The Alan Turing Institute

June 2022

Data-centric engineering at the Turing: the story so far

Delivering safer, smarter engineering

The Alan Turing Institute

This report was prepared by Katy Henderson, in collaboration with the Turing Communications and Partnerships teams. Design by Kimberley Van Russelt.

Contents

Introduction

The data-centric engineering programme has been one of the central components in the success of The Alan Turing Institute.

The research and innovation output from the programme to date stands for itself. Already it is being translated into real-world impact, as you can see from activities such as supporting the development of the world's first 3D printed steel bridge, which now crosses a canal in Amsterdam, and our work in aerospace, which is being widely adopted by industry. Not only this, but the programme is driving forward the creation of data-centric engineering as a discipline, and the take-up of data-centric engineering principles. It is ensuring that the science is advanced globally, through setting standards, changing policies, translating principles into industry practice, and working its way into engineering education, training and continued professional development.

The strong partnership that the Turing has established with the Lloyd's Register Foundation has laid the groundwork for how the Institute can and should be engaging with foundations to add real value to the research and innovation landscape. Engineering is vital to every aspect of the built world; data-centric engineering continues to be crucial to the work of the Turing. As we continue to explore how data-centric engineering can be collaboratively developed and applied across themes and programmes, the possibilities for this discipline are boundless. This is only phase one for the data-centric engineering programme.

Adrian Smith
Institute Director and Chief Executive,
The Alan Turing Institute

Vision

The challenge

The concept of using data in engineering is not a new one. It can be traced all the way from the Ancient Greek Antikythera mechanism, which used data to forecast astronomical positions, through to Lord Kelvin's use of data in the 1800s to develop his mechanical predictor of tides. However, the concept of viewing 'data-centric engineering' as a science in its own right is a new one. Encouraging industries to recognise the value that can be achieved by adopting such a perspective is a new challenge.

The Alan Turing Institute (the Turing) and Lloyd's Register Foundation identified this challenge as an opportunity and partnered together to establish the data-centric engineering programme in 2015. Sparked by the Lloyd's Register Foundation's '**Foresight review of big data**', the data-centric engineering programme was created with a vision of embedding a data-centric approach across all aspects of engineering and related industry.

The goal of the programme was to formalise data-centric engineering as a discipline that would make data science, mathematical and statistical techniques fundamental to engineering practice, making engineering safer and smarter, and leading to a world in which all that is engineered – buildings, transport, energy systems and much more – is better designed, built and maintained, more energy efficient and ultimately safer to use and live in.

The foundations

The Lloyd's Register Foundation's mission is to engineer a safer world by funding research, innovation and education. This perfectly complements the Turing's mission to make great leaps in data science and artificial intelligence (AI) research to change the world for the better. These missions formed the foundation of the **data-centric engineering programme** and supported the development of aims and **Grand Challenges** to help guide the creation, structure and direction of the programme.

The main aims of the first phase of the data-centric engineering programme (aligned to the first grant from the Lloyd's Register Foundation) were to:

- Establish the first generation of research leaders and grow a data-centric engineering community.
- Lead the way, not just in research but also in application and policy.
- Start a global conversation to develop a unified vision for how data-centric methods can empower engineering research, education and professional practice.
- Improve public understanding of the importance of data science and AI across engineering.
- Train and develop future generations of data-centric engineering researchers and practitioners.

Three overarching 'Grand Challenges' were designed to structure the programme and to meet the data-centric engineering needs of society and industry: 'Resilient and robust infrastructure', 'Monitoring of complex systems' and 'Data-driven design under uncertainty'. These Challenges were selected to provide the greatest societal impact. Projects were then initiated with consideration as to how they could deliver on the path to achieving these Challenges.

Underpinning these Grand Challenges, the programme adopted a 'Group Leader' structure, in which established and rising academics were appointed to lead teams of researchers undertaking foundational research in specific engineering and data science areas. The Group Leaders were also tasked with attracting new funding to the emerging field, establishing wider scientific collaborations and engaging with industry and policy-makers to translate research into practice. This open engagement approach has been key to achieving a broad and solid base for data-centric engineering research and practice.

“Within the next five to 10 years, we are going to witness step changes in sensor technology, data-driven intelligent systems, computer science, and algorithms for data analysis, impacting all aspects of the business life cycle – from design to manufacturing, maintenance to decommissioning. Big data is going to bridge the gap from monitoring ‘what is’ to predicting ‘what if?’. Through our partnership with The Alan Turing Institute on the data-centric engineering programme, Lloyd’s Register Foundation has become a major supporter of international research in this field, and brought together global thought leaders to ensure advances in mathematics, computer science and big data can be applied to improve the performance of infrastructure and make the world a safer place.”

Richard Clegg
Chief Executive, Lloyd’s Register Foundation

Achievements

The community

The Turing

The development of a data-centric engineering community began at the Turing and has since spread across both industry and academia. Ideally positioned as a hub with a UK-wide network of university partners, the Turing embraces a core ethos of collaborative working and shared knowledge. Over the course of the programme, a community of data scientists, mathematicians, statisticians, computer scientists, software engineers and engineers has grown within the Institute. The community shares knowledge by working on joint projects and community building activities such as regular reading groups and seminars, workshops, digital communication channels and a community newsletter.

Through the Turing's university partners, and other universities both within the UK and internationally, the data-centric engineering community is intrinsically connected with academia. For example, Imperial College London researchers are the 'Grand Challenge Leaders' for the data-centric engineering programme's Grand Challenges, and Group Leaders represent UK universities including Oxford, Cambridge, Warwick, Southampton, Newcastle, University College London and Queen Mary University of London.

