

Data Week Online 2020

Making data work for everyone

bristol.ac.uk/golding

Data Week Online 2020

The Jean Golding Institute

- A central hub for data science and data-intensive research
- One of 5 University of Bristol research institutes
- Connect multidisciplinary experts across the University and beyond
- Events, training, funding, Ask JGI, The Alan Turing Institute

Our priorities

1. Societal challenges
2. Data visualisation
3. Reproducibility & data governance
4. Fundamental research

Making data work for everyone



bristol.ac.uk/golding

Data Week Online 2020

Date	Event	Speaker
Monday 15 June	Data science and COVID 19 & Data Week Introduction	Kate Robson Brown, JGI Director
Monday 15 June	Intermediate Python	Advanced Computing Research Centre
Tuesday 16 June	Talk: Working at and with The Turing Institute: experiences as a Fellow	Jon Crowcroft, Turing Fellow & University of Cambridge
Tuesday 16 June	Talk: increasing engagement with data	Michael Green, Luna 9
Tuesday 16 June	Introduction to data analysis in Python	Advanced Computing Research Centre
Wednesday 17 June	Do you want to be a data Rockstar?	Luke Stoughton, The Information Lab
Wednesday 17 June	Applied data analysis in Python	Advanced Computing Research Centre
Thursday 18 June	Talk: New data on COVID-19 is undermined by old statistical problems	Gibran Hemani, University of Bristol
Thursday 18 June	Managing sensitive research data: from planning to sharing	Library Research Services
Thursday 18 June	Introduction to deep learning	Advanced Computing Research Centre
Friday 19 June	Deep Learning for Health and Life Sciences	Valerio Maggio, University of Bristol
Friday 19 June	Tour of the Tidyverse	Max Kronborg, Mango Solutions
Friday 19 June	Best practices in software engineering	Advanced Computing Research Centre

