

# Exploring the opportunities, responsibilities, and challenges of AI in research: A service provider's perspective

**Simon Porter**

AI UK Fringe 2024 Workshop

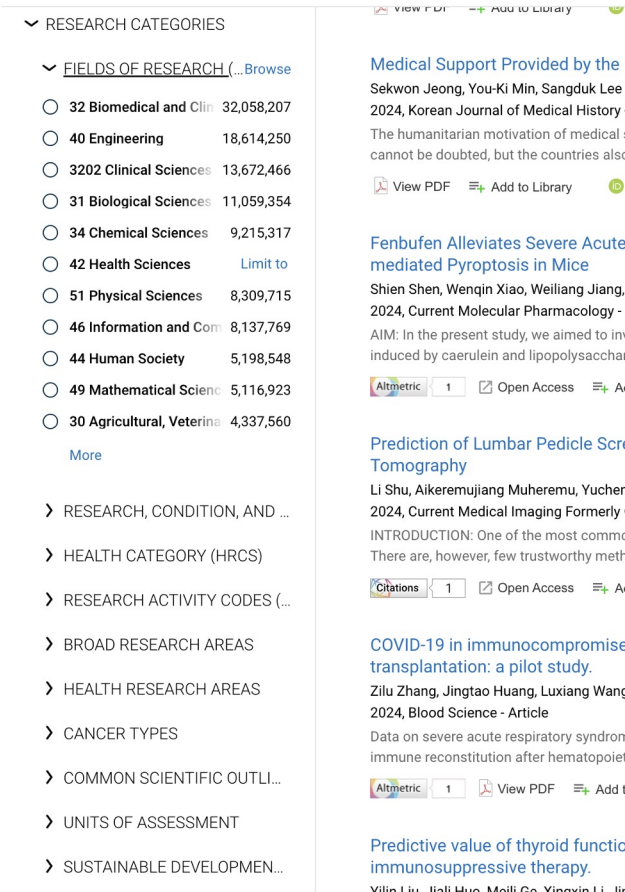


# Our AI Story



# Before the current AI trend was Machine Learning...

- Dimensions
  - **Classification schemes: RCDC, SDGs, FoR Codes**
  - Clustering analyses
  - Data annotation and enhancement
- Writefull
- LLMs
  - Dimensions
  - ReadCube
  - OntoChem & Metaphacts



RESEARCH CATEGORIES

FIELDS OF RESEARCH (...Browse)

32 Biomedical and Clin	32,058,207
40 Engineering	18,614,250
3202 Clinical Sciences	13,672,466
31 Biological Sciences	11,059,354
34 Chemical Sciences	9,215,317
42 Health Sciences	Limit to
51 Physical Sciences	8,309,715
46 Information and Com	8,137,769
44 Human Society	5,198,548
49 Mathematical Scienc	5,116,923
30 Agricultural, Veterina	4,337,560

More

RESEARCH, CONDITION, AND ...

HEALTH CATEGORY (HRCS)

RESEARCH ACTIVITY CODES (...)

BROAD RESEARCH AREAS

HEALTH RESEARCH AREAS

CANCER TYPES

COMMON SCIENTIFIC OUTLI...

UNITS OF ASSESSMENT

SUSTAINABLE DEVELOPMEN...

Medical Support Provided by the  
Sekwon Jeong, You-Ki Min, Sangduk Lee  
2024, Korean Journal of Medical History  
The humanitarian motivation of medical  
cannot be doubted, but the countries also

View PDF Add to Library

Febufen Alleviates Severe Acute  
mediated Pyroptosis in Mice  
Shien Shen, Wenqin Xiao, Weiliang Jiang,  
2024, Current Molecular Pharmacology -  
AIM: In the present study, we aimed to in  
induced by caerulein and lipopolysacchar

Altmetric 1 Open Access

Prediction of Lumbar Pedicle Scre  
Tomography  
Li Shu, Aikeremujiang Muheremu, Yucher  
2024, Current Medical Imaging Formerly  
INTRODUCTION: One of the most commo  
There are, however, few trustworthy meth

Citations 1 Open Access

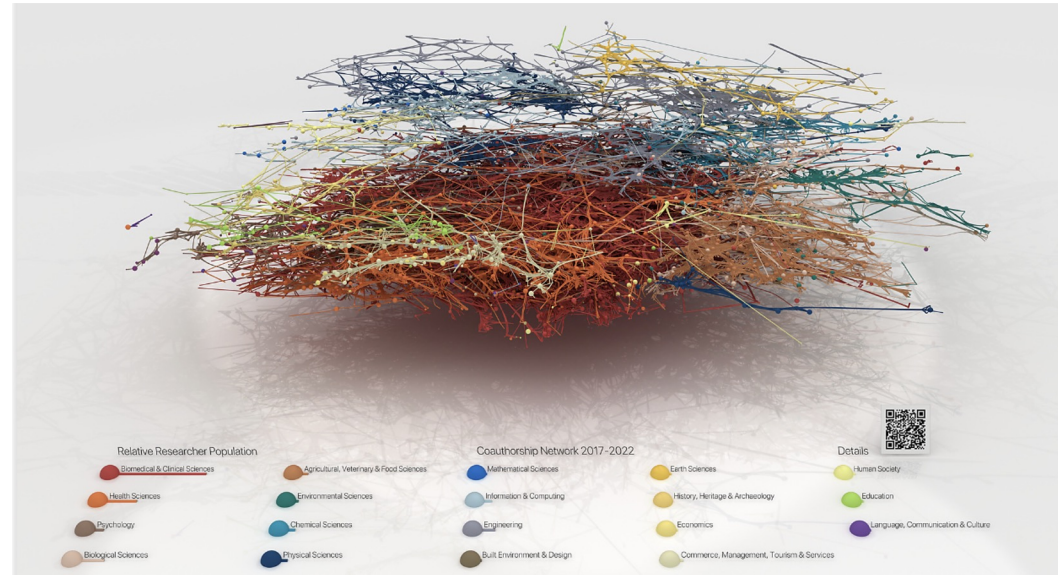
COVID-19 in immunocompromise  
transplantation: a pilot study.  
Zilu Zhang, Jingtao Huang, Luxiang Wang  
2024, Blood Science - Article  
Data on severe acute respiratory syndrom  
immune reconstitution after hematopoiet

