MACQUARIE University

Strategic Approaches for Technologically enabled Research and Data Science

Brian Ballsun-Stanton

12 August 2020 • Faculty of Arts

Introduction

Dr Brian Ballsun-Stanton – Solutions Architect (Digital Humanities)

- PhD UNSW 2012 in the Philosophy of Data.
- BS and MS at Rochester Institute of Technology in Information Technology.
- $2,262,449 in Grants and Prizes
- 23 Projects across the faculty since 2017 including Philosophy, Security Studies, and Ancient History.

(a) 63 field data collection projects

(b) onwork.edu.au, indexing 2000 citations

(c) MQ's Partnership with Google Arts & Culture. 15,820 views in July 2020

Outline of Talk

MACQUARIE University

Technologically Enabled Research

Illustrative Examples

Macquarie's Strategic Transformation

Issues for consideration

MACQUARIE University

Technologically Enabled Research

MACQUARIE University

Technology:

> ... technologies either as ***objects*** *(the stethoscope, the rifle),* ***practices*** *(disciplinary techniques),* ***knowledge*** *(medicine and penology), modes of* ***organization*** *(hospital, school, prison), or frequently, by the* ***conglomeration*** *of these. (Lagdameo 2019)*

Data:

> *Data as objective* ***measure****, Data as subjective* ***observation****, Data as encoded human* ***communication*** *(Ballsun-Stanton 2012)*

System:

> *According to the cybernetician,* ***the purpose of a system is what it does.*** *This is a basic dictum. It stands for a bald fact, which makes a better starting point in seeking understanding than the familiar attributions of good intentions... (Beer 2002)*

Ethical Considerations

Reduce the number of redundant studies because we don't have the data

- 'As open as possible, as closed as necessary.' *(European Commission 2016)*
- The replication crisis and our duties to future scholarship *(Hochstrasser 2020; National Academies of Sciences, Engineering, and Medicine et al. 2019; Reed 2014; Franco, Malhotra and Simonovits 2014)*

Reduce harm to participants, researchers, institutions

- Automation
- NHMRC Compliance
- Reduce impact of breaches

Figure 2: QLD Health breach. *(Bavas 2019)*

Figure 3: Consequences of drives left in a safe. *(Davis 2020)*

Optimising for human attention

Humans to figure out what – computers to do over and over again.

- Data should only need to be entered once!
- Graphs should automatically generate on changes to data!
- Submitting papers to a different journal should not take a week of re-typesetting.

FAIR and Reproducible Research

- *Excel delinda est!* Let us code our analyses in a testable way! *(Krugman 2013; Ross and Ballsun-Stanton 2020; Bruford et al. 2020)*
- Directly and specifically reward data reuse.
- Promote reproducing studies as a valid MRes option. *(Spring et al. 2018)*

Optimising for novel capabilities

MACQUARIE University

New practices + new knowledge + new tools

- Human + Computer can extend capabilities of both. *(Vinge 1993)*
- Technology may make novel research possible – with an awful lot of blind alleys first.

Solving the allocation problem is hard

- We need to reward research which might fail.
- We need to make failure as cheap, fast, and easy as possible. Reputationally expensive to fail *after* winning the ARC grant.

We increase what we reward – but are the system's rewards aligned to what we say we want? *(Beer 2002)*

MACQUARIE University

Illustrative Examples

Impact: Social Media Research for NSW government

MACQUARIE University

Useful advocacy to solve specific research problems

- Automated collection: 1.3 million social media posts.
- Python's Natural Language Processing to collect, process, describe, and visualise data.

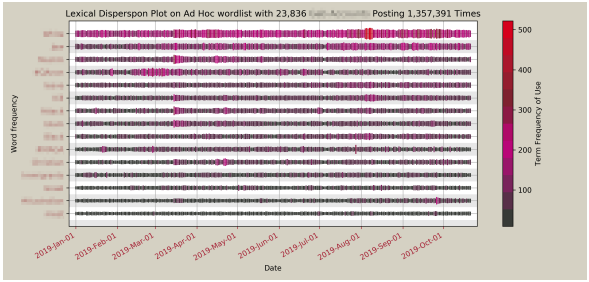

Figure 4: Blurred Lexical Dispersion plot from this research

High impact

- Policy recommendations backed by thorough and reproducible data.
- Presented to Commonwealth committee this week, state committees in July.

Engagement: Google Arts and Culture and Fabricius

- *artsandculture.google.com/partner/macquarie-university*
  - Google Arts and Culture:
    'is to preserve and bring the world's art and culture online so it's accessible to anyone, anywhere.' (Google Arts and Culture 2020)
  - 15,820 views from 142 countries on MQ's Arts and Culture pages in July.
- *g.co/fabricius*
  - Google Cloud Machine Learning + Ancient Egyptian Hieroglyphs.
  - 230 press stories around the world.
  - Approving tweet to us from Australian ambassador to Egypt.

MACQUARIE University

Macquarie's Strategic Transformation

Focused investment in Research result: 'Improved access to, and quality of, shared research facilities and infrastructure with a focus on providing institution-level facilities'

- Deploying technology at Macquarie is difficult –
- Persuading joyful adoption is harder.

Digital Transformation measure: 'Adoption of new technology and ways of working.'

- Being a *useful* advocate to solve *specific problems*.
- Demonstrate that new systems promise higher reward relative to increased risk.

Focused investment in Research result: ‘Improved upon the results of ERA and EI 2018’

- Reduce exposure to data breaches, scandals, and disruptions to research continuity through data-loss.
- Increase technical options for high engagement and opportunities for impact.
- Improve paper writing and collaboration technologies.

