



Director's Report

Management Committee Meeting

17 August 2021

1. Summary

Since the last Management Committee, we have firmly established the Centre for Data Science as a Tier 1 Centre. We have grown our membership across the University, extended and consolidated our partnerships, expanded our research activities and contributed to major University research endeavours.

Research highlights include the addition of two Domains (Government Systems, Sports Data Science), the appointment of six Postdoctoral Research Fellows, the approval of First Byte projects for ECRs and the current call for First and Second Byte projects for ECRs and MCRs respectively, the expansion of our HDR students and postdoctoral researchers, and the pending appointment of a Professor of Practice as a Data Scientist in Residence.

Engagement highlights include fruition of Research Constellations with the Australian Institute of Sport and the Queensland Academy of Sport, and with the Australian Bureau of Statistics. Each of these comprises an externally funded senior appointment (a Senior Research Fellow and a Chair, respectively) and a set of PhD scholarships. Other co-funded HDR scholarships and postdoctoral research positions have also been established in collaboration with external partners.

A meeting of the Centre's Advisory Board is planned for November 2021, following from the previous meeting held in November 2020.

2. Centre Profile

Leadership

The CDS Core Leadership Team has remained stable since the commencement of the Centre and comprises satisfactory discipline and gender balance:

- Distinguished Professor Kerrie Mengersen (Director), Mathematical Sciences
- Professor David Lovell (Deputy Director), Computer Science
- Professor James McGree, Mathematical Sciences
- Professor Chris Drovandi, Mathematical Sciences
- Associate Professor Richi Nayak, Computer Science.

The Leaders and Co-Leaders of the CDS Research Programs have remained stable for the past year. They are listed below:

Research Programs	Current Lead	Co-Lead
Data for Discovery	James McGree	Moe Wynn
Models and Algorithms	Chris Drovandi	Timothy Moroney and Matt Simpson
Data Focused Decision Making	David Lovell	Margot Brereton
Applied Data Science	Richi Nayak	Simon Denman

The Centre has established two new Domains, in Government Systems and e-Research, and is in the process of establishing another new Domain in Sports Data Science. The Co-Leaders for the other CDS Domains have remained stable for the past year. The full list of CDS Domain Co-Leaders is as follows.

Domains	Co-Lead	Co-Lead
Health and Biological Systems	Dale Nyholt	Dimitri Perrin
Environment and Natural Systems	Kate Helmstedt	Jinglan Zhang
Business and Industry Systems	Adam Clements	Ashish Bhaskar
Social Systems	Char-Lee Moyle	Timothy Graham
Government Systems	Gentry White	Bernadette Giuffrida
e-Research	Matt Bellgard	Andrew Gibson
Sports Data Science	Paul Wu	TBC

Administrative Appointments

Joyce Wang gave notice on 8th February 2021 and resigned as Centre Manager a month later. Plans for a replacement were initiated in February and Becki Cook joined as Centre Manager on 31st May, 2021. This process took much longer than expected and placed a substantial constraint on the functioning of the Centre during the first half of the year.

The Centre has appointed 6 Postdoctoral Fellows aligned to the 4 Research Programs. Another Postdoctoral Fellow, Lina Xu, is funded by the QUT Business School and is working closely with Adam Clements and James McGree on Centre aligned research.

The appointment of the part-time Digital Communications Coordinator, Amy Rawson, has been extended to September, 2021. Amy works on the Australian Data Science Network (ADSN) website and organises ADSN related events.

Positions for two Project Officers funded by the CDS have been initiated to lead the development of research networks and communications (one for CDS and the other for the ADSN). These are still pending approval.

Centre Participants

Since the last Management Committee in March 2020, the Centre has continued to grow its Centre Participants.

The Centre is in the process of reclassifying its members. In accordance with initial recommendations, only the Core Team and Program/Domain Co-Leaders were classified as Chief Investigators and all others were classified as Participants. In accordance with subsequent recommendations, Participants were reclassified as Associate Investigators and a membership nomenclature was introduced. More recently, new criteria have been introduced by the University and most of our members have been classified as Chief Investigators. Some members have elected to remain as Associate Investigators for various reasons.