Altmetric 1 View PDF Add to Library

Predictive value of thyroid functio  
immunosuppressive therapy.  
Yilin Liu, Jiali Huo, Meili Gu, Xinyin Li,

# Before the current AI trend was Machine Learning...

- Dimensions
  - Classification schemes: RCDC, SDGs, FoR Codes
  - **Clustering analyses**
  - Data annotation and enhancement
- Writefull
- LLMs
  - Dimensions
  - ReadCube
  - OntoChem & Metaphacts



# University of Tokyo



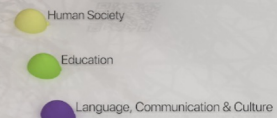
## Relative Researcher Population



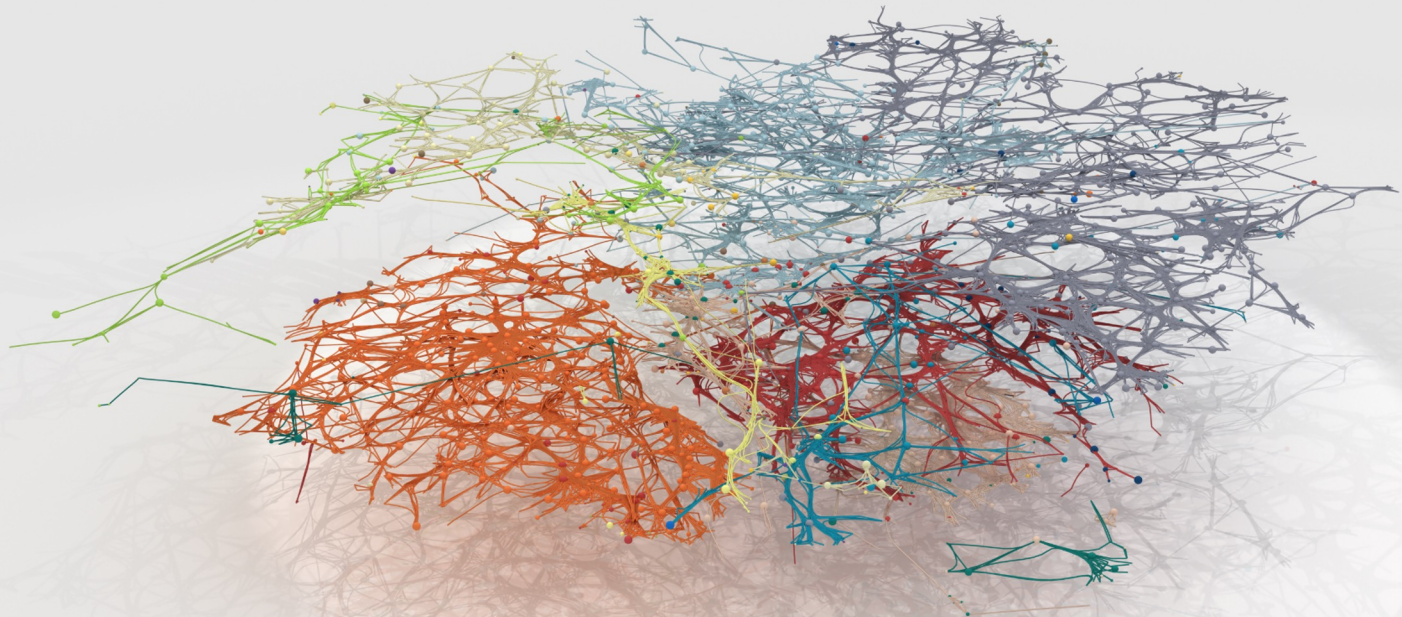
## Coauthorship Network 2017-2022



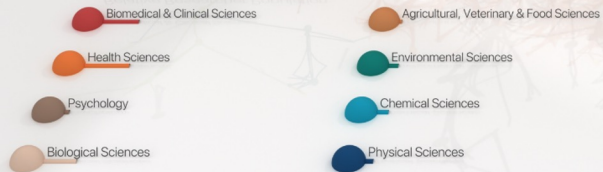
## Details



# University of Technology Sydney



## Relative Researcher Population



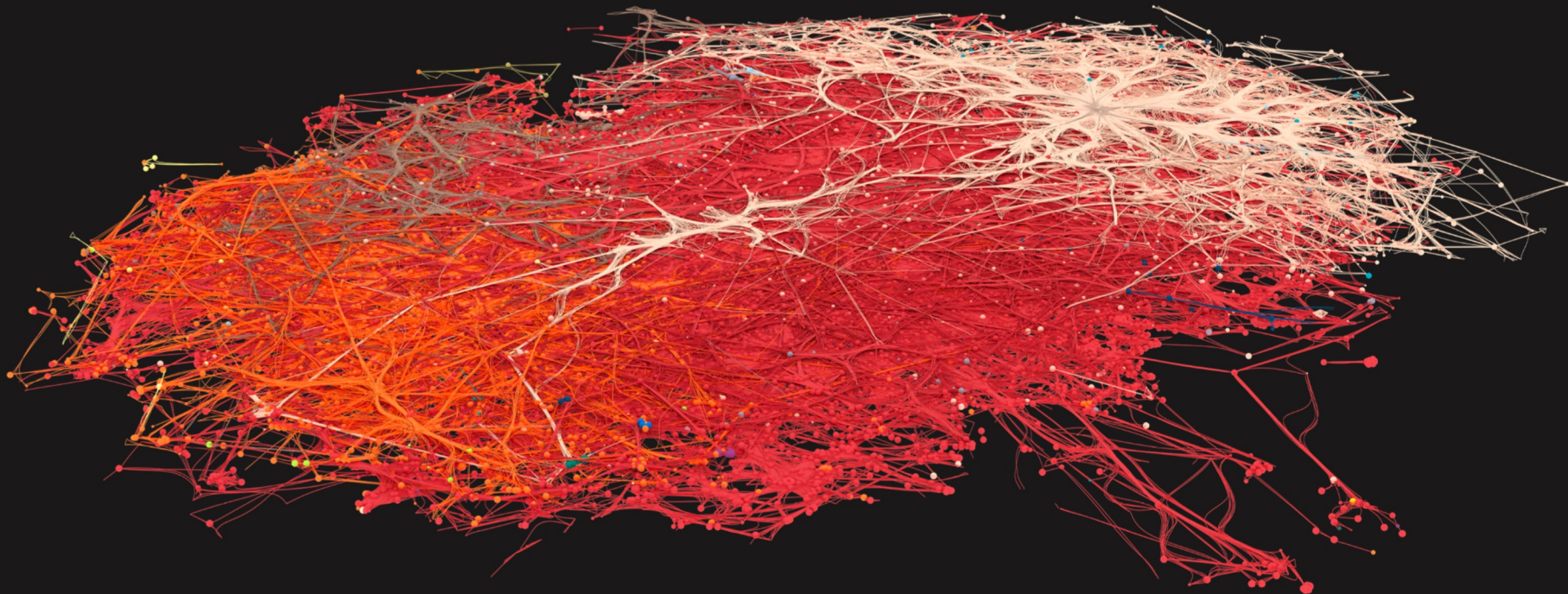
## Coauthorship Network 2017-2022



## Details



# University of California, San Francisco



## About the Researchers

The network above represents a connected graph of 26,279 co-authored researchers affiliated to the University of California, San Francisco from 2017-2022, making up 94% of all affiliated researchers over this time period. Each researcher has been colour coded by the 2-digit FoR 2020 code they are most associated with.

## About the Clusters

387 research clusters were identified in the network above. To make the network easier to read, collaborations between clusters are not displayed, although they do play a significant role in the layout of the network. Clusters of 20 or more researchers can be explored further in the associated [Fishbone record](#) (linked in the QR code top right).

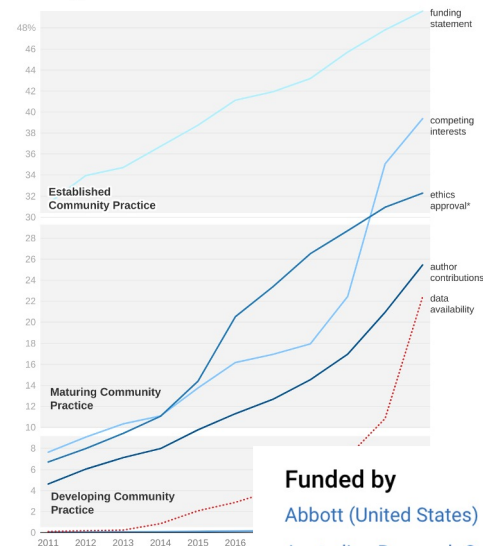