The first generation

The Turing has directly employed 46 postdoctoral research associates and research assistants to work on projects during the first phase of the programme with an additional 21 researchers employed through programme funding at partner universities. These early career researchers

are the first explicitly employed in the field of data-centric engineering. They have gone on to new positions in academia and industry, taking the skills, knowledge and contacts acquired through their work on the programme into their new roles. These are the first generation of data-centric engineering researchers.

Furthermore, existing members of the data-centric engineering programme are being recognised by industry for work in this field:

- Adam Sobey (University of Southampton) is co-lead of the Marine and Maritime Group. He has been invited to hold a non-executive director position with Theyr, an organisation focused on high-precision weather and ocean forecasting.
- Myriam Neaimeh (Newcastle University), Vehicle Grid Integration Group Leader, has been appointed to national roles on UK Power Networks' Innovation Council and in the Office for Zero Emission Vehicles, inputting into policies and setting standards at a national level.
- Andrew Duncan, Group Leader and Theme Leader for Digital twins complex systems engineering, has been appointed Senior Research Scientist at metaverse technology company Improbable.

A number of PIs and programme leads have received high-profile research appointments and awards, reflective of the high level of recognition they have within the scientific community. Mark Girolami (Data-Centric Engineering Programme Director from 2017 to 2021) and Jennifer Whyte (Data-Centric Engineering Grand Challenge Leader) have both been awarded Royal Academy of Engineering Research Chairs. Mark Girolami was elected to the Sir Kirby Laing Professorship of Civil Engineering at the University of Cambridge where he succeeds the Lord Robert Mair and was appointed the Turing's first Chief Scientist in 2021. Julie McCann (Grand Challenge Leader) is Deputy Director of the PETRAS National Centre of Excellence for IoT Systems Cybersecurity. Omar Matar (Strategic Leader) and Leroy Gardner (Researcher) were elected to the Fellowship of the Royal Academy of Engineering. Theo Damoulas (Deputy Programme Director 2019-21 and Group Leader) was awarded a highly selective Turing AI Acceleration Fellowship from EPSRC and promoted to Professor of Computer Science and Statistics at Warwick.

Researcher spotlight

Lassi Roininen
Data-Centric Engineering Postdoctoral Research Associate

- Research focuses on developing methods in uncertainty quantification for large-scale inverse problems.
- Has now **established his own research group** at LUT University, Finland, aiming to use deep mathematical ideas for machine learning with applications in industry, finance and science.
- Identifies the **strong international collaborations** within the data-centric engineering programme as having key importance in his ongoing research output, industrial and scientific collaborations, and significant grant funding, especially from the Academy of Finland.
- Is now **actively fostering the worldwide dissemination of data-centric engineering** by heading up the LUT Master's Programme in data-centric engineering, and the MSc double-degree programme at LUT and Università della Svizzera italiana, Switzerland.
- Is PI for a project which supports African PhD training via doctoral double-degree programmes in applied mathematics between LUT and Bahir Dar University (Ethiopia), University of Rwanda and the African Institute for Mathematical Sciences in Rwanda.

“My collaboration with the programme was a major contributor towards getting a tenure track position at LUT University in 2018. I’m now in the process for becoming a full tenured professor by autumn 2022."

Researcher spotlight

Virginia Aglietti
Data-Centric Engineering PhD Student

- Spent the final two years of her PhD (University of Warwick) working on data-centric engineering projects.
- Research focused on developing scalable Bayesian models for spatio-temporal processes within the London Air Quality project and on incorporating causality within decision-making algorithms to improve digital twins’ forecasting capability.
- Says that joining the data-centric engineering programme **crucial for her career development**, by enabling **building of collaborations** with researchers inside and outside the Turing.
- The programme **supported visits** to the University of Sydney and Data61 for a four-month period, which **led to publications** in top machine learning conferences.
- Completed a **fully-funded internship** at Microsoft Research, which opened up several industry collaborations.
- Is now utilising all she has acquired from her time on the programme in her **new role in industry**, as a Research Scientist at DeepMind.

“The programme gave me the opportunity to work within a diverse and inclusive environment that has fostered and nurtured my interests in a field that has traditionally been male-dominated. It has given me the motivation to further pursue my research agenda and work on problems that have a direct societal and industrial impact.”

Researcher spotlight

Chris Oates
Data-Centric Engineering Group Leader

- Group Leader on the programme from 2017 to 2021.
- Led on projects including probabilistic numerics, inverse problems, and statistical techniques for engineering with advanced materials.
- The programme was **integral to Chris building international research collaborations** and establishing the annual ProbNum workshops – now the de facto international conference for Probabilistic Numerics.
- Research outputs include: Bayesian Probabilistic Numerical Methods published in the 2019 SIAM Review (one of the three most downloaded papers in the journal in 2020-2022); two powerful algorithms for Bayesian computation, called 'Stein Thinning' and 'KSD-Bayes', both to feature in the Journal of the Royal Statistical Society: Series B.
- Research laid the foundations for **successful grant applications** to the British Heart Foundation-Turing call and to EPSRC.
- Was **promoted directly to Professor** from Senior Lecturer in 2018, in part due to the recognition that Chris' data-centric engineering-related research was generating significant interest and seemed set to make considerable impact.
- In 2021 was elected to Chair Elect of the Computation section of the International Society for Bayesian Analysis (ISBA).

"I would not normally have been exposed to the cutting-edge engineering applications I found at the Turing within the academic environment at Newcastle. This was only made possible by my collaboration with the Turing and the data-centric engineering programme."

