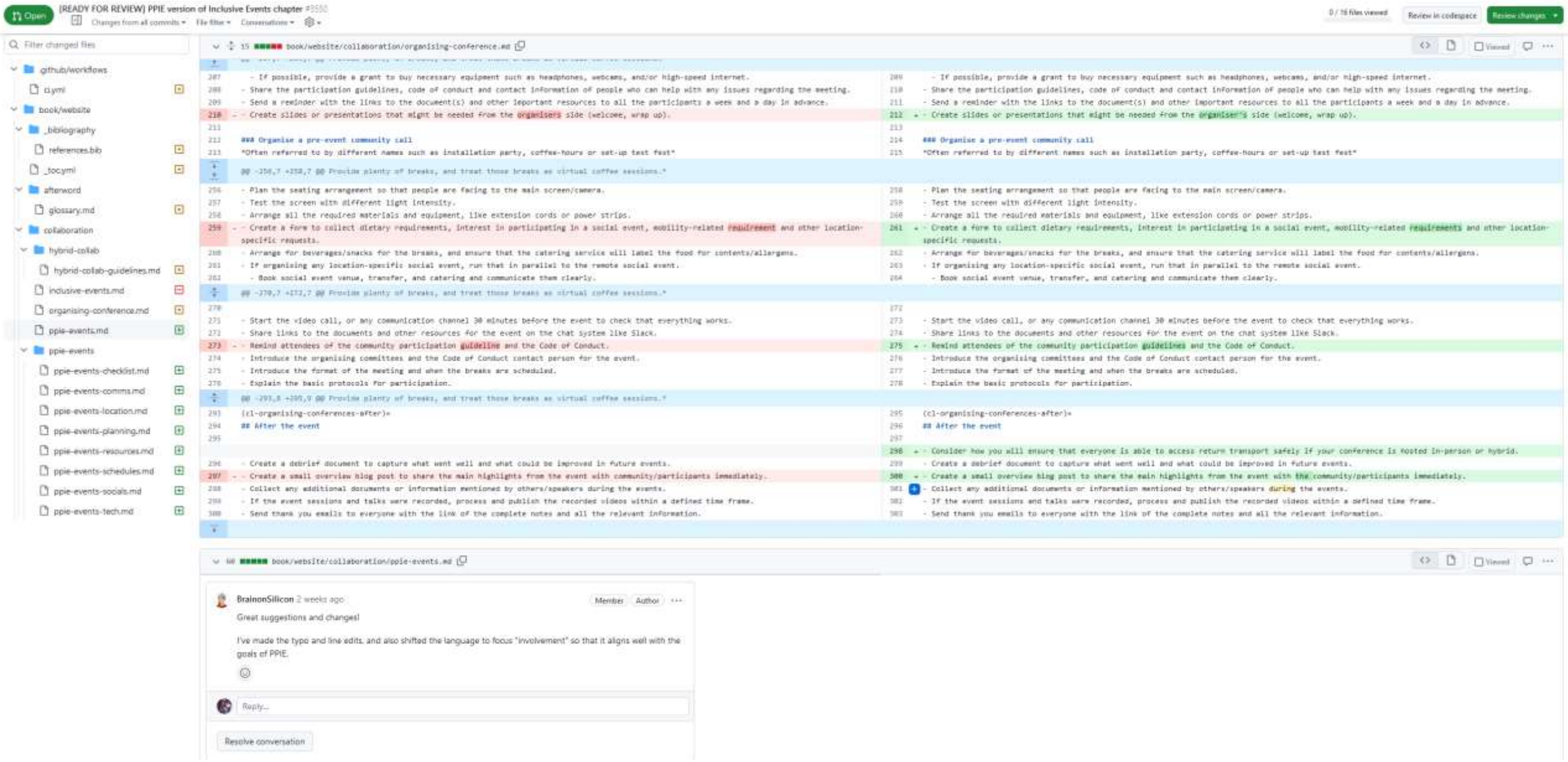
The Turing Way is an **open source book project** that involves and supports a **diverse research community** in ensuring that reproducible and ethical **data science** is accessible and comprehensible **for everyone**.