Participation Type	Participants (Sep 2020)	Participants Aug 2021)
Core Team (CIs)	5	5
Research Program and Domain Co-Leaders (CIs)	13	16 + 1 TBC
Associate Investigators and DataScience@QUT (AIs)	85	86
Centre Affiliate	3	4
Post-Doctoral Fellows	6	6
Students	55	69
Key Partners	8	8
DataScience.Info subscribers	208	246
Data Scientist in Residence	0	1 TBC

A breakdown of CIs and AIs by Schools/Faculties is as follows.

Areas	CIs/AIs (Sep. 2020)	CIs/AIs
School of Mathematical Sciences	26	28
School of Computer Science	17	18
Business School	16	13
School of Information Systems	11	12
School of Electrical Engineering and Robotics	9	9
School of Civil and Environmental Engineering	5	8
eResearch	2	2
Law School	2	2
School of Biomedical Sciences	2	2
School of Built Environment	2	2
School of Civil & Environmental Engineering	2	2
School of Communication	2	2
School of Public Health and Social Work	2	2
IHBI	1	1
QUT Academy of Learning and Teaching	1	1
School of Biology and Environmental Science	1	1
School of Earth and Atmospheric Sciences	1	1
School of Mechanical Medical and Process Engineering	1	1
AD Research	0	1

3. Activities/Events

Covid-19 has provided challenges for the Centre in conducting planned events. Despite this, most planned activities have been undertaken, mostly through virtual means.

Data Science in the News

The Data Science in the News webinar series has continued in collaboration with the Queensland Academy of the Arts and Sciences. In June 2021, the series moved from a fortnightly to a monthly program. Each webinar convenes a panel of four CDS participants and colleagues to provide evidence-based comments on the relationship between data science and a topical issue in the news. This outreach webinar series is intended for a wider audience to increase the understanding of how data science can impact our everyday lives. Recordings of the webinars are accessible via the Centre's [website](#).

Below are the webinars held since the last Management Committee meeting:

- Going for Gold: Data Science and the Olympic Games
- Data under attack: cybercrime and data science
- Over the Horizon
- Bringing Interdisciplinary Science to Bear on Aged Care
- The data science of COVID-19 and vaccinations
- (Data) sharing is caring: The opportunities and challenges of open access data
- How data and data science are reshaping cultural institutions
- Data Science in the News: headline edition
- Journalism old and new
- How we value the humanities

Data Science Under the Hood

The primary aim of this seminar series is to disseminate, to the DataScience@QUT network, fundamental and cutting-edge data science methods that have been developed across several disciplines. This will service to expand the knowledge base of our Centre Participants, to foster interdisciplinary data collaborations by increasing awareness and highlighting challenges. The talks are designed to be accessible to the wider DataScience@QUT community. Assoc Prof Chris Drovandi is the convenor of this series. Recordings of the webinars are accessible via the Centre's [website](#).

Below are the seminars held since the last Management Committee meeting:

- 26 Aug 2021: Data Fusion and Knowledge Discovery using Factorization methods
- 29 July 2021: Manifold Learning and its incorporation to a dimensionality reduction technique to learn data features
- 13 May 2021: A telescopic introduction to multilevel Monte Carlo for simulation and inference
- 8 April 2021: Population Data: Getting to the Heart of Variability
- 25 Nov 2020: Analysing ecoAcoustic data for environmental monitoring
- 29 Oct 2020: Enabling digital transformation of organisations through process mining
- 24 Sept 2020: Human interaction with data

Distinguished Visitors Seminar Series

The CDS has established a Distinguished Visitors Seminar Series. In the first instance, members of our Advisory Committee have been invited to give seminars. Due to covid-19 restrictions, all speakers to date have been virtual. The seminars have been recorded and are available on the CDS Website. Presentations so far have included:

- Professor Antonietta Mira, University della Svizzera italiana; QUT Adjunct Professor
Title: Intrinsic Dimensions
- Professor Ruth Baker, University of Oxford, UK
Title: Quantifying Electrotaxis
- Professor Nial Friel, University College Dublin, Ireland
Title: A short overview of stochastic block models with application to the assessment of competitive balance in the English premier league
- Professor Ruth Baker, University of Oxford, UK
Title: Quantifying the effect of electric fields on single-cell motility
- Associate Professor Joris Bierkens, Delft University of Technology, The Netherlands
Title: Spectral theory and asymptotic variance of piecewise deterministic samplers

Early Career Seminar Series

Our ECRs (HDR students and Postdoctoral Researchers) have also initiated an ‘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.