Researcher spotlight

Eky Febrianto
Data-Centric Engineering Postdoctoral Research Associate

- Research focus on the development of the **statistical finite element method** and its application to structural health monitoring of large infrastructures such as bridges.
- Actively engaged in **international collaboration** with several academic institutions (University of Cambridge, Imperial College London, York University, Queen's University Canada, Czech Academy of Sciences) and industry (Network Rail, EPIC Games).
- Won **best postdoctoral presentation award** at the UK Association for Computational Mechanics Conference 2021 for his work on "Digital-twinning of self-sensing railway bridge using statistical finite element method".
- The programme sponsored his participation in Advance HE's **Diversifying Leadership** programme, which aims to support young researchers from underrepresented backgrounds in becoming future leaders in their field.
- Now **starting his own research group** at James Watt School of Engineering, University of Glasgow, focusing on computational mechanics, computational geometry and computational statistics with application to engineering problems.

"The data-centric engineering programme has connected me to the international community of data-centric engineering and has facilitated collaborations to address important engineering problems. Not only this but The Alan Turing Institute has supported my development as an early career academic through participation in international conferences and leadership training."

Researcher spotlight

Ruchi Choudhary
Data-Centric Engineering Group Leader

- Leads the Built Environment group on the data-centric engineering programme.
- Research aims to develop new mechanisms to leverage data science and simulation to support energy-efficient built environments.
- Working with the programme has had **transformational impact** on the Group's digital twins work, thanks to the collaboration with data-centric engineering researchers and the Turing's Research Engineering team.
- By connecting with the programme and its community, each member of Ruchi's research team has been able to work effectively with data models. Most have **co-authored papers** with other researchers at the Turing.
- Being a part of the programme has also provided the opportunity to **work across programmes**, developing a project with the Turing's urban analytics programme that has **broadened the research** and resulted in new collaborations.
- A direct result of being part of the programme has been the creation of **new international partnerships** for the Group in Singapore.

"Having my research group based at the Turing has opened us up to a range of new collaborations, across teams, programmes, and internationally. The data-centric engineering programme has played a very important role in pushing our research forward."

Researcher spotlight

Jennifer Whyte
Data-Centric Engineering Grand Challenge Leader

- Grand Challenge Leader on the programme from 2017 for the challenge 'Data-driven design under uncertainty'.
- Leads on projects including design change in digital twins, data-driven design assurance, retrofit in the built environment, and data-driven design in civic infrastructure.
- **Interdisciplinary research** has brought together mathematicians, engineers from civil engineering and engineering systems, and design researchers.
- Presented research at the TUM Georg Nemetschek Institute Symposium on Artificial Intelligence in the Built World, and published in **leading journals** such as Automation in Construction.
- Grand Challenge research has **informed policy t**hrough Jennifer's membership of the Digital Framework Task Group, the Transforming Construction Advisory Group, and the Construction Leadership Council.
- Involvement in the programme laid the foundations for **successful grant applications** to Innovate UK and EPRSC.
- Is now Head of the School of Project Management in the Engineering Faculty at the University of Sydney, which aims to be a world-leader in project management for the international research community and senior project sponsors.

"Through working with the Turing, I have been able to collaborate with both excellent peers and the upcoming generation of researchers. By connecting data-centric approaches with the design of engineering systems, these collaborations have significantly enriched my work. As my team of postdocs have moved on to new roles they are growing an international network of collaborators at leading institutions, and are now disseminating this interdisciplinary approach. This was all made possible by the data-centric engineering programme."

A global discipline

To have greatest impact, the programme has always recognised that data-centric engineering has to be adopted on a global basis. From its foundational community within the UK, the programme has worked to develop international connections to drive forward data-centric engineering activity across the world.

So far, eight international memoranda of understanding have been signed across four continents with: the Advanced Manufacturing Centre at the University of Sheffield and Maritime Digital Hub in the UK, the Canadian Statistical Sciences Institute, the Oden Institute (the University of Texas at Austin), the Finnish Centre for Artificial Intelligence (FCAI), Visual Intelligence (Norway), the University of Western Australia and the University of Sydney. These agreements formally recognise shared ambitions around embracing data-centric methods in engineering. They encourage the co-development of activities, collaborative research, sharing knowledge, and hosting events and exchange visitors.

Active projects include the collaboration with University of Western Australia researchers on redeveloping the 'finite element method' (see right), and the Fundamentals of Statistical Machine Learning Group working with FCAI using machine learning to improve the reliability of wireless communication systems. The Turing is also working with Data61, the data and digital specialist arm of Australia's Commonwealth Scientific and Industrial Research Organisation (CSIRO), to support the development of strategies for AI use and implementation. This collaboration aims to mutually benefit both the UK and Australia in using AI to help deliver on each country's national challenges and priorities.

The programme has also developed collaborative international projects with partners based in Amsterdam, San Francisco and Singapore.

In 2019, the programme hosted an International Workshop on Data-Centric Engineering at MIT. This event brought together world-leading researchers to collectively articulate a unified vision for data-centric engineering as a research discipline and its use within professional practice.

Australian collaborations: connecting across the globe

The programme's work with researchers in Australia has created an important connection with the natural resources and offshore industries sectors, specifically with safety and environmental concerns. The programme has memoranda of understanding with both the DARE (Data Analytics for Resources and Environments) Centre at the University of Sydney and TIDE (Transforming Energy Infrastructure Through Digital Engineering) at the University of Western Australia. Both centres are funded by the Australian Research Council and are aiming to boost data science skills in the natural resources sectors and offshore energy environments, respectively.

One of the initial outputs of the collaborations with Australia has been the redesign of the 'finite element method' by a group of researchers from the Turing, University of Cambridge and University of Western Australia. The method is being applied to furthering understanding of internal ocean waves (solitons), which regularly occur on Australia's North West Shelf and are a threat to critical offshore infrastructure such as wind turbines.

