

- [CEO's Update](#)
- [Research Update](#)
- [Achievements](#)
- [Recent Publications](#)
- [Innovation Exchange](#)
- [PHD Profile](#)
- [Education & Training](#)
- [In the News](#)

CEO's Update

As we enter the last eight months of the D2D CRC, it is very pleasing that we are able to highlight some of the significant achievements the CRC team has made in the last quarter. Of particular note are the Australian and United States trials and operational deployments of the Apostle (Fivecast Insight) system, the user trials of the Investigation Management System (IMS), the completion of the Law and Policy program's 'Open Source Data and Information' project and several awards received by D2D CRC researchers and project teams.

It has also been great to see the increasing number of students completing their PhD studies in the last quarter with six D2D CRC students now having completed their studies and one student being awarded a doctorate (Dr Stanley Shanapinda).

Finally, this eNewsletter has further information on the Data Science Competency Framework (DSCF) and associated Development Planning Tool (DPT). We have revamped the tool based on feedback from initial user trials and the new DPT is an excellent resource for organisations looking to build and develop their data science teams. The DSCF and DPT are freely available to Commonwealth departments – if you are interested, I encourage you to contact me or our Education and Training Manager, Regine Richelle.

Sanjay Mazumdar
CEO

Research Update

The research and development teams have continued to focus on maturing capabilities and refining these through trials and demonstrations to end users. Some highlights include:

Beat the News

Engineering and Devops Team:

- Data science team completed initial development of a topic discovery framework for short text.

Disease prediction stream (U of A):

- Published a paper on characterising seasonal influenza in Australia in PLoS Computational Biology Journal.
- Published a report outlining an approximate Bayesian computation (ABC) method to forecast influenza in Australia.
- Developed prototype model for individual behaviour for application in web-based disease forecasting systems.

Social disruption prediction stream (U of A):

- Produced a dashboard demo product demonstrating Bayesian method on events in Australian cities during 2017-18.
- Conducted case study on echo chambers in social media discussion around the 2017 Australian same-sex marriage survey, and mapped discussion versus echo chambers in online social networks.
- Explored the use of network structures in related tweets for use in event prediction.

The proactive approach to potential social disruption stream (UniSA):

- Finished building two risk analysis models based on topic modelling and evolution tree respectively.
- PhD student, Jeff Anshah published a paper, titled SensorTree: Bursty Propagation Trees as Sensors for Protest Event Detection in the International Conference on Web Information Systems Engineering (WISE) 2018.
- PhD student, Sha Lu has submitted a paper, titled Effective Outlier Detection based on Bayesian Network and Proximity to 2018 Institute of Electrical and Electronics Engineers International Conference on Big Data (IEEE BigData).

Apostle

Engineering and Devops Team:

- Commenced new trials with Australian and United States national security agencies.
- Continued to support existing operational usage within Australian agencies.
- Prototyped additional ingestors, and new improved user experience.

Knowledge mining stream (UNSW):

- New method for event mining from text streams based on learning consistency rules, paper submitted to Association for the Advancement of Artificial Intelligence Conference 2019 (AAAI).
- New model for identifying issues of interest and stance (orientation towards issues) for politically motivated groups from social media text.
- Prototyped interface for displaying and ranking events, groups associated with events, and threat indicators.

Picturing knowledge stream (ANU):

- Presented work on linguistic style transfer of image captions at the 2018 Computer Vision Pattern Program Conference.
- Designed a visualization system for finding differences in graphs and submitted a demo paper.
- Designed a novel system for generating comparative summaries of document collections.

Semantic indexing stream (UTS):

- Submitted three papers to AAAI on topics of question answering, semantic role labelling, and embedding for semantic relation classification.

Integrated Law Enforcement

Engineering and Devops Team:

- Conducted user trials with two law enforcement agencies.
- Demonstrated federated search tool and installed search product at customer site.
- Implemented advance security controls in IMS for future trials.

Data curation foundry stream (UNSW):

- Designed, implemented and demonstrated features of a digital assistant for analysis of investigation data (offence types, evidence items, allegations and event types).
- Designed and implemented component that supports natural language query over investigation data.
- Improved event type recognition using vector space model.
- Completed the event entity extraction and recognition application programming interfaces.