HDR Academy

CDS has established a HDR Academy to provide a networking forum to build a community of HDRs interested in all things data science. The Academy is convened by Associate Professor Richi Nayak, with the assistance of CDS Postdoctoral Researchers. The Academy comprises around 55 HDR students at present.

The Centre has held online get-to-know-you meetings and two face-to-face meetings. In the face-to-face meetings, student participants have had the opportunity to present their research and themselves to their peers.

Data Science for Social Good

The Data Science for Social Good program has been established by the CDS since the last Management Committee meeting. This pro-bono program is an initiative to provide hands-on experience for CDS members to engage with external groups and use their data science skills to solve some real-world challenges with social impact. This program invites small projects involving problems or challenges having impact on social good from not-for-profit organisations. Small groups of CDS members, comprising students, data scientists and mentors, will aim to deliver solutions to the problems in the form of report, codes, apps, algorithm, models etc, depending on the needs of the projects. The program will be beneficial to both the students and the organisations.

The program was workshopped with a wide range of University members and support staff. The first set of projects is due to commence in September 2021. The inaugural partners are Cancer Council Queensland and FareShare.

Peers and Pizza initiative

The Peers and Pizza initiative aims to support the health and wellbeing of CDS HDR students. Sessions aim to connect students and to provide a platform for students to discuss wellness and other challenges faced by HDR students. Each session included strategies and resources for self-care and coping with HDR life.

4. Research Initiatives

Data Scientist in Residence

The Centre has initiated the appointment of a Professor in Practice as its inaugural CDS Data Scientist in Residence. Pending final approval and acceptance, the appointee, Dr Rachel Thomas, will take up this role for 12 months at 0.5FTE. Dr Thomas previously held a position as an Associate Professor in research ethics at the University of San Francisco, USA, and is also the co-founder of a well-known data science company, *fast.ai*.

First and Second Byte Programs

In early 2021, the CDS rolled out the inaugural First Byte Program for ECR researchers. The second rollout of this Scheme is open now, with a concomitant Second Byte Program for mid-career researchers. These programs seek to fund researchers to undertake high-quality projects that are strategically aligned with CDS Research Programs and Domains. They particularly encourage research collaboration between CDS Participants and the wider data science community, and encourage research that is likely to lead to further grant funding. First and Second Byte projects are funded to a maximum of \$10,000 and \$20,000, respectively.

12 First Byte projects were approved in the first round and have commenced. See

- [CDS First Byte Funding Program 2021](#)
- [CDS First Byte Funding Program 2021 - Round 2](#)
- [CDS Second Byte Funding Program 2021](#)

Co-Funded Scholarships

The Centre is committed to expanding the capacity of new HDR scholarships in the Data Science community at QUT. In the 2020/21 budget, we proposed up to 28 co-funded scholarships. These have proved very popular with our external partners. However, due to restrictions on international students due to covid-19, as well as national initiatives such as the National AI Centre (led by Data61), it has become increasingly difficult to recruit the required number of students with the required skills. We have therefore recently converted a large proportion of the scholarship allocation to additional postdoctoral research positions and First and Second Byte projects.

Co-funded scholarships have been established or are under negotiation with the following organisations:

- Australian Institute of Sport and Queensland Academy of Sports (3)
- Australian Bureau of Statistics (3)
- Services Australia (4)
- Children's Health Queensland (1)
- Oracle (1)
- Food Agility CRC (1)

Five PhD scholarships have also been awarded for the Major Infrastructure Long-Term Monitoring of Benefits and Impacts Program (formerly the Queens Wharf Brisbane Study). This is a substantive cross-University endeavour and includes supervisors: Connie Susiliwati, Char-lee Moyle and Amanda Beatson (QUT Business School), Darren Wraith (School of Public Health), John Scott (School of Justice), James McGree and Paul Corry (School of Mathematical Sciences), Ashish Bhaskar (School of Transport).

PhD Top-Up Scholarships

The Centre offered 5 PhD top-up scholarships as part of the QUT 2021 scholarship round. The same number of top-up scholarships are being offered in the current round. See

- **Centre for Data Science PhD top-up scheme for 2021**

Australian Data Science Network (ADSN)

The CDS has successfully established and is the lead node in the Australian Data Science Network (ADSN). The ADSN seeks to connect the concentrations of expertise in data science across Australia, in order to improve communication, encourage collaboration, expand opportunities, and promote our individual and collective capabilities. It provides a unique opportunity to enlarge the profile of CDS and extend our research and engagement opportunities. The ADSN now comprises 23 Partners across Australia.

