

Luke Thorburn

My core research interests lie at the intersection: algorithms \cap conflict. My background is in probability and statistics.

🌐 lukethorburn.com

📧 [@LukeThorburn@social.coop](mailto:mail@lukethorburn.com)

✉ mail@lukethorburn.com

🔄 [@luke-thorburn](#)

Education

Oct 2021 – Present

PhD in Computer Science

King's College London

GROUP UKRI CDT in Safe & Trusted AI

TOPIC Bridging-Based Ranking with Diverse Approval

SUPERVISORS Maria Polukarov, Carmine Ventre, Aviv Ovadya

Researching how spatial models such as the left-right political spectrum (and higher-dimensional variants) can be used in recommender systems to perform bridging-based ranking.

Jul 2017 – Jul 2019

Master of Science (with Distinction)

The University of Melbourne

SPECIALISATION Statistics & Stochastic Processes

MINOR Pure Mathematics

SUPERVISOR Kostya Borovkov

THESIS *On integral representations of fractional Brownian motion*

GRADE First Class Honours

Jan – Feb 2018

AMSI Summer School

Australian Mathematical Sciences Institute

Mar 2014 – Dec 2016

Bachelor of Science

The University of Melbourne

MAJOR Mathematics & Statistics

SPECIALISATION Statistics & Stochastic Processes

GRADE First Class Honours

Feb 2014 – Dec 2016

Diploma of Languages (Japanese)

The University of Melbourne

Feb 2012 – Nov 2013

International Baccalaureate Diploma

St Leonard's College

GRADE 43/45 (equivalent to notional ATAR of 99.60)

Affiliations

Mar 2025 – Present

Council on Technology and Social Cohesion

Member

May 2024 – Present

AI & Democracy Foundation

Research Fellow

Jan 2024 – Present

Build Up

Associate

Jun 2023 – Present

MINT Lab, Australian National University

Affiliate

Mar 2023 – Present **GETTING-Plurality Research Network, Allen Lab for Democracy Renovation, Harvard University**
Member

Mar – May 2024 **Center for Constructive Communication, MIT Media Lab**
Research Affiliate

Employment

May – Aug 2024 **Meta**
Research Scientist Intern

Mar 2016 – May 2024 **Resilient Youth Australia**
Data Scientist

Nov 2022 – Mar 2023 **Ofcom (via research consultancy Pattnr)**
Technical Analyst

Apr 2011 – Aug 2022 **Future Problem Solving Program Australia**
International Evaluator & Future Scene Writer

Dec 2019 – Jan 2022 **Hunt Lab for Intelligence Research, University of Melbourne**
Research Associate (Machine Learning & Statistics)

Jan – Jul 2021 **Population Interventions Unit, University of Melbourne**
Data Visualisation

Jun 2016 – Jul 2018 **CSIRO Manufacturing (Biomedical)**
Research Assistant

May – Jul 2017 **The University of Melbourne**
Research Assistant

Dec 2014 – Feb 2015 **Mercer**
Stochastic Modelling

Teaching

Feb 2019 – Jun 2020 **The University of Melbourne**
Mathematics & Statistics School Tutor

Aug – Oct 2017 **Newman College**
Academic Tutor

Jun 2014 – Nov 2019 **St Leonard's College**
Student Study Assistant

Mar 2014 – Nov 2017 **Self-employed**
Private Academic Tutor

Service

Nov 2024 – Present **Meta**
Community Alignment Advisory Group

Aug 2024 – Mar 2025 **Knight–Georgetown Institute, Georgetown University**
Working Group

Jun 2022 – Oct 2023 **Goodenough College**
Treasurer, Members' Council

- Jun 2020 – Dec 2021 **newDemocracy Foundation**
Research & Computational Support
- Jan 2015 – May 2020 **Future Problem Solving Program Australia**
Moderniser of Evaluator Resources
- Dec 2015 – Dec 2017 **Young Australians Cancer Initiative**
Director
- Jun – Oct 2014 **iGEM UniMelb**
Videographer
- Feb – Oct 2012 **St Leonard's College**
Peer Support Leader
- Feb – Nov 2011 **St Leonard's College**
SCOPE Leader

Awards & Honours

- 2021 **Rae and Edith Bennett Travelling Scholarship**
The University of Melbourne
- 2019 **Mathematics and Statistics School Scholarship**
The University of Melbourne
- 2018 **Maurice H. Belz Scholarship**
The University of Melbourne
- 2016 **Undergraduate Research Opportunities Program (UROP) Scholar**
Biomedical Research Victoria
- Leaders in Communities Award (LiCA)**
The University of Melbourne
- 2015 **Keith Frampton Roll of Honour (Member)**
Future Problem Solving Program Australia
- Highest Mark, Weather & Climate Systems**
The University of Melbourne
- 2014 **Second Place, Introductory Microeconomics Film Competition**
The University of Melbourne
- 2013 **Year 12 IB Dux**
St Leonard's College
- Full College, Music and House Colours**
St Leonard's College
- Old Collegian's Award for Positive Involvement**
St Leonard's College
- 2012 **Monash Prize for Excellence & Year 11 IB Dux**
St Leonard's College
- E. Paul Torrance Youth Award**
Future Problem Solving Program Australia