Narrative visualisation stream (UniSA):

- Continued work with Defence Science and Technology to develop the Narrative Visualisation.
- Had three papers accepted for 4th International Symposium on Big Data Visual and Immersive Analytics, University of Konstanz, October 17-19 2018, Germany:
 1. *Immersive Visualisation of Geo-Temporal, Narratives in Law Enforcement*, Andrew Cunningham, J Walsh, B Thomas
 2. *Evaluating Navigation Techniques for 3D Graph, Visualizations in Virtual Reality*, A Drogemuller, A Cunningham, J Walsh, M Cordell, W Ross, B Thomas
 3. *Tangible Braille Plot: Tangibly Exploring Geo-Temporal Data in Virtual Reality*, J Walsh, A Cunningham, R Smith, B Thomas

Entity linking stream (UniSA):

- Designed and implemented a graph functional dependency-based entity linking method, and compared it with other state of the art methods. Results showed this method is advantageous in both effectiveness and interpretations.
- Compared Washington University's open IE4 implementation frameworks and ClauseIE, designed noise removal rules, and improved the efficiency and accuracy of entity property and relationship extractions from texts.

Law & Policy

- Dr Stanley Shanapinda, a D2D CRC doctoral scholarship holder, was awarded his PhD by UNSW titled *Advance Metadata Fair: The Retention and Disclosure of Location Information for Law Enforcement and Security and the Impact on Privacy*.
- A suite of five reports were delivered in the UNSW-led project entitled *Using 'Open Source' Data and Information for Defence, National Security and Law Enforcement* (project lead Associate Professor Lyria Bennett Moses, UNSW Law School). Comment has been invited from end-user agencies before the reports are finalised.
- A workshop was held in Brisbane between Data61 researchers, international researchers and the La Trobe research team to launch the project on Compliance by Design (CbD) and Compliance through Design (CD) Solutions to Support Automated Information Sharing (project lead Professor Pompeu Casanovas, La Trobe Law School).

Brenton Cooper
CTO

[Any questions? Please don't hesitate to contact Brenton Cooper](#)

Recent Publications

A full list of formal publications can be found at:

- [Apostle](#)
- [Beat the News](#)
- [Integrated Law Enforcement](#)
- [Law & Policy](#)

Notable Recent Publications Include:

Narrative and Spatial Memory for Jury Viewings in a Reconstructed Virtual Environment
C Reichherzer, A Cunningham, J Walsh, M Kohler, M Billingham, B Thomas.

Manipulating visibility of political and apolitical threads on Reddit via score boosting
M Carman, M Koerber, J Li, K Choo, H Ashman.

DataSynapse: A Social Data Curation Foundry
A Beheshti, B Benatalah, A Tabatabaee, H Motalaari-Nezhad, M Barukh, R Nouri.

Open Set Adversarial Examples
Z Zheng, L Zheng, Z Hu, Y Yang.

Analysing TV Audience Engagement with Twitter: Incremental Segment-Level Opinion Mining of Second Screen Tweets
G Katz, B Heap, W Wobcke, M Bain, S Kannangara.

SensorTree: Bursty Propagation Trees as Sensors for Protest Event Prediction
J Anshah, W Kang, L Liu, J Liu, J Li.

Attentive Sequence to Sequence Translation for Localizing Clusters of Interest by Natural Language Descriptions
K Ning, L Zhu, M Cai, Y Yang, D Xie, F Wu.

Generating Connected Random Graphs
C Gray, L Mitchell, M Roughan.

Machine Learning to Support Social Media Empowered Patients in Cancer Care and Cancer Treatment Options
D De Silva, W Ranasinghe, T Bandara, A Adkari, N Mills, L Idmalagoda, D Alahakoon, N Lawrentschuk, R Persad, E Osipov, R Gray, D Bolton.

Learning to Propagate Labels: Transductive Propagation Network for Few-Shot learning
Y Liu, J Lee, M Park, S Kim, E Yang, S Hwang, Y Yang.

Opinion Fraud Detection via Neural Autoencoder Decision Forest
M Dong, L Yao, X Wang, B Benatalah, C Huang, X Ning.

Data-Augmented Regression with Generative Convolutional Network
X Ning, L Yao, X Wang, B Benatalah, S Zhang, X Zhang

Innovation Exchange Update

The Big Data Connect Program has officially wrapped up now with great feedback from all participants involved (Promet Valve Australia, Frontier Microscopy, Coopers Brewery and SenteK). These projects were very diverse with unique applications of data science in each. The outcomes were extremely pleasing, so much so that SenteK have requested an additional project and other participants have also expressed interest in further engagement. Specifically, SenteK have extended their involvement with a further study to automatically estimate irrigation events and the impact on crop roots. D2D CRC is also working with Zileak on another soil-based project that aims to predict contamination levels from soil spectra. Both projects are progressing well and have the potential to extend further.