## About the Classifications

The 2020 Field of Research codes used in this analysis have been assigned to publications using the approach detailed in "Recategorising research: Mapping from FoR 2008 to FoR 2020 in Dimensions" ([https://doi.org/10.1162/qss\\_a\\_00244](https://doi.org/10.1162/qss_a_00244).) Note: some research areas are not well represented in the network due to single author

# Before the current AI trend was Machine Learning...

- Dimensions
  - Classification schemes: RCDC, SDGs, FoR Codes
  - Clustering analyses
  - **Data annotation and enhancement**
- Writefull
- LLMs
  - Dimensions
  - ReadCube
  - OntoChem & Metaphacts

Evolving Science Trust Markers 2011-2021



## Funded by

Abbott (United States)

Australian Research Council

Department of Health and Social Care

Government of Canada

NIHR Academy

National Eye Institute

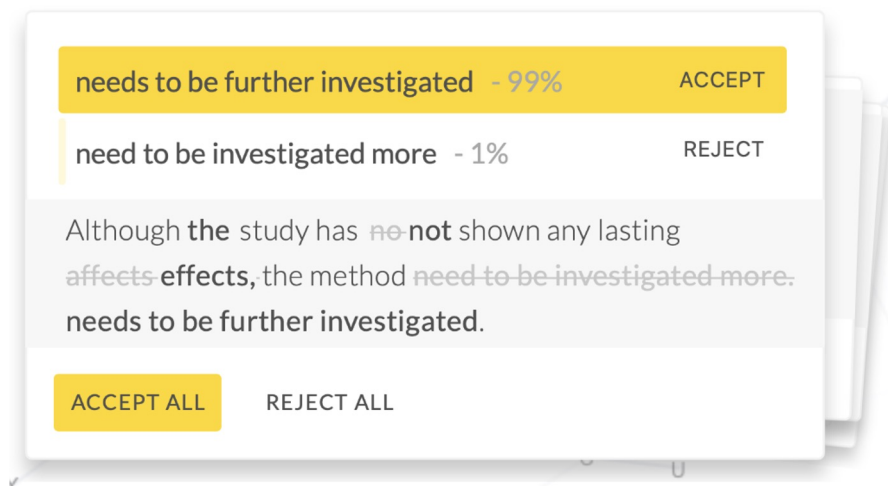
National Health and Medical Research Council

\*The percentage of ethics papers are calculated over pu the ethics trust marker looks at those papers that include ethics approval 'somewhere in the text'  
Chart: Dimensions Research Integrity • Source: Dimensions



# Before the current AI trend was Machine Learning...

- Dimensions
  - Classification schemes: RCDC, SDGs, FoR Codes
  - Clustering analyses
  - Data annotation and enhancement
- **Writefull**
- LLMs
  - Writeful
  - Dimensions
  - ReadCube
  - OntoChem & Metaphacts

A screenshot of a text analysis interface. It shows a yellow bar at the top with the text "needs to be further investigated - 99%" and "ACCEPT" on the right. Below it is a white bar with "need to be investigated more - 1%" and "REJECT" on the right. The main text area contains the sentence: "Although the study has not shown any lasting affects effects, the method need to be investigated more. needs to be further investigated." At the bottom, there are two buttons: "ACCEPT ALL" (yellow) and "REJECT ALL" (white).