@TuringWay @kirstie_j

https://doi.org/10.5281/zenodo.10992603

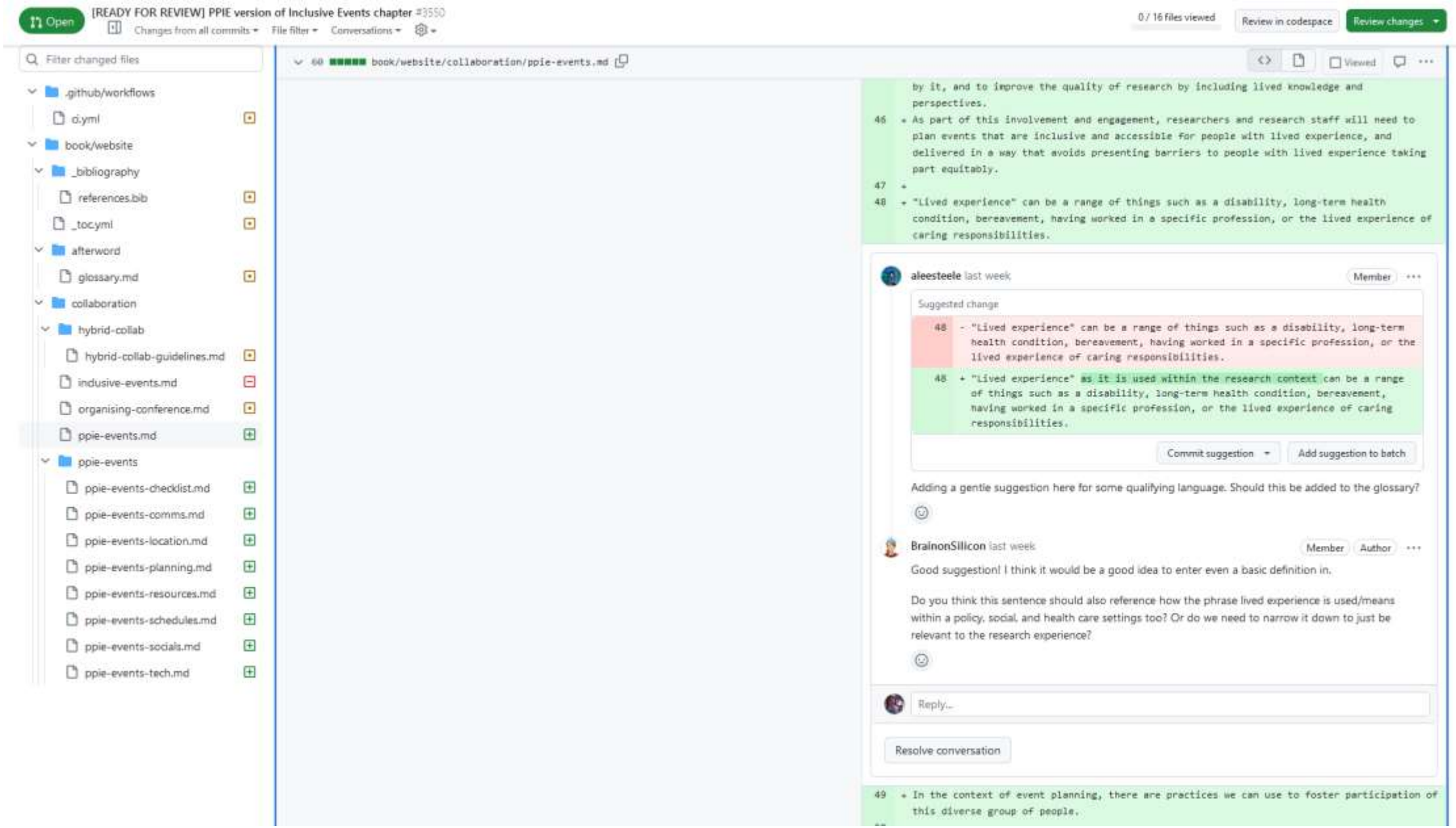
https://the-turing-way.netlify.app/index.html
https://github.com/the-turing-way/the-turing-way