In the health analytics program, the hospital outcome project is progressing well with more papers submitted to high profile journals and several presentations made. The previous focus was on cardiovascular outcomes, but this is now being extended to all medical conditions and procedures. This could have far reaching impacts on patient safety moving forwards. D2D CRC has received interest about this project from several organisations, including SA Health, and are now looking to deploy at operational deployments of the outcomes.

The 'Actionable In-time Insights' (AI2) project is also progressing with trials monitoring mental health patients and their medication/appointment adherence. This work is also being extended to an application that can assist clinicians in ensuring patient health and well-being.

PhD Profile

Minjeong Shin

Australian National University

Minjeong Shin is a PhD candidate exploring the visualization of big data by constructing knowledge graphs and generating visual event summaries. In addition to her current research, Minjeong holds a Masters in Computer Science and a Bachelors of Computer Science from KAIST, South Korea.

Minjeong worked as a research engineer for LG in South Korea before coming to Australia in 2015. Since 2016 she has worked on the D2D CRC Apostle project as a research engineer at Australian National University before deciding to commence her PhD.

She was awarded the inaugural D2D CRC Women in Data Science award for creating a positive contribution to the data science field and the role of women within it. Minjeong also delivered a software prototype to D2D CRC in her first six months.

Education and Training Update

Current Workforce

Big Data Leaders Series

The next workshops are currently being planned, with topics including the challenges of building new architectures with legacy systems planned.



Short courses and seminars

In the last three months, D2D CRC organised several seminars in Adelaide and two in-house seminars for Australian Taxation Office staff. The seminars were well received by the audience and attendance continues to be strong. The focus was to involve D2D CRC PhD students in each of them.

The first seminar was on *Big Data Analytics in Health*. This seminar focussed on D2D CRC health analytics projects and their potential to improve the quality of healthcare in Australia. It contained two presentations, one by a D2D CRC staff member (Dennis Horton) and one by a PhD student (Dennis Liu).

D2D CRC organised a seminar in August on *Using Deep Reinforcement Learning to Tackle Complex Problems*, presented by Alex Long (D2D CRC PhD student) as well as the *Responsible Application of Data Analytics*, presented by Professor Roger Clarke.

In September, Dr Stanley Shanapinda, the first D2D CRC student to receive their doctorate, presented the results of his research project, *Advanced Metadata Fair - Big Location Data, Privacy and Law Enforcement*. His presentation summarised the strengths and opportunities of the regulatory framework and analysed the powers it grants law enforcement agencies to collect and use Big Location Data vis-a-vis the privacy tests introduced in 2017, in relation to court cases and the functionality of data analysis.

Future Workforce

PhD and Honours Scholarship Update

One new student started a D2D CRC PhD scholarship.

Two more students submitted their thesis (a total of six students have now submitted), and one student, Dr Stanley Shanapinda was awarded his doctorate.

On October 18 D2D CRC organised a meet and greet event in Canberra with 14 D2D CRC PhD students and five national security agencies. The students showcased their research and networked with various agency representatives. The success was guaranteed by great engagement from partners and exciting perspectives in the sector for our students.

Planning is underway to develop a PhD personal development day at the D2D CRC annual conference in March 2019 as well as a celebration event for D2D CRC students completing their PhDs.

D2D CRC is organising a public speaking workshop for PhD students to be held in November, tailored to their needs.

Internships

The last two summer interns have been recruited, who will start a ten week internship on 1 December 2018.

Data Science Competency Framework (DSCF) and Development Planning Tool (DPT)

Interest in the DSCF and DPT continues to grow, including discussions with several organisations looking to trial the framework and tool.

Feedback from the pilot undertaken in 2017 has been implemented to further enhance the framework and tool. Both are now ready to be used by the public sector and can be freely trialled by Commonwealth agencies.

The beta version of the DPT is being trialled by a government department in Canberra.

For more information on the Data Science Competency Framework or for a trial of the framework and Development Planning Tool, please contact regine.richelle@d2dcr.com.au.

Football and Science Combine — helping kids to increase kids participation in STEM.

[Read More](#)



Funding boost for South Australian AI Business

[Read More](#)