# Responsibilities



# With experience comes perspective...

- Responsible AI
  - **The challenges of introducing technology in our space and the adverse effects**
  - Speed to market versus making the right decisions
  - Principles
  - Transparency

Effect of climate on tuberculosis

### Summary

Climate and air quality have a potential to regulate the incidence of tuberculosis, with precipitation, atmospheric pressure, relative humidity, temperature, and wind speed having varying effects on the incidence of tuberculosis [1]. The study found evidence of delayed effects of climate on PTB incidence in Brunei, with varying degrees of magnitude, direction and timing [2]. The prevalence of tuberculosis was influenced by average temperature, rainfall, wind speed, and air pressure, but not by average humidity [3]. Higher temperature extremes increase the risk of tuberculosis outbreaks within groups by increasing physiological stress and dispersal of males, which can lead to rapid intensification of climate-change impacts in natural populations [4].

### Top publications related to your interest

[Climate and air pollution alter incidence of tuberculosis in Beijing, China](#) [1]

Chun Yan Zhang, Ang Zhang  
2019, *Annals of Epidemiology* - Article

Citations 30 Add to Library 10.1016/j.annepidem.2019.07.003

**Excerpt** Summary & highlights Abstract

... on tuberculosis incidence by improving the concentrations of inhalable particulate matter and sulfur dioxide are potential regulators of the incidence of tuberculosis. The improved air quality contributes to the health of the population. **The impact of climatic indicators on the incidence of tuberculosis was mainly regulated by the environment.** To formulate preventive and regulatory strategies for tuberculosis based on different climatic and air quality conditions, tuberculosis has long threatened human health and is a major public health issue.

# With experience comes perspective...

- Responsible AI
  - The challenges of introducing technology in our space and the adverse effects
  - **Speed to market versus making the right decisions**
  - Principles
  - Transparency

Effect of climate on tuberculosis

### Summary

Climate and air quality have a potential to regulate the incidence of tuberculosis, with precipitation, atmospheric pressure, relative humidity, temperature, and wind speed having varying effects on the incidence of tuberculosis [1]. The study found evidence of delayed effects of climate on PTB incidence in Brunei, with varying degrees of magnitude, direction and timing [2]. The prevalence of tuberculosis was influenced by average temperature, rainfall, wind speed, and air pressure, but not by average humidity [3]. Higher temperature extremes increase the risk of tuberculosis outbreaks within groups by increasing physiological stress and dispersal of males, which can lead to rapid intensification of climate-change impacts in natural populations [4].

### Top publications related to your interest

[Climate and air pollution alter incidence of tuberculosis in Beijing, China](#) [1]

Chun Yan Zhang, Ang Zhang  
2019, *Annals of Epidemiology* - Article

Citations 30 Add to Library 10.1016/j.annepidem.2019.07.003

**Excerpt** Summary & highlights Abstract

... on tuberculosis incidence by improving the concentrations of inhalable particulate matter and sulfur dioxide are potential regulators of the incidence of tuberculosis. The improved air quality contributes to the health of the population. **The impact of climatic indicators on the incidence of tuberculosis was mainly regulated by the environment.** To formulate preventive and regulatory strategies for tuberculosis based on different climatic and air quality conditions, tuberculosis has long threatened human health and is a major public health issue.

# With experience comes perspective...

- Responsible AI
  - The challenges of introducing technology in our space and the adverse effects
  - Speed to market versus making the right decisions
  - **Principles**
  - Transparency

- **Responsibility & Respect:** We aim to use AI systems in a socially beneficial and inclusive manner, consider the impact on people and take steps to prevent harm to society and the environment.
- **Fairness:** We take action to mitigate the potential for bias and inequities.
- **Transparency:** We disclose when an AI system is being used and explain how our solutions work in an understandable manner.
- **Accountability:** We create accountability through human oversight.
- **Privacy and Data Security:** We respect personal privacy and data protection laws.
- **Copyright:** We respect existing copyrights and intellectual property rights.

# With experience comes perspective...



- Responsible AI
  - The challenges of introducing technology in our space and the adverse effects
  - Speed to market versus making the right decisions
  - Principles
  - Transparency

**QSS**  
an open access journal

Check for updates

Citation: Porter, S. J., Hawley, L., & Hook, D. W. (2023). Recategorising research: Mapping from FoR 2008 to FoR 2020 in Dimensions. *Quantitative Science Studies*, 4(1), 127–142. [https://doi.org/10.1162/qss\\_a\\_00244](https://doi.org/10.1162/qss_a_00244)

DOI: [https://doi.org/10.1162/qss\\_a\\_00244](https://doi.org/10.1162/qss_a_00244)

Peer Review: <https://www.researchprotocols.org/2023/10/e37000>

Received: 31 August 2022  
Accepted: 26 December 2022

Corresponding Author: Daniel W. Hook [d.hook@stat.adelaide.edu.au](mailto:d.hook@stat.adelaide.edu.au)

Handling Editor: Luo Wehman

Copyright: © 2023 Simon J. Porter, Lezan Hawley, and Daniel W. Hook. Published under a Creative Commons Attribution 4.0 International (CC BY 4.0) license.