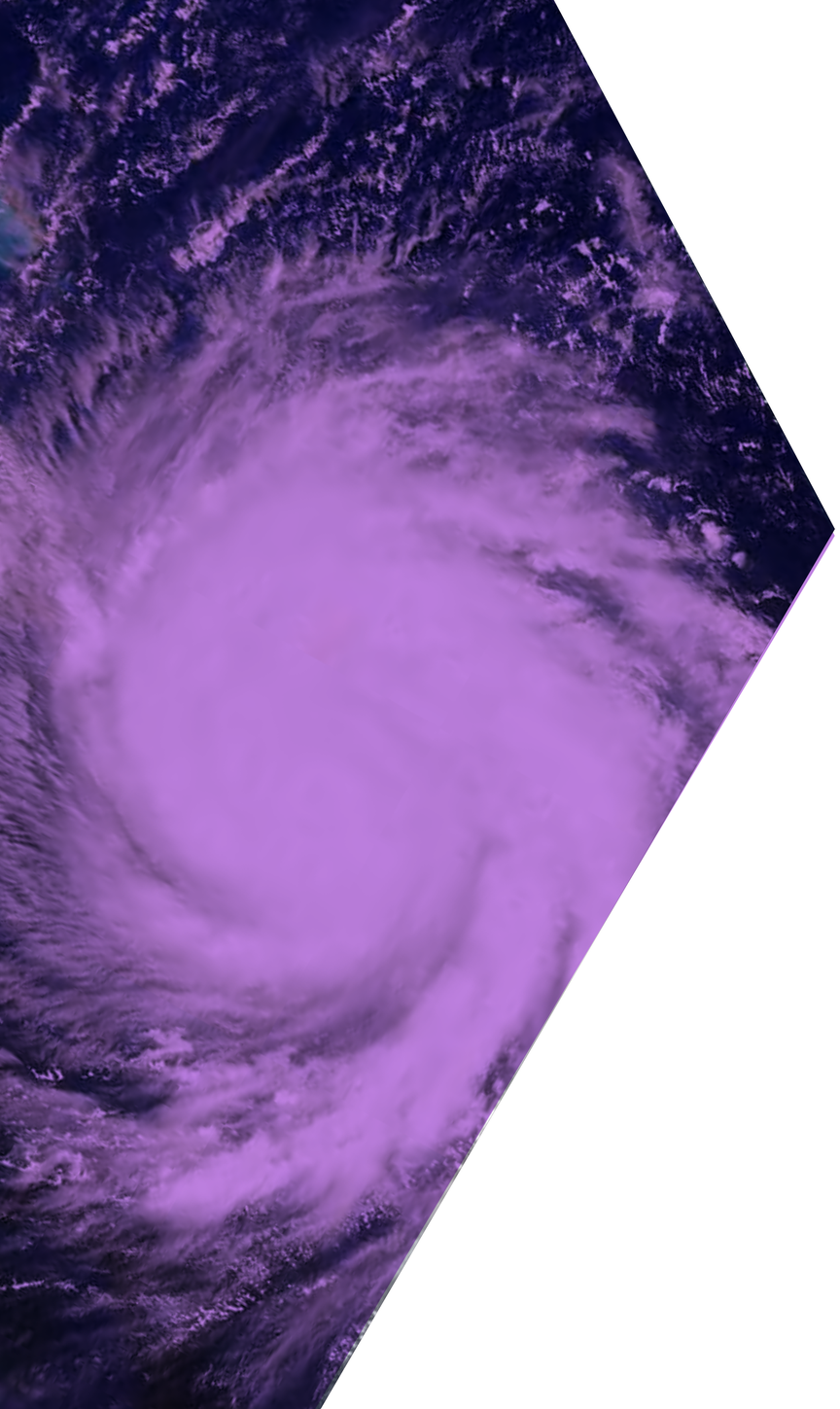
An independent discipline

The fruits of this international network can be seen growing independently of the programme. Data-centric engineering is establishing itself as an academic discipline. LUT University in Finland now has a MRes course in data-centric engineering. Queen Mary University of London has a pilot 'Centre for Doctoral Training in Data-Centric Engineering' that aims to upskill people working in industry, through building a EngD portfolio while staying in employment. The University of Exeter has its own 'Data-Centric Engineering Group'. The term 'data-centric engineering' is becoming more frequently used outside the Turing: Google searches show independent uses of the term growing year-on-year.

"Certain engineering professions are being completely transformed because of our ability to make measurements and gather data in resolutions and volumes that we've never had before. Consequently, there's an urgent need for the engineering professions to gain a much more sophisticated level of the use and exploitation of data – both professionally and in terms of how we deliver education in the engineering sciences."

Mark Girolami
Chief Scientist, The Alan Turing Institute

Taken from '**The data-centric engineering revolution**' in *Civil Engineering Surveyor*



Taking the lead

The data-centric engineering programme has positioned itself as a leader in encouraging the widespread adoption of data-centric principles, methods and tools within engineering practice and policy, ensuring that uptake is truly embedded within the sector.

Changing policy and practice

Central to the Turing's mission is connecting researchers with businesses and public and third sector organisations so that its research can be applied to real-world problems.

The Grand Challenge and Group Leaders have been central to leading the way for changes in policy and practice within the UK. These leaders have provided input into inquiries and calls for evidence and reports, and joined committees and advisory groups for organisations such as the British Standards Institution, National Infrastructure Commission, Parliamentary and Scientific Committee, House of Lords select committees, Department for Digital, Culture, Media and Sport, Committee on Standards in Public Life, Construction Leadership Council, Greater London Authority, Foreign Commonwealth and Development Office, Office for Zero Emission Vehicles, UK Research and Innovation, HM Treasury, the Cabinet Office, Joint Biosecurity Centre, National Grid, and Royal Academy of Engineering.

