



Centre for
Data Science

Director's Report

Steering Committee Meeting

27 March 2023

Strategy

1. Review of Strategic Plan

CDS members were invited to comment on the Centre Strategic Plan at the 2022 CDS Showcase and Retreat, through a series of structured discussion sessions. The CDS Core Team subsequently reflected on these comments and undertook a review of the Strategic Plan, the corresponding strategies and the associated KPIs.

A copy of the updated Strategic Plan is attached. Any feedback from the steering committee is welcomed. This plan will also be discussed at the upcoming Scientific Advisory Committee meeting which will be held on 27 April 2023.

2. Focus on collaboration

A major strategic aim of the Centre is to increase opportunities for collaborative research in data science. This is facilitated through a number of Centre strategics.

Importantly, CDS reiterates the value it places on the principle of mutual attribution across various QUT Centres, Schools and Faculties.

The Centre is also committed to creating meaningful collaborations with external entities. These include research partners, leaders, members of the Australian Data Science Network, and the Queensland AI Hub.

3. Promotion of equity and diversity

The promotion of equity and diversity in data science continues to be a key strategy in creating an active and inclusive representation of Data Science across QUT. Our funding schemes (e.g., PhD top-ups, First Byte grants) continue to explicitly seek and consider information about the potential for the student-supervisor-project combination to promote diversity and inclusion.

i. **International Women's Day and Women in Data Science**

The Centre hosted a Wikithon on Monday 6 March, and a book club on Wednesday 8 March for International Women's Day. The Wikithon aimed to develop wiki pages for women in data science to raise their profiles. The book club discussed the book *Invisible Women: Data Bias in a World Designed for Men*.

As the lead note for the Australian Data Science Network, CDS also supported the Women in Data Science (WiDS) event [Meet Australia's Rising WiDS Stars](#) which shared the stories of 25 women in data science from across Australia through a marathon five-hour zoom relay.

ii. **Aboriginal and Torres Strait Islander engagement**

In 2022, the Centre initiated an Aboriginal and Torres Strait Islander HDR Pathways Program. The Program aimed to provide research experience for undergraduate students and prospective postgraduate students. In 2023, three students participated in the program and they share their experiences in this [news story and video](#).

4. Increasing capability in data science

i. First and Second Byte Schemes

The First and Second Byte Schemes exemplify the Centre's commitment to increasing capability in data science. After the success of previous schemes the Centre has awarded funding for 9 [First Byte Projects](#) and 7 [Second Byte Projects](#) of \$10K each.

We are currently undertaking a review of outcomes and impacts from the previous rounds of funding.

Preliminary findings of 10 responses so far include:

- 9 publications (4 published, 4 in review, 1 in draft).
- 1 RA enrolled in an MPhil
- 1 RA secured a postdoctoral appointment
- Collaborations across QUT which have led to 2 publications (with different groups)
- Three new cross-disciplinary applications for externally funded research projects.

ii. Leadership

Some of our Program and Domain co-leads have recently stepped down due to other commitments. In light of this, and consistent with the Centre's ethos of inclusion and diversity, we have initiated a renewal program for Centre leaders. Recruitment is underway for the Social Systems Domain Co-Lead, Responsible Data Science Co-Lead, Human Centred AI co-lead and Models and Algorithms co-lead.

The Centre has also initiated a strategy of including postgraduate students and ECRs in the leadership teams for the Programs and Domains. Currently, Dr Aiden Price and Dr Catherine Kim have joined the Government Systems Domain and the Environment and Natural Systems Domain.

iii. Program and Domain funding

To support activities within the Centre's research programs and domains, we are currently inviting applications for 2 rounds of funding for program and domain activities up to \$10K each.

5. Centre Structure

Due to the rising interest in data sovereignty and governance, as well as unbiased AI, the CDS initiated a new research Program on Responsible Data Science. ECR Dr Bernadette Hyland has taken up the role of co-lead of this program. Major strategic directions of the Program include: (i) participation in government policy related to data and data science strategies, (ii) partnerships with Aboriginal and Torres Strait organisations about data sovereignty, literacy and governance and (iii) engagement with international entities focused on these issues.