The MIT Press

RESEARCH ARTICLE

## Recategorising research: Mapping from FoR 2008 to FoR 2020 in Dimensions

Simon J. Porter<sup>1</sup>, Lezan Hawley<sup>2</sup>, and Daniel W. Hook<sup>1,2,3</sup>

<sup>1</sup>Digital Science & Research Solutions Ltd., London, UK  
<sup>2</sup>Centre for Complexity Research, Imperial College London, London, UK  
<sup>3</sup>Department of Physics, Washington University in St. Louis, St. Louis, MO, USA

**Keywords:** categorisation, ANZSRC FoR Codes, Dimensions, machine learning

**ABSTRACT**

In 2020 the Australia New Zealand Standard Research Classification Fields of Research Codes (ANZSRC FoR codes) were updated by their owners. This has led the sector to need to update their systems of reference and has caused suppliers working in the research information sphere to need to update both systems and data. This paper focuses on the approach developed by Digital Science's Dimensions team to the creation of an improved machine-learning training set, and the mapping of that set from FoR 2008 codes to FoR 2020 codes so that the Dimensions classification approach for the ANZSRC codes could be improved and updated.

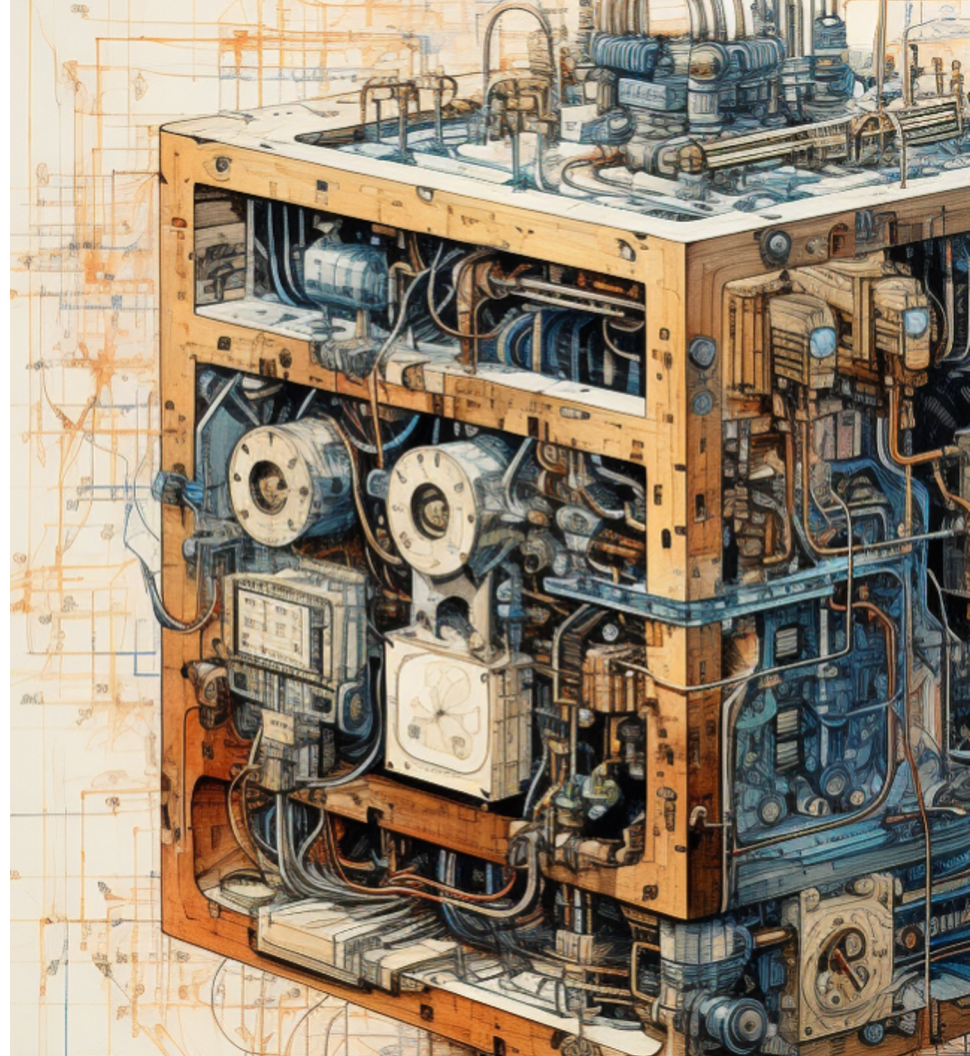
**1. INTRODUCTION**

In 2020 the organizations behind the Australian and New Zealand Standard Research Classification (ANZSRC)—the Australian Bureau of Statistics (ABS), Stats NZ, the Australian Research Council (ARC), and the New Zealand Ministry of Business, Innovation & Employment (MBIE), completed a review of all three components of the classification, including the Fields of Research (FoR) Codes. Over recent years the FoR codes have become an increasingly popular classification system due to their broad subject coverage and well-formulated three-level structure. The FoR codes have been used extensively in Digital Science products such as Almetrics, Symplectic Elements, Figshare, and Dimensions. In particular, Dimensions makes use of machine-learning approaches to classify research objects across its database at a per-object level rather than at a journal level. This was a logical requirement of the Dimensions database as it includes not only publications but also grants, patents, policy documents, clinical trials, and datasets, none of which come with a journal “wrapper” and which, for the purposes of analysis, must have a standardized classification approach applied across them (Hook, Porter & Hawley, 2018). When the change in coding was announced, it was clear that a mapping from the FoR 2008 Codes to the FoR 2020 Codes would be needed and this opened up the opportunity to revisit the machine-learning training set that had originally been used to develop the FoR 2008 classification technology. This paper focuses on explaining the methodology for the creation of a new machine-learning training set for Dimensions.