Making data work for everyone

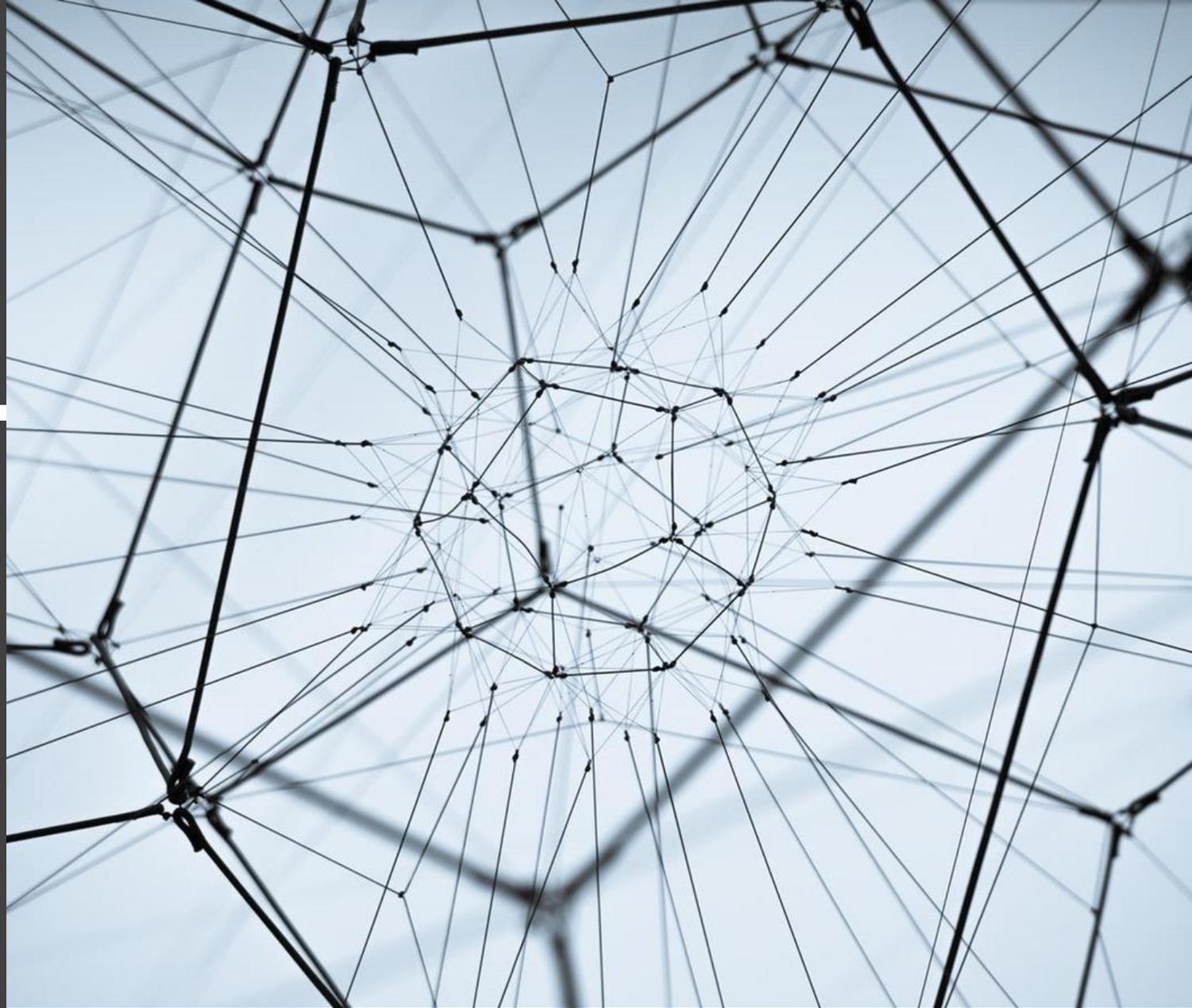
bristol.ac.uk/golding

**The
Alan Turing
Institute**

Turing Fellows

Research Project Stories

**Professor Jon Crowcroft, FRS
Researcher-at-large**



Take homes

- We do foundational research
 - In Data Science, machine learning, AI...
- Can look at papers to date to see that
 - <https://www.turing.ac.uk/research/publications>
- Strategic Partners give us convening power
 - For challenging problems and,
 - key, for interesting & useful data
- We have some innovative ways of working
 - Because we can, but also because we must.

Revolutionise healthcare

A vision for personalised medicine through machine learning-driven diagnosis and treatment plans, and clinicians operating with augmented intelligence



Using machine learning techniques to help improve treatment plans for people living with the challenging symptoms of **cystic fibrosis**



Funding research bringing data science into new treatment options for **cardiovascular disease** with the **British Heart Foundation**

Deliver safer, smarter engineering

Through spearheading new skills, standards, and education, we will deliver a safer, smarter era of data-centric engineering



Digital twin of world's largest 3D printed structure to inform design, track performance, and feed into future 3D structures



Developing machine learning algorithms and data science platforms to understand and **improve air quality** over London

Manage security in an insecure world

Developing new analytic technology and tools that can help us to manage security in an insecure world, at the cutting edge of data science and AI research



Understanding the **mechanics that cause conflict** and identifying multi-scale population areas that are at risk of conflict



Drawing on advances in **AI and machine learning** to address **data wrangling** issues, to help automate the data analytics process

Shine a light on our economy

Applying data science and AI techniques to understand the economy and monitor the real-world impact of business and policy decisions



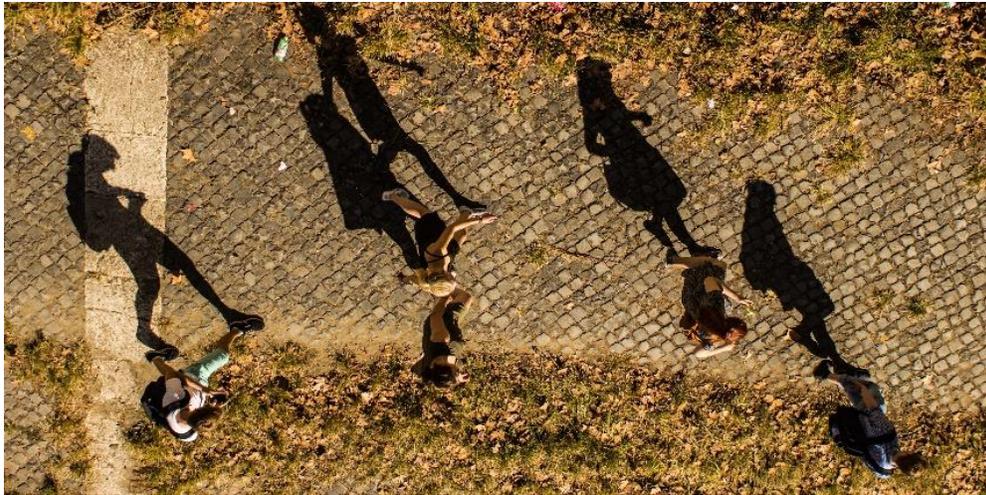
Investigating the impact new technology has on **employment and output**



Understanding how **economic outcomes** in different regions are linked and how **shocks spread**