Performance

Feedback from the Centre's Annual Report indicated that the Centre was tracking well.

Feedback on the Centre's performance was also solicited from CDS members at the 2022 Annual Retreat. In early 2023, these comments were compiled and shared with Centre leaders in group meetings. They formed a basis for discussing and determining key priorities and activities in each of the Programs and Domains for the coming year.

Risk Management

Risks have been identified as being around either people or data.

1. People

People risks include health and wellbeing. Risk management includes supporting staff wellbeing and supporting employment by activities such as: Peers and Pizza, HDR Academy, Seminars on behaviours and code of conduct.

2. Data

Data risks are managed by data governance mechanisms, data ethics workshops teaching good behaviour and practice surrounding data.

Finance and Resourcing

1. Finance

A copy of the Centre's Budget including 2022 actuals and 2023 forecast is attached.

In 2022 the Centre expended \$1,557,754 of \$1.6M.

The Centre looks to expend the full \$1.5M in 2023.

The Centre is reviewing the 2024-2025 budget regarding future appointments and scholarships.

2. Resourcing

i. **Administrative Appointments**

Project Officer Desmond Frescon's appointment ended on 20 February 2023.

ii. **Academic Appointments**

No new academic appointments have been made since the last Steering Committee Meeting.

iii. **Centre Participants**

Centre participants have remained stable since the last steering Committee Meeting.

The Centre is considering co-funded postdoc positions for 2024-2025.

Academic Matters

1. Educational programs

Acknowledging the Faculty and University strategies around external education programs, CDS has recognised a persistent and growing call for advanced profession education courses in data science. To this end, a call out has gone to Centre Participants for interest in developing such courses.

Brand

The Centre Director is aware of the Centre's impact on the University's brand, reputation and marketing strategies. CDS leaders aim to abide by and enhance these strategies.

Initiatives and Engagement Activities

1. Engagement Activities

The Centre continues to hold a number of events through a combination of in-person, virtual and mixed modes.

i. **Data Science in the News**

10 webinars were held in 2022 Data Science in the News webinar series. Topics have been: Inflation and the Australian Economy, Advancing communities in Central Western Queensland, NAIDOC Week, The War in Ukraine, the Power of Data in the 2022 Election, Future of Floods, Post Pandemic, is life just permanently digital?, Emerging Trends in Data Science and AI, Cybersecurity and Men's Health.

One seminar has been held in 2023 (2023 Through a Data Lens) with a second planned for the end of April.

Recordings of the webinars are accessible via the Centre's [website](#).

ii. **Data Science Under the Hood**

This seminar series aims to share fundamental and cutting-edge data science methods that have been developed across difference disciplines

- 25 February 2022 - What We Don't Talk About When We Talk About Machine Learning
- 29 April 2022 - Improving Monte Carlo Estimates with Control Variates
- 20 June 2022 – Mathematical models of coral reef re-growth
- 21 July 2022 – Deep learning for medical imaging applications
- 27 October 2022 - Exploring how older adults make sense of wellbeing data

Recordings are available on the [website](#).

In 2023 this series will be refreshed to "Data Science in the Wild" and will be coordinated by David

Warne. Data Science in the Wild will explore how various problems can be solved through applied data science methods.

iii. Distinguished Visitors Seminar Series

One seminar has been held so far in 2023 by visiting fellow Dr Rob Mahar on Making SMART prophylaxis and treatment decisions. This is on the Centre's [website](#).

iv. Early Career Seminar Series

Our ECRs (HDR students and Postdoctoral Researchers) have continued to lead the 'Early Career Bayes Seminar Series' to enable ECRs to give invited presentations. In addition to being valuable for knowledge exchange, this series provides opportunities for ECRs to build their profile despite covid-19 restrictions to travel and face-to-face meetings.