It is perhaps unusual to focus on the methodology for the creation of a machine-learning training set in an academic paper. However, the creation of training sets is something of a dark art and with the rising use of AI and machine-learning approaches becoming increasingly characteristic in scholarly information infrastructure we argue that greater transparency is in the interests of the communities served by these systems. Calls for responsible research

[https://doi.org/10.1162/qss\\_a\\_00244](https://doi.org/10.1162/qss_a_00244)

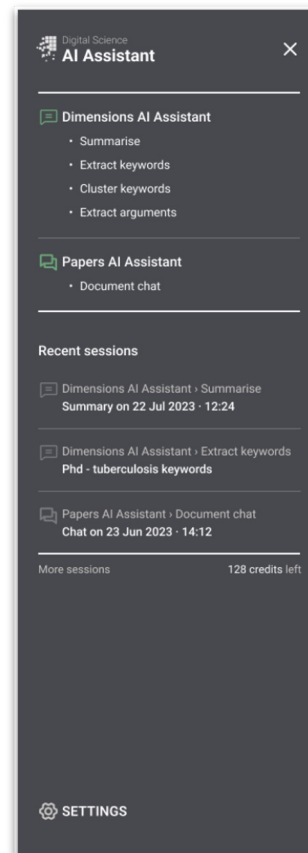
# Challenges



# Designing our ecosystem to meet our aspirations



- **Consistent AI experience across Digital Science Products**
- Thoughtful approach - Place the Researcher in control
- Practical use cases





# Designing our ecosystem to meet our aspirations

- Consistent AI experience across Digital Science Products
- **Thoughtful approach - Place the Researcher in control**
- Practical use cases



# Designing our ecosystem to meet our aspirations



- Consistent AI experience across Digital Science Products
- Thoughtful approach - Place the Researcher in control
- Practical use cases (both positive and negative)



<https://arxiv.org/abs/2401.04022>



# Opportunities



# The beginning of our LLM Journey...



- Dimensions
  - Classification schemes: RCDC, SDGs, FoR Codes
  - Clustering analyses
  - Data annotation and enhancement
- Writefull
- LLMs
  - **Writeful + Overleaf**
  - Dimensions
  - ReadCube
  - OntoChem & Metaphacts



## Dimensions AI Assistant (beta)

Effect of climate on tuberculosis

### Summary

Climate and air quality have a potential to regulate the incidence of tuberculosis, with precipitation, atmospheric pressure, relative humidity, temperature, and wind speed having varying effects on the incidence of tuberculosis [1]. The study found evidence of delayed effects of climate on PTB incidence in Brunei, with varying degrees of magnitude, direction and timing [2]. The prevalence of tuberculosis was influenced by average temperature, rainfall, wind speed, and air pressure, but not by average humidity [3]. Higher temperature extremes increase the risk of tuberculosis outbreaks within groups by increasing physiological stress and dispersal of males, which can lead to rapid intensification of climate-change impacts in natural populations [4].

### Top publications related to your interest

[Climate and air pollution alter incidence of tuberculosis in Beijing, China \[1\]](#)

Chun Yan Zhang, Ang Zhang  
2019, *Annals of Epidemiology* - Article

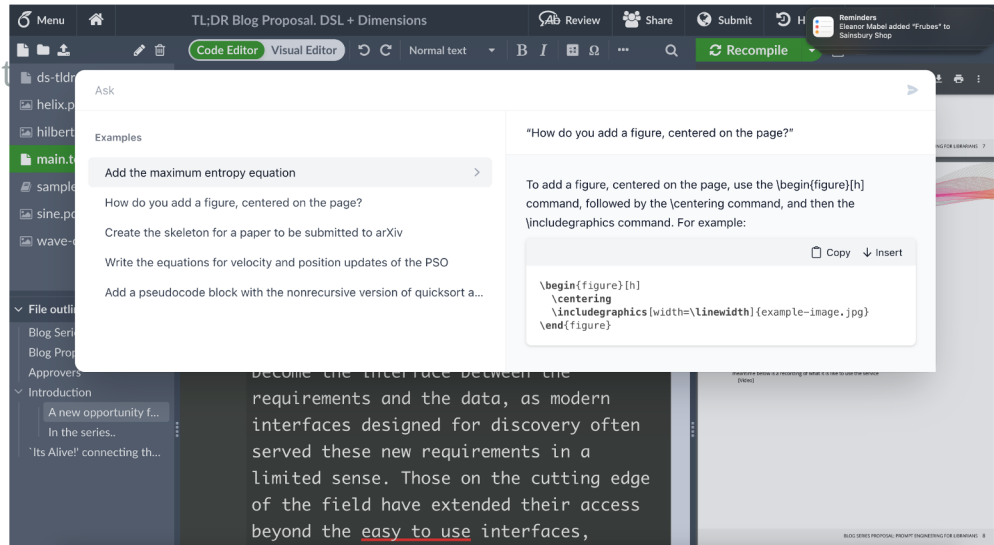
Citations 30 Add to Library 10.1016/j.annepidem.2019.07.003

**Excerpt** Summary & highlights Abstract

... on tuberculosis incidence by improving the concentrations of inhalable particulate matter and sulfur dioxide are potential regulators of the incidence of tuberculosis. The improved air quality contributes to the health of the population. **The impact of climatic indicators on the incidence of tuberculosis was mainly regulated by the environment.** We formulate preventive and regulatory strategies for tuberculosis based on different climatic and air quality indicators. Infectious disease, tuberculosis has long threatened human health and is a major public health issue.