Make algorithmic systems fair, transparent, & ethical

Delivering fair, transparent, and ethical algorithms through bringing together cutting edge technical skills with expertise in data ethics, privacy, and policy



Data Ethics Group in partnership with Nuffield Foundation and ICO, understanding the ethical and societal implications of data



A new approach to **fairness in algorithm-led decisions**, by looking at the causes of factors that can result in biased decisions

Design computers for the next generation of algorithms

Adapt and refine computer hardware in order to meet the needs of the next generation of algorithms and data scientists



Computer scientists, clinicians, and scientists working to improve the ability of **computers to recognise tumour and cancerous cells**



Training neural networks, and developing related hardware, to be better at **translating millions of words** of online text

Supercharge research in the sciences and humanities

Research organisations are creating enormous sources of data and there are opportunities for data science and AI to assist management and processing



The **British Library** has digitised millions of pages from its collections, facing processing and analysis challenges



Automating processing of **Diamond Light Source's** synchrotron and electron microscope data using **machine learning**

Foster government innovation

Data science and AI can improve the design and provision of public services and inform policy-making across all levels of government



Turing researchers made recommendations to the National Infrastructure Commission for a national **'digital twin' of UK infrastructure**



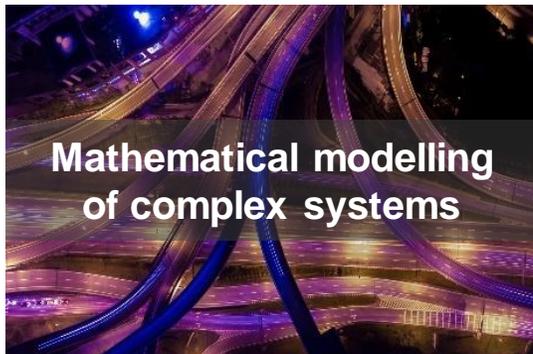
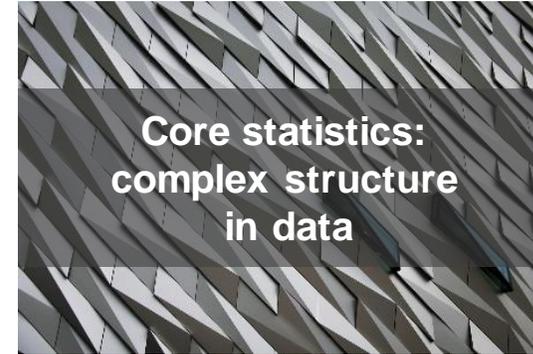
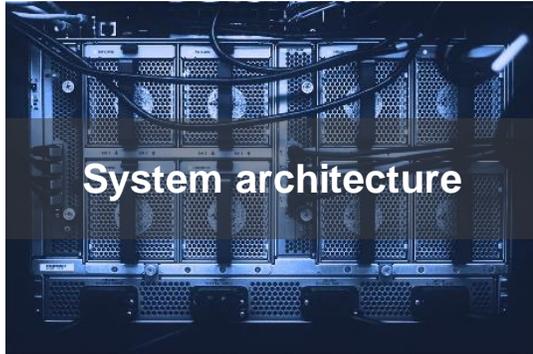
Turing researchers **submitted evidence** to the Science and Technology Committee's inquiry into **algorithms in decision-making**

Paths to Glory

a.k.a. impact



Core capabilities



Research interest groups

- Privacy-preserving data analysis
- Data science and digital humanities
- Logics for data science
- Urban analytics
- Sports and well-being
- Online machine learning
- Fairness, transparency, privacy
- Data Ethics Group
- Topology and geometry
- Low-dimensional structure in data
- Protocol governance: Blockchain & beyond
- Sampling algorithms for data analytics
- Social data science
- Natural language processing
- Mental health
- Data and inequality
- High dimensional statistics

Data Study Groups

Week-long data science sprints

Day 1: Problems are presented by partners

Researchers branch into groups for each problem and discuss different approaches

Day 2-4: Brainstorming, modelling and problem solving

Collaboration and discussion with industrial representative

Day 5: Progress and recommended routes forward are presented.

Output: Report and possible extension of collaboration



Theory&Methods

Challenge Fortnights – new!!!