Grants & Funding

- Apr 2025 – Oct 2026 **Deliberative AI Specifications and Infrastructure**
With the AI & Democracy Foundation
SPONSOR ARIA
VALUE TBA
- Jan 2024 – Dec 2025 **Quantifying the Contribution of Online Platforms to Polarization**
Collaboration between Build Up and Sciences Po
SPONSOR Project Liberty
VALUE USD 300,000
- Oct 2021 – Sep 2025 **Computational Social Choice and Machine Learning for Ethical Decision Making**
Centre for Doctoral Training in Safe & Trusted AI
SPONSOR UKRI Engineering & Physical Sciences Research Council
VALUE ≈ GBP 26,000 tuition + GBP 19,000 stipend / year
- Jan – Jul 2023 **Momentum Building for Bridging**
Project Funding from Plurality Research Network Conference
SPONSOR Plurality Institute
VALUE USD 2590.26 (amount determined via quadratic voting)
- Jun – Dec 2021 **Transfer Learning for Argument Operations**
Artificial Intelligence for Decision Making Initiative (Phase 2)
SPONSOR Office of National Intelligence, Commonwealth of Australia
VALUE AUD 50,000
- Aug – Nov 2020 **Augmenting Corpus Analysis through Argument Mining and Networked Notes**
Artificial Intelligence for Decision Making Initiative (Phase 1)
SPONSOR Office of National Intelligence, Commonwealth of Australia
VALUE AUD 20,000

Publications

ACADEMIC PAPERS

Using Collective Dialogues and AI to Find Common Ground Between Israeli and Palestinian Peacebuilders

A Konya, L Thorburn, W Almasri, OA Leshem, A Procaccia, L Schirch, and M Bakker
arXiv, 2025.

Prosocial Media

EG Weyl, L Thorburn, E de Keulenaar, J Mchangama, D Siddarth, and A Tang
arXiv, 2025.

Views on AI aren't binary — they're plural

T Bristow, L Thorburn, and D Acosta-Navas
AAAI/ACM Conference on AI, Ethics, and Society, 2024.

The Moral Case for Using Language Model Agents for Recommendation

S Lazar, L Thorburn, T Jin, and L Belli
arXiv, 2024.

Toward Democracy Levels for AI

A Ovadya, L Thorburn, K Redman, F Devine, S Milli, M Revel, A Konya, and A Kasirzadeh
Pluralistic Alignment @ NeurIPS, 2024.

Societal Sorting as a Systemic Risk of Recommenders

L Thorburn, M Polukarov and C Ventre
ACM Conference on Recommender Systems, 2024.

Views on AI aren't binary – they're plural

T Bristow, L Thorburn, and D Acosta-Navas
arXiv, 2024.

Mapping the space of social media regulation

N Lubin, K Mayberry, D Moses, M Revel, L Thorburn, and A West
MIT Science Policy Review, 2024.

Research Agenda for Sociotechnical Approaches to AI Safety

S Curtis, R Iyer, C Domenico Kirk-Giannini, V Krakovna, D Krueger, N Lambert, B Marnette, C McKenzie, J Michael, E Miyazono, N Mima, A Ovadya, L Thorburn, and D Turan (alphabetical)
SSRN, 2024.

Bridging Systems

A Ovadya and L Thorburn
Knight First Amendment Institute, 2023.

SHAPE: A Framework for Evaluating the Ethicality of Influence

E Bezou-Vrakatseli, B Brückner and L Thorburn
European Conference on Multi-Agent Systems, 2023.

Error in the Euclidean Preference Model

L Thorburn, M Polukarov and C Ventre
International Joint Conference on Artificial Intelligence, 2023.

Polarization as Probabilistic Dependence

L Thorburn, M Polukarov, and C Ventre
International Workshop on Computational Social Choice, 2023.

Error in the Euclidean Preference Model

L Thorburn, M Polukarov and C Ventre
International Conference on Autonomous Agents and Multiagent Systems, 2023.

Optimizing Language Models for Argumentative Reasoning

L Thorburn and A Kruger
International Workshop on Argumentation & Machine Learning, 2022.