# The beginning of our LLM Journey...

- Dimensions
  - Classification schemes: RCDC, SDGs, FoR Codes
  - Clustering analyses
  - Data annotation and enhancement
- Writefull
- LLMs
  - **Writeful + Overleaf**
  - Dimensions
  - ReadCube
  - OntoChem & Metaphacts



# The beginning of our LLM Journey...



- Dimensions
  - Classification schemes: RCDC, SDGs, FoR Codes
  - Clustering analyses
  - Data annotation and enhancement
- Writefull
- LLMs
  - Writeful + Overleaf
  - **Dimensions**
  - ReadCube
  - OntoChem & Metaphacts



## Dimensions AI Assistant (beta)

Effect of climate on tuberculosis

### Summary

Climate and air quality have a potential to regulate the incidence of tuberculosis, with precipitation, atmospheric pressure, relative humidity, temperature, and wind speed having varying effects on the incidence of tuberculosis [1]. The study found evidence of delayed effects of climate on PTB incidence in Brunei, with varying degrees of magnitude, direction and timing [2]. The prevalence of tuberculosis was influenced by average temperature, rainfall, wind speed, and air pressure, but not by average humidity [3]. Higher temperature extremes increase the risk of tuberculosis outbreaks within groups by increasing physiological stress and dispersal of males, which can lead to rapid intensification of climate-change impacts in natural populations [4].

### Top publications related to your interest

[Climate and air pollution alter incidence of tuberculosis in Beijing, China \[1\]](#)  
Chun Yan Zhang, Ang Zhang  
2019, *Annals of Epidemiology* - Article

Citations 30 Add to Library 10.1016/j.annepidem.2019.07.003

**Excerpt** Summary & highlights Abstract

... on tuberculosis incidence by improving the concentrations of inhalable particulate matter and su  
are potential regulators of the incidence of tuberculosis. The improved air quality contributes to the  
**The impact of climatic indicators on the incidence of tuberculosis was mainly regulated by the en**  
formulate preventive and regulatory strategies for tuberculosis based on different climatic and air  
infectious disease, tuberculosis has long threatened human health and is a major public health issu

# Upcoming Dimensions sneak peak...





## Medical Support Provided by the UN's Scandinavian Allies during the Korean War

Sekwon Jeong, You-Ki Min, Sangduk Lee

2024, Korean Journal of Medical History - Article

The humanitarian motivation of medical support from the three Scandinavian countries during the Korean War cannot be doubted, but the countries also had to be politically sensitive during this period.... [more](#)

 Add to Library

 Summarize

# Upcoming Dimensions sneak peak...



## Medical Support Provided by the UN's Scandinavian Allies during the Korean War

Sekwon Jeong, You-Ki Min, Sangduk Lee

2024, Korean Journal of Medical History - Article

The humanitarian motivation of medical support from the three Scandinavian countries during the Korean War cannot be doubted, but the countries also had to be politically sensitive during this period.... [more](#)



View PDF



Add to Library



Summarize

TL;DR   Key highlights   Top keywords



During the Korean War, Scandinavian countries provided medical support with a humanitarian motivation, but also had to be politically sensitive. They focused on treating wounded soldiers, but also helped civilians whenever possible. This support ultimately led to the establishment of a modern medical system in Korea.



# Upcoming Dimensions sneak peak...

## Medical Support Provided by the UN's Scandinavian Allies during the Korean War

Sekwon Jeong, You-Ki Min, Sangduk Lee

2024, Korean Journal of Medical History - Article

The humanitarian motivation of medical support from the three Scandinavian countries during the Korean War cannot be doubted, but the countries also had to be politically sensitive during this period.... [more](#)



View PDF



Add to Library



Summarize

TL;DR Key highlights Top keywords



1. The humanitarian motivation of medical support from Scandinavian countries during the Korean War.
2. The medical support provided by Scandinavian countries was not only for military purposes, but also intended to help civilians.
3. The medical support activities from Scandinavian countries were flexible and adapted to the changing trends of war.
4. The establishment of the National Medical Center in Seoul after the war was influenced by the common humanitarian purpose of rescuing and treating civilians, and the continued support from Scandinavian countries helped develop the modern medical system in Korea.

# Upcoming Dimensions sneak peak...

## Medical Support Provided by the UN's Scandinavian Allies during the Korean War

Sekwon Jeong, You-Ki Min, Sangduk Lee

2024, Korean Journal of Medical History - Article

The humanitarian motivation of medical support from the three Scandinavian countries during the Korean War cannot be doubted, but the countries also had to be politically sensitive during this period... [more](#)



View PDF



Add to Library



Summarize

TL;DR

Key highlights

Top keywords



1. Medical support

2. UN

3. Scandinavian allies

4. Korean War

5. Humanitarian motivation

6. Political sensitivity

7. Military purpose

8. Civilian treatment

9. Relief

10. Modern medical system

What is the effect of climate change on malaria?



**Dimensions Research GPT**

✓ Talked to ai-api.dimensions.ai ▾

The impact of climate change on malaria has garnered significant attention due to the disease's substantial burden and its transmission sensitivity to environmental conditions. Drawing from the latest research findings, we explore the complex interplay between climate change and malaria transmission, focusing on various aspects such as modeling efforts, socio-economic factors, and regional specificities. This synthesis is based on the top results from a collection of studies, which provide strategic insights into how climate variability and socio-economic development influence malaria dynamics.