One field in which the programme is leading the way is the marine and maritime sciences, which do not yet widely use data science and AI. The Turing is developing a roadmap for AI in the marine and maritime industries and the programme has signed a memorandum of understanding with the Maritime Digital Hub, with the intention of being a facilitator to signpost stakeholders in the field towards relevant funding, collaborations and activities.

Fostering public understanding

Leading a public conversation about how data science and AI can improve engineering safety has also been an important goal for the programme, and this has been achieved with the support of the Turing's Communications and Engagement team.

The Turing has upskilled researchers with media training, and has actively reached out to media outlets to provide interesting stories about the programme's work and impact.

The result has been that over the course of the first phase of the programme, 274 media articles, features, blogs and videos have been published. Projects and research have appeared in both mainstream and industry press, including the BBC News website, BBC Radio 4 and Radio Cambridgeshire, Financial Times, Independent, Wall Street Journal, Medium, Daily Mail, Time Out, Naval Architect and Physics World.

Mark Girolami, as Programme Director, delivered the '2020 IET-BCS Turing Talk', a prestigious public lecture series, to capacity audiences in London, Manchester and Belfast. The talk, "Digital twins: the next phase of the AI revolution?", has been viewed nearly 3,000 times on the Institution of Engineering and Technology's website.

The mini-bridge: an educational tool

The 'mini-bridge', a scale version of the 3D printed steel bridge that was installed over an Amsterdam canal in 2021, has been developed as a public communications and education model. It was showcased at the 2019 Great Exhibition Road Festival, the 2019 Lloyd's Register Foundation Conference, the 2020 IET-BCS Turing Talk and the 2019 Turing-Accenture Workshop. The mini-bridge has enabled researchers to demonstrate the different aspects of the project in an accessible and interactive way, and the public has engaged with it strongly.

DCEng SUMMIT

Our first global summit

As the programme has progressed, these activities have laid the foundations for the Turing as a leader in, and an advocate for, the potential of a data-centric approach to engineering. The culmination of this activity was the **DCEng Summit** in September 2021.

This major digital event at the end of the first phase of the programme was a global call to action. It brought together world-class thinking on the application of data-centric methods to engineering, the adoption of new standards, and current developments in ethics, policy and regulation.

By combining meaningful, multi-disciplinary dialogue with insights from industry, the two-day event helped to grow the global conversation and community around data-centric engineering. The event welcomed more than 70 international expert speakers from organisations around the globe who contributed to 25 thought-provoking sessions. The conversations sparked at the event are ongoing.

Growth

As understanding of the data-centric engineering discipline has grown, the programme has itself grown year-on-year, establishing new partnerships (academic, in-kind and financial) and leveraging new sources of funding to support future growth.

The programme has established relationships and worked with several industry partners on projects, including lead industry players and start-ups in the UK and international companies. These include US software and data analytics company Splunk, UK start-up Cervest, Rolls-Royce and NATS. Projects work with a huge range of partners including IBM, Greater London Authority, Microsoft, Procter & Gamble, National Rail and Transport for London. The programme has also seen researchers leading on grant applications with a range of funding bodies such as EPSRC, ESRC, ERC, Innovate UK, NERC and Research England.

In addition to leveraging funding from the AI for Science and Government programme, the data-centric engineering programme has also been involved in several large EPSRC grants. These include CoSInES (Computational Statistical Inference for Engineering and Security), DataSig (Unparameterised multi-modal data, high order signatures, and the mathematics of data science), PREMIERE (PREdictive Modelling with quantIfication of uncERtainty for MultiphasE systems), and ROSEHIPS (Revolutionising Operational Safety and Economy for High-value Infrastructure using Population-based SHM). Applications were subject to intense peer review scrutiny, and these awards reflect the quality of the research and pertinence of application.

In total, additional funds (grants and other cash income) connected to data-centric engineering projects led by the programme's leads and project PIs comes to £83.6m. This represents a more than 8x return on investment for the initial £10m provided by the Lloyd's Register Foundation.

NATS: an evolving partnership

A particularly strong relationship has been developed with NATS (formerly National Air Traffic Services). The partnership has gone from strength to strength, growing from an initial data study group to a successful joint submission to UKRI for a £12m Prosperity Partnership. The scientific vision for this exciting partnership is the creation of the technical building blocks and conceptual roadmap for a safe and assured transition to automation in air traffic control. The central challenge will be building a digital twin of UK airspace. Through the NATS partnership, the programme is delivering direct impact through the application of data-centric methods to the real world.

Programme statistics

94
projects

300+
researchers engaged on projects

377
published (and accepted for publication) research papers and conference abstracts

274
news articles, features, blogs, podcasts and radio interviews

8
MoUs signed on four continents (in USA, Canada, Finland, Australia, Norway, UK)

68
research associates/assistants funded by the programme to date

Use of the term 'data-centric engineering' in non-Turing contexts (number of Google results for the term 'data-centric engineering' that do not include the word 'Turing')

2018	2019	2020	2021	2022
13	102	211	539	744

The science

Making an impact and pushing forward collective change such that engineering across the globe becomes safer and smarter is the ultimate mission of the programme.

Scientific impact on a five-year scale can be difficult to determine, but the programme has already made significant headway in establishing activities that are delivering impact and will continue to do so in the longer term.

Research projects to date have covered an incredibly broad range of activity. Work includes: retrofitting older buildings and working with new housing developments, improving the UK's electric vehicle charging network, refining aircraft engine design, developing new ways of farming both urban and rural, predicting rainfall and water supply in Singapore, maintaining the UK's railway bridges, preventing tailings dam collapse in Brazil, supporting military engineers in post-hurricane rescue efforts, assessing the safety of new infrastructure materials, improving safety on escalators, recognising how cycle networks alleviate traffic congestion, improving understanding of air-polluted urban areas, and much more.