The IC Rating Scale as a Measure of Analytic Rigor

L Thorburn, T van Gelder, C Shahan Brugh, S Kim, M Kolb and B Zulauf
SSRN, 2022.

Using argument mapping to improve clarity and rigour in written intelligence products

A Kruger, L Thorburn and T van Gelder
Intelligence and National Security, 2022.

Towards Argument Embeddings

L Thorburn
Online Handbook of Argumentation and Artificial Intelligence, 2021.

Tools to Ease the Choice and Design of Protein Crystallisation Experiments

N Rosa, M Ristic, L Thorburn, GJ Abrahams, B Marshall, CJ Watkins, A Kruger, A Khassapov and J Newman
Crystals, 2020.

Better Feeds: Algorithms That Put People First

A Moehring, A Cooper, A Narayanan, A Ovadya, E Redmiles, J Allen, J Stray, J Kamin, L Sigerson, L Thorburn, M Motyl, M Eslami, N Farid Johnson, N Lubin, R Iyer, Z Arnao
Knight—Georgetown Institute, 2025.

Shaping the Future of Social Media with Middleware

L Hogg, R DiResta, F Fukuyama, R Reisman, D Keller, A Ovadya, L Thorburn, J Stray, and S Mathur
McCourt School of Public Policy, Georgetown University, 2024.

Evaluating recommender systems in relation to the dissemination of illegal and harmful content in the UK

R Armitage, J Gallacher, J Gu, A Hogan, P Howard, B Kollanyi, R Kuchta, L Neudert, and L Thorburn (alphabetical)
Ofcom, 2023.

The Impact of Evidence on Decision Making

A Kruger, T van Gelder and L Thorburn
Centre of Excellence for Biosecurity Risk Analysis, 2021.

Analytic Rigour in Intelligence

A Barnett, T Primoratz, R de Rozario, M Saletta, L Thorburn and T van Gelder (alphabetical)
Hunt Lab for Intelligence Research, 2021.

Enhancing human analysis of large corpora via argument mining and backlinks—a prototype

L Thorburn
Hunt Lab for Intelligence Research, 2021.

GENERAL PRESS

Underexplored Ways to Regulate Social Media

N Lubin, K Mayberry, M Revel, L Thorburn
Tech Policy Press, 2024.

Experiments are the Best Kind of Transparency

L Thorburn, J Stray and P Bengani
Tech Policy Press, 2024.

How to redesign social media algorithms to bridge divides

L Thorburn and A Ovadya
The Conversation, 2023.

What is Media Diversity and Do Recommender Systems Have It?

P Bengani, J Stray and L Thorburn
Tech Policy Press, 2023.

What's the Difference Between Search and Recommendation?

L Thorburn, J Stray and P Bengani
Tech Policy Press, 2023.

Making Amplification Measurable

L Thorburn, J Stray, and P Bengani
Tech Policy Press, 2023.

When You Hear “Filter Bubble”, “Echo Chamber”, or “Rabbit Hole” – Think “Feedback Loop”

L Thorburn, J Stray, and P Bengani
Understanding Recommenders, CHAI, 2023.

Is Optimizing for Engagement Changing Us?

L Thorburn, J Stray, and P Bengani
Understanding Recommenders, CHAI, 2022.

Algorithmic Risks to Collective Decision-Making

L Thorburn
The Future of AI Regulation, 2022.

A Menu of Recommender Transparency Options

P Bengani, J Stray, and L Thorburn
Tech Policy Press, 2022.

How to Measure the Effects of Recommenders

L Thorburn, J Stray, and P Bengani
Understanding Recommenders, CHAI, 2022.

What Will “Amplification” Mean in Court?

L Thorburn, J Stray, and P Bengani
Tech Policy Press, 2022.

What's Right and What's Wrong with Optimizing for Engagement

P Bengani, J Stray, and L Thorburn
Understanding Recommenders, CHAI, 2022.

What Does it Mean to Give Someone What They Want?

L Thorburn, J Stray, and P Bengani
Understanding Recommenders, CHAI, 2022.

How Platform Recommenders Work

L Thorburn, P Bengani and J Stray
Understanding Recommenders, CHAI, 2022.

Charting our way back to normal

T Blakely, T Wilson, L Thorburn and N Grills
Pursuit, 2021.