Ways of Working measure: ‘Staff engagement and retention, specifically in career development’

- Use best methods in data skills training *and community building* from around the world.
- Showing faculty and staff how to solve *small* problems with code creates a community and wins that translate into career development.
- Guide researchers into considering what tools to learn and use, rather than reaching for comfortable, slow, default.

MACQUARIE University

Issues for consideration

Many strategic issues

MACQUARIE University

Valuing researcher/staff time, risk-taking, and building general capability

- How do we create space for staff to engage with novel techniques?
- How do we treat researcher and staff time-value versus money?
- How do we get our system to reward general skill/capability building?

Pay attention to university feedback and reward loops

- What, specifically, do we reward – and does it support better data practices?
- What are our feedback loops? How long to learn about failure?

Caution via ‘Goodhart’s Law’: ’When a measure becomes a target, it ceases to be a good measure.‘ *(Strathern 1997; Goodhart 1975)*

Thank you!

A citable resource containing LaTeX source code and this compiled presentation is available at: *https://osf.io/y5c7n/*

References 1

Ballsun-Stanton, Brian. 2012. ‘Asking About Data: Exploring Different Realities of Data Via the Social Data Flow Network Methodology’. PhD diss., University of New South Wales.

Bavas, Josh. 2019. ‘Queensland Health launches investigation after medical files found on busy Brisbane road’. *ABC News* (June). *https://www.abc.net.au/news/2019-06-26/medical-files-dropped-on-busy-brisbane-road/11249460*.

Beer, Stafford. 2002. ‘What is cybernetics?’ *Kybernetes. The International Journal of Cybernetics, Systems and Management Sciences* 31, no. 2 (January): 209–219. ISSN: 0368-492X. doi:*10.1108/03684920210417283*.

Bruford, Elspeth A, Bryony Braschi, Paul Denny, Tamsin E M Jones, Ruth L Seal and Susan Tweedie. 2020. ‘Guidelines for human gene nomenclature’ [in en]. *Nature genetics* 52, no. 8 (August): 754–758. ISSN: 1061-4036, 1546-1718. doi:*10.1038/s41588-020-0669-3*.

References 2

Davis, Jessica. 2020. *UW Medicine Hit with Lawsuit for Breach Impacting 974K Patients.* Accessed: 2020-8-9, February. *https://healthitsecurity.com/news/uw-medicine-hit-with-lawsuit-for-breach-impacting-974k-patients*.

European Commission. 2016. *Open Access*. *https://ec.europa.eu/research/participants/docs/h2020-funding-guide/cross-cutting-issues/open-access-data-management/open-access_en.htm*. Accessed: 2020-8-8, July.

Franco, Annie, Neil Malhotra and Gabor Simonovits. 2014. 'Social science. Publication bias in the social sciences: unlocking the file drawer' [in en]. *Science* 345, no. 6203 (September): 1502–1505. ISSN: 0036-8075, 1095-9203. doi:*10.1126/science.1255484*.

Goodhart, C. 1975. 'Problems of monetary management: the UK experience in papers in monetary economics'. *Monetary Economics* 1.

References 3

Google Arts and Culture. 2020. *About Google Cultural Institute.* Accessed: 2020-8-8. *https://about.artsandculture.google.com/*.

Hochstrasser, Judith. 2020. *Non-reproducible because of missing data.* Accessed: 2020-6-18, June. *https://www.horizons-mag.ch/2020/06/04/non-reproducible-because-of-missing-data/*.

Krugman, Paul. 2013. 'The excel depression'. *The New York times* 18. ISSN: 0362-4331.

Lagdameo, Federico José T. 2019. 'Heidegger, Foucault, And An Affordance Theory Of Technology'. PhD diss., Ateneo De Manila University.

References 4

National Academies of Sciences, Engineering, and Medicine, Policy and Global Affairs, Committee on Science, Engineering, Medicine, and Public Policy, Board on Research Data and Information, Division on Engineering and Physical Sciences, Committee on Applied and Theoretical Statistics, Board on Mathematical Sciences and Analytics et al. 2019. *Reproducibility and Replicability in Science* [in en]. National Academies Press, October. ISBN: 9780309486163. *https://www.nap.edu/read/25303/*.

Reed, Richard. 2014. *Again, again, and again – our first (of many) replication issues.* doi:*10.1108/ajb-01-2014-0008*.

Ross, Shawn A, and Brian Ballsun-Stanton. 2020. 'Introducing Preregistration of Research Design to Archaeology'. In *Digital Heritage and Archaeology in Practice,* edited by Ethan Watrall and Lynne Goldstein. Gainesville, FL: University Press of Florida.

References 5

Spring, Jonathan, Mary Maller, Enrico Mariconti and Jeremiah Onaolapo. 2018. *An untapped resource to reproduce studies.* Accessed: 2018-11-27, July. *https://www.benthamsgaze.org/2018/07/31/an-untapped-resource-to-reproduce-studies/*.

Strathern, Marilyn. 1997. ‘‘Improving ratings’: audit in the British University system’. *European review* 5, no. 3 (July): 305–321. ISSN: 1062-7987, 1474-0575. doi:*10.1002/(SICI)1234-981X(199707)5:3<305::AID-EURO184>3.0.CO;2-4*.

Vinge, Vernor. 1993. ‘Technological Singularity’. *Whole Earth Review.* *https://frc.ri.cmu.edu/~hpm/book98/com.ch1/vinge.singularity.html*.