Grand Challenge – Resilient and robust infrastructure

This challenge aims to:

- Use data to ensure the safe operation of major infrastructure systems.
- Improve predictive modelling.
- Improve sensor design and operation.
- Improve robustness against shocks and unexpected incidents.

The 3D printed bridge: a showcase for data-centric engineering

The world's first 3D printed steel bridge, built in collaboration with MX3D and Autodesk, and located in central Amsterdam across the Oudezijds Achterburgwal canal, has become a multi-faceted showcase for data-centric engineering.

The bridge is a result of a cross-disciplinary team of experts collaborating on the future design of public spaces.

The bridge was not only printed by robotic arms, but is also equipped with an innovative sensor network, linked to a 'digital twin' computer model that monitors its performance in real-time. Sensors built into the bridge are constantly gathering data on strain, displacement, vibration, air quality and temperature. Everyone that walks, runs, or cycles over the bridge generates data, which will help researchers to monitor the bridge's structure and how it is being used.

The bridge has the potential to reform how urban infrastructure is designed, built and maintained. 3D printed steel is a new material, so its use could influence conventional construction methods and impact the building industry. The digital twin will help engineers to understand how 3D printed steel might be used for larger scale and more complex building projects.

©Adriaan de Groot

London Air Quality and Project Odysseus: flexible tools, new applications, fresh impact

In 2017, the programme started research into using air quality sensors across London to develop machine learning algorithms and data science platforms to understand and improve air quality. More accurate air pollution forecasts would help local citizens be better informed, and the insights gained from the research would also support the design and evaluation of government policy.

As the pandemic took hold in spring 2020, the London Air Quality project was paused. The team pivoted to begin monitoring activity on the streets of London. The goal was to understand how lockdown was affecting city life, and what interventions were needed to allow the city's nine million people to keep socially distanced.

Named Project Odysseus, and working with the Greater London Authority (GLA) and Transport for London (TfL), the team modified its air pollution algorithms, feeding them with data from London's traffic cameras and sensors to estimate pedestrian and vehicle densities and distances.

The result of the work was a piece of software that the authorities could use to visualise the anonymised, near real-time data, allowing them to monitor pedestrian density and make social distancing interventions where required. TfL implemented over 700 such interventions at the height of the pandemic's first wave, such as moving bus stops, widening pavements and closing parking bays. The Turing's tool provided key data for those decisions.

Looking forward, the researchers are hoping to work with the GLA to monitor high street activity as London recovers from the pandemic, to help understand how social and commercial activity have been affected.

"This collaboration has not only succeeded in knowledge transfer, but has also created a lasting legacy that we intend to build on."

Paul Hodgson
Senior Manager for City Data in the GLA's City Intelligence Unit

Grand Challenge – Monitoring of complex systems

This challenge aims to:

- Use statistical techniques to maintain the operation of vital and complex systems.
- Improve health monitoring.
- Improve predictive maintenance.
- Improve anticipation of rare events.

Foundational tools: the mathematics underpinning real-world applications

Rigorous theoretical research and practices are essential for delivering reliable and robust data science solutions within engineering applications.

A 'Mathematical Foundations' stream was set up to facilitate and develop links between theoretical research and application, building a bridge between the two and projects set up to focus on developing underpinning mathematical tools

One of the foundational tools that has been radically redesigned by the programme is the finite element method (FEM). This predictive tool has been used in engineering and the physical sciences for more than 70 years. The new version reconsiders the FEM from a statistical viewpoint, and allows data to be integrated with the FEM. This work was led by the data-centric engineering programme, with collaborators from an international consortium of researchers.

The new FEM tool has been deployed in looking at internal ocean waves, which are a threat to critical offshore infrastructure such as wind turbines. Research at the University of Western Australia has demonstrated how the FEM's improved prediction methods have provided significant benefit in understanding the waves' impact on engineering design and safety, and on the operations of the offshore energy industry.

"Until now, a missing ingredient in finite element methods has been their integration with observed data – data that can update our understanding of the true process, quantify uncertainty and inform decision making. Working together, researchers from the Turing and UWA have developed a widely applicable methodology for the synthesis of data with FEM – the implications are significant, and we are only starting to explore what is possible from an engineering perspective."

Phil Watson
Director, ARC Research Hub for Transforming Energy Infrastructure Through Digital Engineering

Streamlining jet engine design and manufacture

To analyse the performance of jet engines, aerospace engineers use computer models to simulate the engine's components and the intricate, super-heated airflow through them. But the sheer complexity of these models means that they can take days or even weeks to run, slowing down the speed with which engineers can test new designs.

Researchers led by Andrew Duncan and Pranay Seshadri, both Data-Centric Engineering Group Leaders, have been working with Rolls-Royce to use statistical methods from data science to streamline these models. A key achievement has been the development of algorithms that rapidly home in on the model variables that are most important to the problem. For instance, if the engineers want to make the engine's fan blades more efficient, this new technique will tell them which of the blades' 300+ design variables to focus on. The overall result is that engineers can reduce the number of variables in their models, so that the models run quicker, speeding up the development of more efficient engines. These will use less fuel, resulting in a lower carbon footprint, and Rolls-Royce is now using this technique in the design of its future jet engines.

The principles of this work can also be applied to the engine manufacturing process, providing a potential way to cut waste and costs. And looking ahead, the researchers say that their tools could speed up component design in more radical flight concepts such as zero-emission planes.

"This work has the potential to change the way we design and manage our manufacturing processes."