Foundational Mirror for Data Study Groups

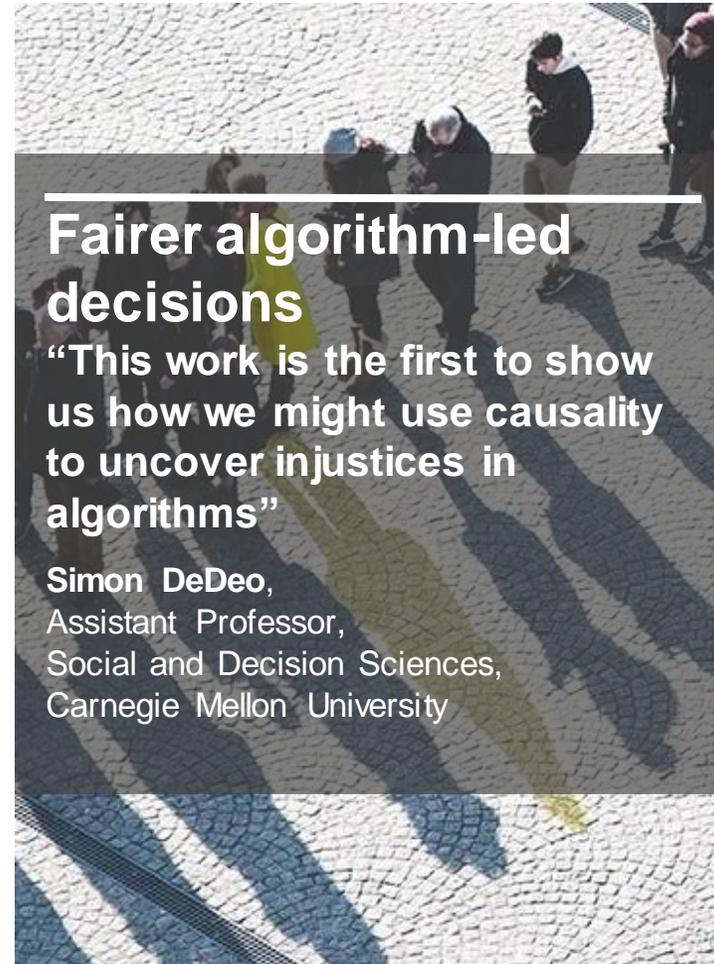
- Φ -ML: Physics-Informed Machine Learning
- Prediction Algorithms with a Causal Interpretation
- Theory of Deep Learning

Some Impact stories



Making simulations simpler
“We bring research code out of the academy and onto the desktop”
James Hetherington,
Director of Research Engineering,
The Alan Turing Institute

Imperial College London 



Fairer algorithm-led decisions
“This work is the first to show us how we might use causality to uncover injustices in algorithms”
Simon DeDeo,
Assistant Professor,
Social and Decision Sciences,
Carnegie Mellon University



A right to explanation
“Engaging and explaining what is going on with algorithmic decisions may help avoid repercussions for the UK economy.”
Sofia Olhede,
Director of Centre for Data Science
UCL

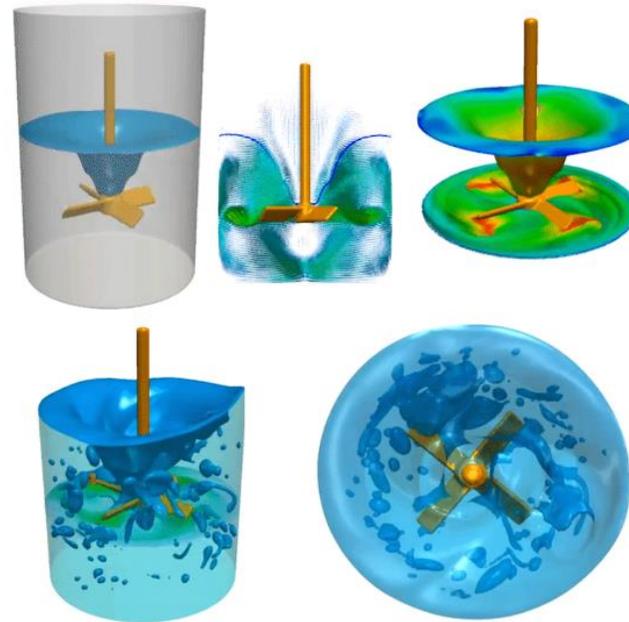
Making simulations simpler



Software engineers at the Turing, in collaboration with partners at Imperial College and UCL, have developed a user interface which aims to make the simulation process more user-friendly, for both academic and industrial communities.

“We bring research code out of the academy and onto the desktop”

James Hetherington, Director of Research Engineering at the Turing



Impact

- Lowering barrier to entry
- Increasing innovation opportunities
- Reproducibility of results
- Tackling key challenge of the Turing’s data-centric engineering programme
- Core code of the app open source

Fairer algorithm-led decisions



Turing researchers from diverse fields have produced a new way of approaching fairness in algorithm-led decisions, by looking at the causes of certain factors that can sometimes result in biased decision-making.