Since the previous meeting we have had:

- 20 October 2022, Quan Vu (University of Wollongong) presented on Warped Gradient-Enhanced Gaussian Process Surrogate Models for Exponential Families with Intractable Normalizing Constants.
- 2 November 2022, Sebastiano Grazi from University Warwick presented on Methods and applications of PDMP samplers with boundary conditions.

v. HDR Academy

The HDR Academy aims to connect HDR students and provide tailored workshops, opportunities and information to support their research journey.

An end of year event was held on 25 November.

The HDR Academy will be holding its first event for 2023 soon with Industry guest Anton Lord from LeapIn.

vi. QUTBANDS

We have continued to engage with the student group QUTBANDS to cross promote data science activities and opportunities.

vii. Networks and Communications

Twitter followers have grown from 1772 followers (as of 7 October 2022) to 1913 followers (as of 16 March 2023). [LinkedIn](#) followers have increased from 688 followers (7 October 2022) to 983 followers (as of 16 March 2023).

News stories regarding the Centre are also posted on the CDS Website [here](#).

The Centre continues to lead the Australian Data Science Network and an inaugural conference was held at QUT in November 2022. This was attended by 80 people from across Australia.

viii. **SMC Down Under: AMSI-AustMS Workshop on Sequential Monte Carlo**

The Centre is providing support for a workshop on Sequential Monte Carlo which has been cofunded by AMSI and AustMS. This workshop will be held at QUT on 10-13 July 2023.

2. **Research Initiatives**

i. **Scholarships**

The Centrally funded Centre Scholarship for 2023 is currently being assessed. This scholarship was advertised and application have now closed.

All three scholarships for the Hastings Deering Partnership have now been filled, the final student should be commencing shortly.

The remaining 2 scholarships for the ABS partnership have also now been filled and the students commenced on 1 March 2023.

The Centre has decided not to offer top up scholarships in 2023.

The Centre has awarded 1 honours scholarship for 2023.

3. **Major Initiatives**

i. **NGGP Sports Data Science and AI: Creating the Skillsets, Toolsets and Mindsets for Transformation, Innovation, &/or Trajectory 2032**

Globally, sports data analytics was worth ~US\$2.20b in 2020 with a CAGR of ~21.8% by 2028 (Emergen Research). Much of this is driven by growth in sports technology (Sportstech), the intersection of emerging technology with human performance, medicine, and STEM. On-field Sportstech is used to monitor and improve athlete health and performance, and off-field technology involves sports business applications, infrastructure, and engagement.

The global Sportstech industry was worth ~US\$17.9b in 2021 growing to US\$40b by 2026. The Australian Sportstech industry generates US\$3.1b in revenue, with 600+ companies (tripling in the last decade) and employing over 10k people (ATSN Sport Innovation Report, 2022).

This Program will train a sports data scientist cohort to tackle challenges faced by the Australian Sportstech industry. It will meet analytical demand driven by emerging Sportstech technologies for managing athlete health and performance, recovery from injury, and athlete and support staff management, with a focus on personalised outcomes and embracing diversity.

By training multidisciplinary, industry-driven data specialists, the Program will boost the Australian Sportstech sector, accelerate growth and upskilling of new ventures, and support Australia's 2032 Olympic trajectory. Program graduates will be work-ready for sport, and related areas of data-driven human performance, including defence, business, and health.

Chief Investigators	Professor Kerrie Mengersen, Dr Paul Wu, Associate Professor Dimitri Perrin, Associate Professor Divya Mehta, Professor Scott Sisson, Dr Sahani Pathiraja
Universities	UNSW, QUT
Industry partners	Queensland Rugby Union, Platypus Medtech Consulting, Reaction Performance, Again Faster Australia / Fitbox, Microba Life Sciences, Precision Health Alliance, iNSPIRETEK, Paddle Australia, Surfing Australia, Splink, Mitosportz, NVIDIA, Oracle
Student degree type	PhD, MPhil, Coursework Masters (not RTP eligible), Honours