Shahrokh Shahpar
Rolls-Royce Fellow in Aerothermal Design Systems

Equadratures: providing the keys to scalable digital twins

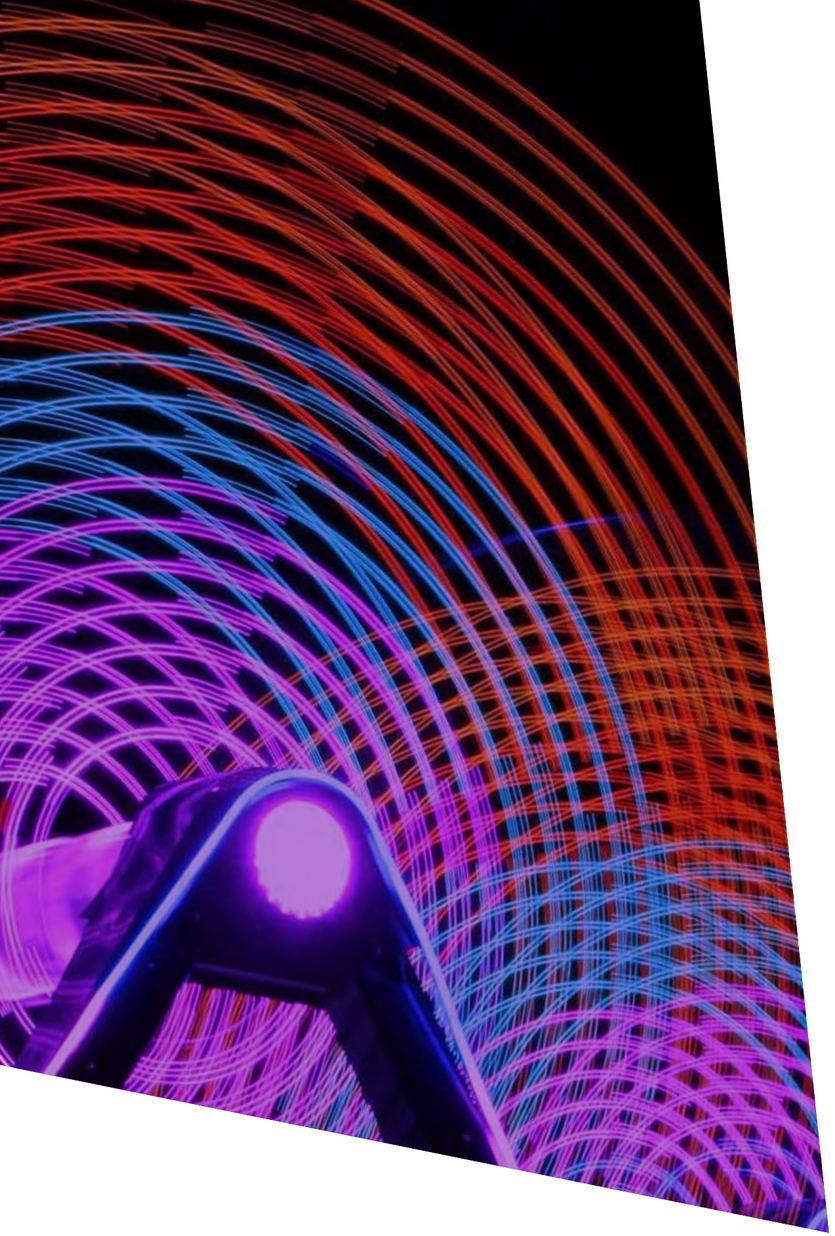
The Equadratures project, developed by the Aeronautics group, is an open-source library for uncertainty quantification, machine learning, optimisation, numerical integration and dimension reduction. In short, it is a library of tools to enable the development of scalable digital twin models.

It is the only existing open platform which enables the rapid deployment of scalable digital twins. This is a necessity in an industry which is built on huge private software frameworks that offer hardly any portability or flexibility. Equadratures already has wide community adoption and support, with frequent third-party contributions to the framework.

The code has been downloaded over 42,000 times to date. The project team has developed workshops to train industry directly in using the code and has held sessions with McLaren, Rolls-Royce, UKAEA and Siemens. The team is now exploring how to engage with a wider range of industries and organisations to broaden how and where the code is used.

"It was one of the best workshops I've attended all year and it was fascinating to hear about where the current state of the art is. It made quite a lot of our current methods look outdated".

Equadratures workshop participant



Grand Challenge – Data-driven design under uncertainty

This challenge aims to:

- Use data to improve the design of engineered systems.
- Create sensor networks and monitoring systems designed for and by data.
- Optimise the collection of data.

Digital twins: transforming predictive technology

Digital twin technologies are increasingly being seen as crucial for the future of engineering. A digital twin is a virtual representation of a physical asset. Data obtained from the operation of the physical asset feeds into the creation of the digital twin, and the twin is then used to test out possible scenarios and improve the design and functioning of the physical asset.

The data-centric engineering programme has positioned itself right on the cutting edge of digital twin technologies, establishing the Turing as a UK leader in this area. Leveraging opportunities provided by Lloyd's Register Foundation, the programme secured additional funding from the AI for Science and Government (ASG) programme to support research into Digital Twins for Complex Systems Engineering.

Programme research to date has provided a suite of foundational tools and methods, building a firm base for the ongoing development and expansion of digital twin usage. This in turn has led to the ASG programme further investing in collaborative cross-theme research on Ecosystems of Digital Twins, which supports the scale-up of digital twin use cases from individual assets to ecosystems, and improves the potential for widespread adoption of these technologies.

The Turing's position as national institute for data science and AI makes it pivotal to the UK in driving forward the innovation and adoption of digital twin technologies. The Turing's next steps will be to further develop the underpinning science, set standards, and work broadly across academia, industry and government to improve digital twin interconnectivity.