https://doi.org/10.5281/zenodo.10992603



https://github.com/the-turing-way/the-turing-way/pull/3550
https://the-turing-way.netlify.app/index.html
https://github.com/the-turing-way/the-turing-way

https://doi.org/10.5281/zenodo.10992603



https://github.com/the-turing-way/the-turing-way/pull/3550
https://the-turing-way.netlify.app/index.html
https://github.com/the-turing-way/the-turing-way

https://doi.org/10.5281/zenodo.10992603

AIM RSF approach is seen as best practise

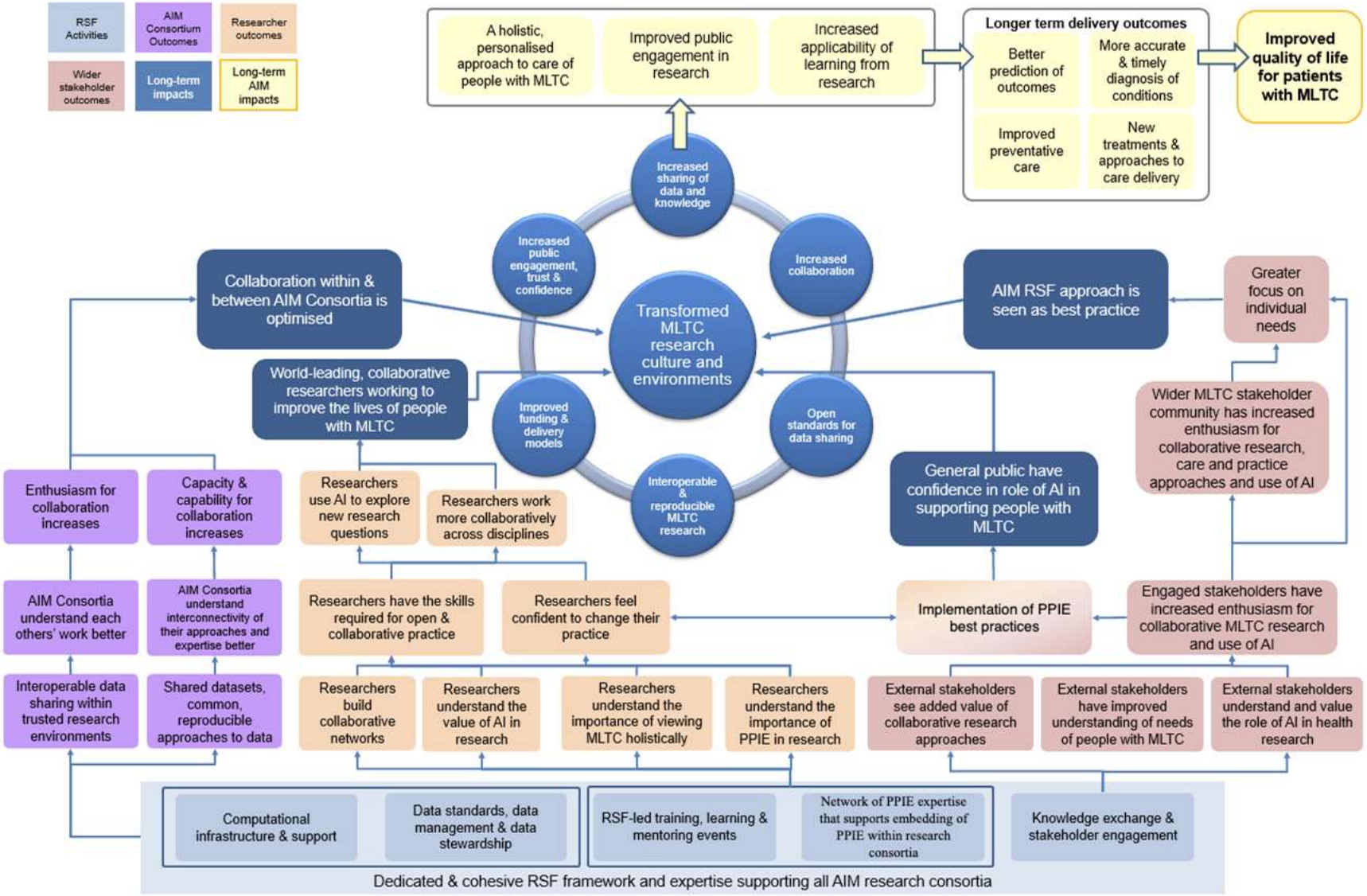
“As a PI, if the RSF isn’t there, it becomes my job to do all this and this becomes an extra burden of work for me. This work may or may not happen, and could delay best practice in this area.” Rafael Perera-Salazar, COMPUTE PI