“This work is the first to show us how we might use causality to uncover injustices in algorithms”

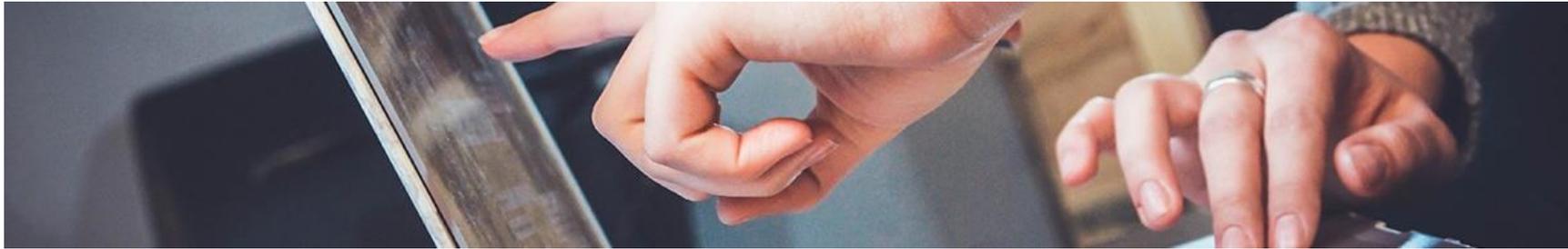
Simon DeDeo, Assistant Professor, Social and Decision Sciences, Carnegie Mellon University



Impact

- Set of technical guidelines for practitioners
- Influenced many leading researchers to use causal methods
- Identifying racial biases in automated US parole procedures
- Potential to have wide-reaching, meaningful effect on everyday life

A right to explanation



Advice from Turing researchers, urging the need for individuals to have a legally-binding right to have automated decisions made about them explained, is helping shape how the new EU general data protection regulations (GDPR) will be implemented.

“Engaging and explaining what is going on with algorithmic decisions may help avoid repercussions for the UK economy.”

Sofia Olhede, Director, UCL Centre for Data Science



Impact

- Helped trigger House of Commons inquiry
- Invited to provide evidence to parliament
- Cited in official EU guidance on automated individual decision-making and profiling in GDPR
- Influenced amendments to UK Data Protection Bill



4 years of being at the Turing

Aside from being handy for the trains to
Europe... 

Interactions & resources

- Privacy-preserving data analysis
- Data science and digital humanities
- Urban analytics
- Fairness, transparency, privacy
- Data Ethics Group
- Low-dimensional structure in data
- Protocol governance: Blockchain & beyond
- Sampling algorithms for data analytics
- Data!
- PhDs
- Papers
- Projects
- Proposals
- Partnerships
- REG
- Feedback
- Fun

Project #1 Maru - Secure Enclaves for AI

- **Funded by D&S Program**
- **Partner with Imperial (not a turing uni then) & Intel**
- **Customer: NCSC - committed&used (tested) code from repo!**
- **software and papers outputs**
- **Influenced design of “Safe Haven” (use for DECOVID!)**
- **intellectual success - liaison via MSR Cambridge**
- **partial failure (sgx;spectre/meltdown)...**
- **followup influence (ARM & RISC-V enclave support different)...**
- **Good example of convening power of Turing**

Project #2 - Trustworthy Digital Identity Systems

- **Funded by Gates Foundation at Turing**
- **partners in World Bank, ID4D, Mosip etc**
- **just closed call for mini-projects (many good submissions!)**
- **redirected some effort to support pandemic work on**
- **secure tracing apps - <https://arxiv.org/abs/2004.04059>**
- **immunity certificate ideas - <https://arxiv.org/abs/2005.11833>**
- **safe haven and secure analytics for DECOVID&tracing...**
- **working on different models of “trustworthiness”**
- **UK is most challenging compared to many developing regions**
- **Great example of value of REG support...**

Research Mapping Exercise&Tools

- Topic Models – sourced from bio, project, papers etc
- PC/RIAC/University Liason expertise
- Crowd sourcing/democratic say
- Bibliometric Studies/NLP
- <https://www.turing.ac.uk/research/research-areas>
- Watch this space...(e.g. on slack channels...)

Data Week Online 2020

Share your participation

#Dataweekonline2020



@JGIBristol

jgi-admin@bristol.ac.uk

bristol.ac.uk/golding

Keep in touch

bristol.ac.uk/golding