### Effects of Climate on Malaria Transmission Dynamics

- **Modeling Studies:** Research by Parham and Michael (2009) utilized dynamic process-based mathematical models to investigate the effects of climate change on malaria transmission, emphasizing the role of rainfall and temperature on mosquito population dynamics and malaria persistence (Parham and Michael, 2009).
- **Socio-economic Interactions:** Béguin et al. (2011) highlighted the opposing eff ●

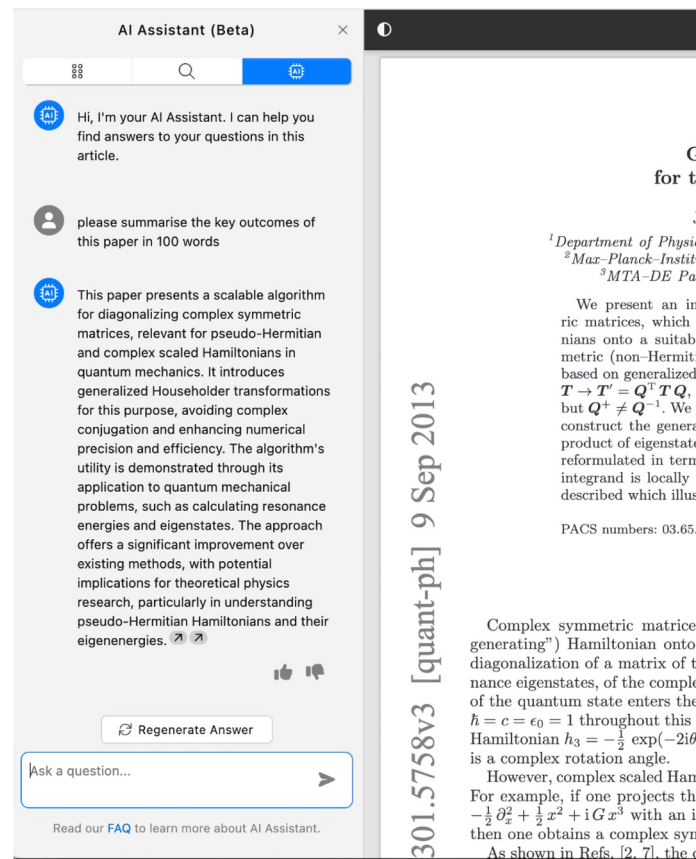


Message Dimensions Research GPT...



# The beginning of our LLM Journey...

- Dimensions
  - Classification schemes: RCDC, SDGs, FoR Codes
  - Clustering analyses
  - Data annotation and enhancement
- Writefull
- LLMs
  - Writeful
  - Dimensions
  - **ReadCube**
  - OntoChem & Metaphacts



AI Assistant (Beta) ×

Hi, I'm your AI Assistant. I can help you find answers to your questions in this article.

please summarise the key outcomes of this paper in 100 words

This paper presents a scalal algorithm for diagonalizing complex symmetric matrices, relevant for pseudo-Hermitian and complex scaled Hamiltonians in quantum mechanics. It introduces generalized Householder transformations for this purpose, avoiding complex conjugation and enhancing numerical precision and efficiency. The algorithm's utility is demonstrated through its application to quantum mechanical problems, such as calculating resonance energies and eigenstates. The approach offers a significant improvement over existing methods, with potential implications for theoretical physics research, particularly in understanding pseudo-Hermitian Hamiltonians and their eigenenergies.

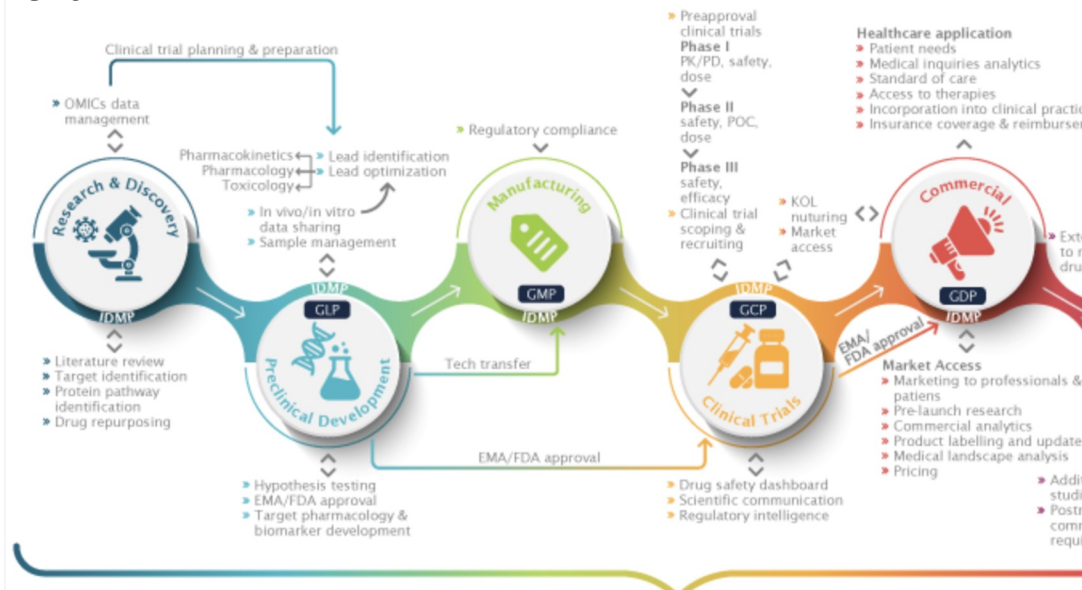
301.5758v3 [quant-ph] 9 Sep 2013

Complex symmetric matrices (generating<sup>†</sup>) Hamiltonian onto diagonalization of a matrix of resonance eigenstates, of the complex of the quantum state enters the  $\hbar = c = \epsilon_0 = 1$  throughout this Hamiltonian  $h_3 = -\frac{1}{2} \exp(-2i\theta)$  is a complex rotation angle. However, complex scaled Hamiltonian For example, if one projects the  $-\frac{1}{2} \partial_x^2 + \frac{1}{2} x^2 + iG x^3$  with an  $i$  then one obtains a complex symmetric. As shown in Refs. [2, 7], the

PACS numbers: 03.65.

# The beginning of our LLM Journey...

- Dimensions
  - Classification schemes: RCDC, SDGs, FoR Codes
  - Clustering analyses
  - Data annotation and enhancement
- Writefull
- LLMs
  - Writeful
  - Dimensions
  - ReadCube
  - **OntoChem & Metaphacts**





# THANK YOU

[www.digital-science.com](http://www.digital-science.com)



 **DIGITAL**  
science