David Lovell is Convenor of the ADSN. A Digital Communications Coordinator has been appointed by the Centre until September 2021 at 2 days a week to help facilitate the Network and its activities. These include:

- The establishment of a website for the ADSN: <https://www.australiandatascience.net/>
- Regular meetings of the ADSN (via zoom)
- A Centre Spotlight Series, which involves a small number of Centre Directors or their representatives introducing their Centre to the Network members
- A Deeper Dive series of workshops, with the first workshop to be held soon on the generation of synthetic data.

Data Science Hackathon

The inaugural CDS Data Science Hackathon, “Decoding Business Statistics”, was held on 5-30 July 2021. The lead organisers were Dr Char-lee Moyle (Business), Professor Timothy Graham (Communications) and Dr Farzana Jahan (CDS Postdoctoral Researcher). The aim of the Datathon was to provide an intensive facilitated training event where teams of data scientists, professionals and subject matter experts could come together to engage in training and mentoring, delve into topical real world challenges, develop meaningful insights and leverage or create novel datasets to provide solutions to problems identified by government or industry. The participants also obtained a Professional Certificate of Completion that can be counted as coursework towards a QUT micro-credential through QUTex, and competed for the CDS Decode Datathon Trophy.

QUT Partners in the Hackathon included CDS, the Centre for Future Enterprise, QUTEx and the QUT Centre for Entrepreneurship. External Partners included Mets Ignited, the ARC Centre of Excellence in Mathematical and Statistical Frontiers, Advance Queensland, StartUp Gladstone and Startup Status.

Research Grants Development

Since the last Management Committee meeting, the CDS has been involved in a number of major grant submissions. These include a CRC for Longevity (unsuccessful but will reapply), a CoE for Shaping the Future of Data Science (EOI under consideration), ARC Industry Transformation Training Centres (one under development (focused on Data Science in Agriculture), one in resubmission (Antipathogenic Materials)) and various university-led industry projects (e.g., Stryker).

5. Industry Engagement and Key Partners

The Centre Industry Engagement Officer, Jessie Roberts, has returned from leave and has taken over from Jeremy Farr-Wharton.

Jeremy was instrumental in finalising a number of major funded opportunities, including:

- The externally funded Chair in Data Science with the Australian Bureau of Statistics (\$186,000 p.a., 3 years)
- The externally funded Research Fellow with the Australian Institute of Sport and Queensland Academy of Sport (\$144K+ p.a., 3 years)
- Stryker consortium (scoping project currently contracted)
- Food Agility projects, with Cameron Ralph: some projects established; agreed level B constellation leader and level D coordinator pulled
- Qld Department of Natural Resources, Mining and Energy (\$200K)
- Australian Institute for Marine Sciences
- Cancer Council Queensland.

Jeremy and James McGree also drove substantive discussions around a proposed ITTC in Data Science in Agriculture. This was deferred to the 2021 round and is now being driven by Jessie and Richi Nayak.

Jessie has also been active in recruiting external partners for the recently submitted CoE EOI on ‘Shaping the Future of Data Science’.

The CDS is also engaging with other groups at QUT to undertake external projects. These include:

- A recently finalised Data and Data Science Strategy for the Aboriginal and Torres Strait Islander Community Health Service, joint with e-Research
- A proposed collaboration with QCOSS, joint with the Centre for Justice.

6. Networks and Communications

Twitter followers have grown from 437 (as of 9 September 2020) to 788 followers (as of 9 August, 2021). News stories regarding the Centre are also posted on the CDS Website [here](#).

During the time since the last Management meeting, communications support from the Faculty and University has been substantially reduced. As noted above, two Project Officer positions have been initiated to help support networking across the Centre and promotion of CDS research. These positions are still pending.

The CDS is engaging in a collaborative activity, the Great Barrier Reef Citizen Science Challenge, with ACEMS for the upcoming National Science Week.

7. Equity and Diversity

The Centre remains committed to its equity, diversity and inclusion statement, which was presented and ratified at the last Management meeting.

We are restricted in monitoring our progress with respect to the stated objectives, since we are unable to obtain diversity (gender, ethnicity) data from HR. However, we aim to discuss with HR appropriate mechanisms for collecting this ourselves, with a view to providing the information in the Annual Report.