Talks

Inferred Preferences

Apr 2025 Knight First Amendment Institute

Sortedness + Recommendation

Apr 2025 MIT CCC
Aug 2024 Meta
Mar 2024 American Philosophical Association
Feb 2024 Australian National University
Dec 2023 Imperial College London
Dec 2023 Build Peace
Oct 2023 Metagov
Sep 2023 Harvard University

Good + Bad Conflict

Mar 2025 ISA

Concreteness

Nov 2024 Build Peace

Mapping the Space of Social Media Regulation

Oct 2024 Ofcom
Oct 2024 Harvard

How to Measure the Effects of Algorithms
Mar 2024 Department of the Treasury (Australia)

Error in the Euclidean Preference Model
Aug 2023 IJCAI

Bridging
Jun 2024 University of Notre Dame
May 2024 New Public
Sep 2023 BBC
Jun 2023 Plurality Institute
Mar 2023 Goodenough College
Jan 2023 Plurality Research Network Conference
Sep 2022 London School of Economics

Metaphors for Public Discourse
Feb 2022 Newspeak House

Argument Processing
Dec 2021 University of Technology Sydney
Sep 2021 Laboratory for Analytic Sciences

Symposia

2025 Artificial Intelligence and Democratic Freedoms
Panelist and attendee.

Annual Convention of the International Studies Association (**ISA**)
Speaker and attendee.

Qualitative Coding Workshop on Good + Bad Conflict
Organiser.

Participatory AI Research & Practice Symposium (**PAIRS**)
Speaker and attendee.

2024 Pluralistic Alignment Workshop @ NeurIPS (**NeurIPS**)
Reviewer.

Digital Democracy Residency
Participant.

Artificial Intelligence and Democratic Freedoms
Discussant and participant.

Build Peace
Speaker and attendee.

Australian AI Safety Forum
Participant.

ACM Conference on Recommender Systems (**RecSys**)
Attendee and poster presenter.

Public Interest Social Media Solutions Workshop
Speaker and participant.

International Conference on Computational Social Science (**IC2S2**)
Attendee.

ACM Conference on Economics and Computation (**EC**)
Attendee and poster presenter.

Workshop on Using Deliberative Technologies in Polarized Contexts
Speaker and participant.

Center for Human-Compatible AI Annual Workshop (**CHAI**)
Attendee and poster presenter.

- Convening on Political Economy of AI
Attendee.
- Trust and Human Connection in the Age of AI and Social Networks
Attendee.
- Shaping the Future of Social Media with Middleware
Participant and speaker.
- Project Liberty Institute Summit
Attendee.
- Convening on AI & the Future of the Public Square
Participant.
- Meeting of the American Philosophical Association (Pacific Division)
Speaker.
- Colloquium on Artificial Intelligence and Catastrophic Risk
Participant.
- Kioloa Workshop on Generative AI and Normative Philosophy of Computing
Participant.
- 2023 **Build Peace**
Attendee and workshop facilitator.
- Summit on AI and Democracy
Attendee and facilitator.
- Psychology of Technology Institute Conference
Attendee.
- Launch of Bridging Responsible AI Divides UK (**BRAID**)
Attendee.
- International Joint Conference on Artificial Intelligence (**IJCAI**)
Speaker.
- ACM Conference on Economics and Computation (**EC**)
Volunteer and attendee.
- International Workshop on Computational Social Choice (**COMSOC**)
Reviewer, attendee and poster presenter.
- Plurality Institute Bridging Symposium
Speaker.
- Center for Human-Compatible AI Annual Workshop (**CHAI**)
Attendee and poster presenter.
- International Conference on Autonomous Agents and Multiagent Systems (**AAMAS**)
Attendee and poster presenter.
- The Ethics of Persuasive Technologies
Attendee.
- Optimizing for What? Algorithmic Amplification and Society
Discussant and participant.
- Tech & Democracy: A Better Tech Future Summit
Attendee.
- Designing Tech for Social Cohesion
Attendee and poster presenter.
- Plurality Research Network Conference
Speaker.
- 2022 **Safe and Trustworthy AI Workshop (STAIws)**
Participant and poster presenter.

Attention: An Interdisciplinary Workshop

Speaker.

International Workshop on Argumentation and Machine Learning (**ArgML**)

Speaker.

ACM Conference on Economics and Computation (**EC**)

Reviewer.

The Race to Regulate AI: Global Comparative Perspectives

Attendee.

Oxford-Goodenough Joint Colloquium on AI and Regulation

Participant.

2021 International Conference on Algorithmic Decision Theory (**ADT**)

Attendee.

Conference for Truth and Trust Online (**TTO**)

Attendee.

2020 International Conference on Computational Models of Argument (**COMMA**)

Attendee.

2019 Effective Altruism Global (**EAG**)

Attendee.

INFORMS Applied Probability Society Conference (**INFORMS-APS**)

Attendee.

2017 Summer Institute on Topics in the History of Economics

Participant.

2016 Conference on Neural Information Processing Systems (**NeurIPS**)

Attendee.

Memberships

Jan 2018 – Jan 2025 Statistical Society of Australia

Apr 2022 – Present Trust & Safety Professional Association

Oct 2021 – Present Newspeak House