“Capturing what all the consortia have done in a rigorous way including the problems they’ve faced is something the consortia would never do themselves” Bruce Guthrie, AIM CISC PI.

- Engage with stakeholders across the ecosystem.
  - Recent presentations at Turing’s AIUK conference.
- Demonstrate a commitment to openness and trustworthy research and innovation.
- Regular connections with NIHR Programme management.

Email: aimrsf@turing.ac.uk, kwhitaker@turing.ac.uk
https://www.turing.ac.uk/research/research-projects/ai-multiple-long-term-conditions-research-support-facility https://doi.org/10.5281/zenodo.10992603

AIM RSF Theory of Change



https://doi.org/10.5281/zenodo.10992603

AI for Multiple Long Term Conditions

- AIM-CISC: Analyse and understand which patterns of multimorbidity are most common, which most affect people's lives, and help improve the quality and safety of care.
  - University of Edinburgh, The Roslin Institute, NHS Lothian, and University College London.
- AI-MULTIPLY: Characterise the dynamic relationships of MLTC and polypharmacy and inform healthcare pathways.
  - University of Newcastle and Queen Mary University.
- Cluster-AIM: Develop and validate population clusters to integrate health and social care using mixed methods.
  - Southampton, Oxford, Kent, Nottingham and Leicester.
- DECODE: Mapping the challenges and requirements for Data-driven, machinE learning-aided stratification and management of multiple long-term COnditions in adults with intellectual DisabilitiEs (ID).
  - Leicestershire NHS Trust, Loughborough University, University of Leicester, and De Montfort University.
- DynAIRx: Develop new, easy-to-use AI and data visualisation tools that help GPs and pharmacists treat patients with MLTC.
  - Liverpool, Leeds, Manchester and Glasgow.
- MELD-B: Identify life-course time points and targets for the prevention of early-onset, burdensome Multiple Long-Term Conditions.
  - University of Southampton, University of Glasgow, Swansea University, Southampton City Council, University of Aberdeen, and King's College London.
- OPTIMAL: OPTIMising therapies, discovering therapeutic targets and AI-assisted clinical management for patients living with complex multiple long-term conditions.
  - University of Birmingham, University of Manchester, University Hospitals Birmingham NHS Trust, NHS Greater Glasgow & Clyde, and University of St Andrews.

AI for Multiple Long Term Conditions: Research Support Facility (AIM RSF), AIM-CISC, AI-MULTIPLY, Cluster-AIM, DECODE, DynAIRx, MELD-B, & OPTIMAL. (2024). AI for Multiple Long-Term Conditions Handbook. Zenodo. https://doi.org/10.5281/zenodo.10867842

https://doi.org/10.5281/zenodo.10992603

Swansea University
Prifysgol Abertawe

The Alan Turing Institute

THE UNIVERSITY of EDINBURGH

Theme 1: Reproducible and interoperable infrastructure

Theme 2: Research-ready data







Theme 3: Community building and training

Theme 4: Public and patient involvement and engagement

Programme management







Theme 5: Sustainability and legacy



The AIM RSF Team

- David Ford
- Ronan Lyons
- Chris Orton
- Reece Labrom
- Ann-Marie Mallon
- Rachael Stickland
- Mahwish Mohammad
- Evelina Gabasova
- Emma Karoune
- Bastian Greshake Tzovaras
- Sophia Batchelor
- Aziz Sheikh
- Monica Fletcher
- Beth Collop
- Kirstie Whitaker
- Sydney Ambrose
- Batool Almarzouq

https://doi.org/10.5281/zenodo.10992603