A Women in Data Science national conference is being planned for September 2021, with 50+ women researchers invited to participate. Participants are CIs and AIs on the proposed CoE in ‘Shaping the Future of Data Science’.

Our funding schemes (e.g., PhD top-ups, First Byte grants) explicitly seek and consider information about the potential for the student-supervisor-project combination to promote diversity and inclusion. Review panels are convened to ensure gender balance.

8. Key Performance Indicators (KPIs)

A status update on our year-to-date KPIs for each of our high-level goals is as below

High Level Goals	Operational Plan	KPI Measures (per annum unless otherwise stated)	YTD KPI (as of 10 Aug 2021)
To create an active Centre that is inclusive and representative of Data Science expertise across QUT	Recruit AIs across QUT	50 AIs	98 CIs and AIs
	Create DataScience@QUT Network	25 DataScience@QUT	23 DataScience@QUT established
	Develop a successful program of joint activities for the Centre	2 DataScience for HDR Networking session	HDR Academy established 2 Academy face to face networking sessions
		9 DataScience Under the Hood seminars	7 seminars held + 3 planned to end of 2021.
Develop equity and diversity policy in line with QUT and professional policies	% of female Centre Participants – 30%	TBC in Annual Report	
	Women in DS networking session	National conference planned for Sept 2021	
To become internationally respected and renowned for world-class, frontier research in data science	Develop a set of internationally leading core research projects aligned to Centre Research Programs	10 core projects	To be reported in the Annual Report
	Appoint CDS postdocs	6 postdocs	7 postdocs (6 funded by CDS, 1 funded by QUT Business School). Next round open now.
	Provide CDS PhD Scholarship top-ups	5 scholarship top-ups	5 scholarship top-ups. Next round open soon.
	Enlist other CDS researchers and students	CDS researchers and students to come from 5 Schools	CDS researchers and students come from 15 Schools (see table above)
		50 peer-reviewed journal articles by Centre Participants	55 peer-reviewed journal articles by Centre Participants
		5 joint publications between different groups at QUT	TBC Annual Report data
Award CDS research excellence prizes	5 CDS awards	5 CDS Awards. Next round open soon	

To amplify impact of CDS through external partnerships in research and translation	Fund and organise a Distinguished Visitors Seminar Program	5 Distinguished Visitors	Virtual Distinguished Visitors Program established
	Develop a Key Partners Program in collaboration with Office of Industry Engagement	4 Data Science @ Work key partners events	Held over due to covid19 restrictions; virtual event planned with Qld AI Hub
	Develop a program of Data Science Workshops aligned to CDS Domains	3 collaborative workshops	4 workshops in 2021
	Identify and publicise external funding opportunities	5 CAT1 joint research grant applications	4 CAT1 grant applications
To advance the Centre and QUT as a national leader in Data Science	Develop and lead an Australia Data Science Network (ASDN)	10 ASDN members	23 ASDN members
	Develop and promote collaborations with international cognate centres	3 international collaborators	TBC
	*Develop an ASDN website	ASDN website	Website established
	*Develop a social media program for CDS and ASDN	100 Twitter followers	788 @QUTDataScience Twitter followers @AustralianData Twitter for ASDN established
To train world class postgraduates, researchers and practitioners within the Centre research culture	Encourage and promote joint research supervision in collaboration with relevant Faculties/Schools	25 HDRs	71 HDRs
	Organise external networking events	4 industry events	1 industry event
To increase participation in Data Science research	Develop and implement a CDS undergraduate VRES program	5 undergraduates in VRES program	None offered in 2021. Centre to review this KPI.
To increase general capability in Data Science	Develop a CDS Professional Sabbatical Program	2 participants	1 participant in 2021 joint with ACEMS, from Dingo.
	Promote and support QUT and national internship programs for CDS ECRs and postgraduates	2 participants	None undertaken in 2021.

9. QUT Annual Reporting

The QUT Annual Reporting for Tier 1 Centres is due soon. Access to data remains a challenge with the new systems and faculty structures.

10. Budget

The 2021 Centre forecast originally indicated a large carryover from 2021 into 2022 (~\$900K) due to late start of staff, postdocs and industry fellows. This provided an opportunity to launch the inaugural First and Second Byte programs and will furthermore provide Domain funding, and support purchase of eResearch computing support later in 2021.