Optimising the world's first underground farm

Thirty-three metres beneath the busy streets of Clapham in London lies a farm that is producing subterranean salad greens. In a repurposed WW2 air raid shelter, the programme's Growing Underground project uses soilless hydroponic technology and LED lighting to grow crops year-round, producing 12 times as many crops per unit area as conventional UK greenhouses.

The farm offers a vision for how food production might be increased for a growing global population without using up valuable land resources. However, farming crops without sunlight is an energy-intensive process, so data-centric engineering researchers have developed a digital twin of the farm, to find ways of maximising crop growth while minimising energy use.

The digital twin is fed with variables including water use, relative humidity, temperature, and CO2 and light levels, from both manual observations and automatic sensors. The researchers can then use the model to identify the combination of variables that most improves crop growth. The digital twin can also make forecasts, helping growers to make decisions about the day ahead. If the model predicts that the farm is likely to be too cold, for instance, the grower might add a temporary heater or tweak the lighting. In turn, the model provides data on how effective the measures were.

Thanks to this digital twin, the farm has reduced the time it takes to grow crops by as much as 50%, and increased yields by almost 25%.

"The underground farm provided the perfect opportunity to test our digital twin technology in a unique environment. Bespoke, data-rich computer models such as this will be crucial for optimising the farms of the future, to maximise their output in a changing climate."

Ruchi Choudhary
Project leader and Data-Centric Engineering Group Leader

Data Study Groups: introducing industry to a new way of working

The data-centric engineering programme kick-started the Turing's 'Data Study Group' (DSG) programme. These intensive, five-day 'collaborative hackathons' bring together organisations from industry, government and the third sector with talented multi-disciplinary researchers from academia. The organisations act as 'Challenge Owners', providing real-world problems to be tackled by small groups of carefully selected researchers. The researchers brainstorm and engineer data science solutions, presenting their work at the end of the week.

The data-centric engineering programme developed and sponsored the prototype DSG format in 2016 in collaboration with Airbus, National Grid, Shell, Siemens, Syngenta and Tata Steel.

Sixty-six researchers were involved in these initial data-centric engineering groups, from which the DSG format has flourished and proved a huge success for the Turing, attracting collaborations from Roche, NHS Scotland, Defence Science and Technology Laboratory, GCHQ, National Cyber Security Centre, the Cabinet Office, Department for Work and Pensions, WWF, NATS, AstraZeneca and many others. DSGs have been key in initiating many of the Turing's industry collaborations, and the ideas generated are often a seed for kick-starting larger collaborative research projects across the Turing.

To date, the Turing has run 12 DSGs with 569 unique participants, registering a total investment of £800,000 in additional funding for data science applications. The DSGs were also an important factor in establishing the Data Science for Social Good programme at the Turing, which works with charities with emerging data science needs.

The *Data-Centric Engineering* journal: sharing world-class research

The *Data-Centric Engineering* journal, an open-access journal published by Cambridge University Press, is an important achievement for the programme.

The vision for the journal, which launched in 2020, is to publish high-quality research that uses data-intensive approaches in any of the engineering sciences, so that these emerging ideas can be accelerated in both research and practice. The journal can be read, redistributed, and re-used without any cost or access barriers, including by stakeholders who might not usually have access to academic publications, such as those in industry or policy fields.

The need to communicate this research effectively to a wider audience is reflected in the types of content the journal publishes, including research articles, perspectives and position papers from both academic and industry authors. The variety of articles and authors cements the journal's role as a source of knowledge for industry and academia alike and as a proponent for driving forward the evolution and take-up of this discipline.

Looking ahead

DCE 2.0

The Turing is now positioning itself for 'DCE 2.0': the second phase of this ambitious programme.

Building on the foundational research undertaken within the three Grand Challenges, the programme will be moving towards a more mission-led research focus that reflects current global priorities for the future of engineering. The programme will establish new research pillars which will aim to have transformational outputs, accelerating impact and exploring possibilities in innovation, translation and commercialisation. This sits within the Turing's developing Science and Innovation (S&I) Strategy, which will see greater emphasis on the effective combination of both innovation and research. The S&I Strategy will see Grand Challenges and targeted missions that will provide research and innovation activities with long-term strategic focus, and galvanise the community to achieve impact towards collective goals.

The research undertaken and partnerships established so far have laid the path for the programme to move towards engaging with a broader range of partners and diversity of funding. The Turing will be working to identify new partners with strategically aligned goals to collaborate on foundational research, research translation, impact, capability building and innovation.

The programme will be working with national academies, learned societies and professional institutions to mainstream data-centric approaches within engineering practice and education and to ensure that engineers are equipped with the right skills, tools, training and support to achieve this.

Cultivating the international relationships established to date, the programme will continue to push forward the global adoption of data-centric engineering. As well as expanding and consolidating existing international collaborations, new strategic partnerships and projects will be sought and established globally. The programme is particularly interested in expanding its work with developing nations. This will widen the reach and impact of data-centric engineering and ensure it is central to driving safer and smarter solutions for global challenges.

Data-centric engineering has enormous potential to deliver transformative and lasting economic and societal impact. The data-centric engineering programme at the Turing, in collaboration with its partners, continues to strive towards using all that the discipline can offer to make the world a safer place.

“The data-centric engineering programme at the Turing has demonstrated the clear potential for expanding the use of data within engineering. We now need to make sure that engineers and engineering are equipped with the skills, partners, methods and support needed to accelerate the take-up of data-centric engineering. This will require collective effort from all engineering stakeholders to ensure that data-centric principles are embedded into engineering training, policy and practice.”

Jim McDonald
President of the Royal Academy of Engineering

The Alan Turing Institute

turing.ac.uk
@turinginst