



MACQUARIE
University
SYDNEY · AUSTRALIA

Seven years of FAIMS Mobile

Shawn A Ross

Office of the Deputy Vice-Chancellor (Research)

June 5, 2019





Transparency and reproducibility

‘Small data’ infrastructure across the data lifecycle

Lessons from FAIMS: summary

Lessons from FAIMS: in-depth

From current practice to better practice

References



MACQUARIE
University
SYDNEY • AUSTRALIA

Transparency and reproducibility

For nearly a decade the reproducibility crisis has featured in the scientific literature [Jasny et al., 2011, Baker, 2016, Munafò et al., 2017]. Low reproducibility rates have emerged from large-scale studies:

- Results from only 39% of psychology studies could be reproduced [Open Science Collaboration, 2015].
- Even lower reproducibility rate in biomedical research [Begley and Ellis, 2012, Prinz et al., 2011].

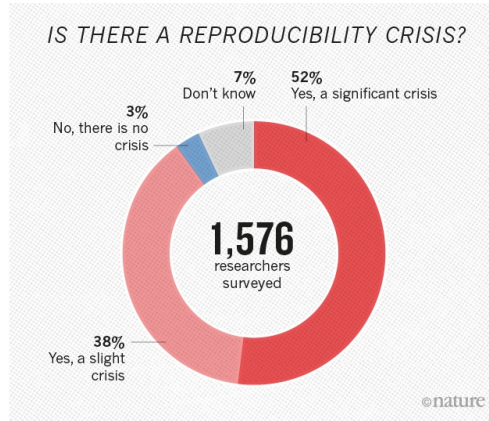


Figure 1: Is there a reproducibility crisis? [Baker, 2016]

MISSING DATA

As research articles age, the odds of their raw data being extant drop dramatically.

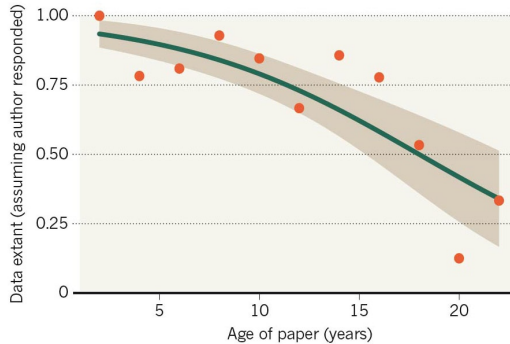


Figure 2: [Vines et al., 2014]



Key guidelines to good practice:

- Findable, Accessible, Interoperable, and Reusable (FAIR) data [Wilkinson et al., 2016, GO-FAIR, 2017].
- Transparency and Openness Promotion (TOP) guidelines [Nosek et al., 2015].
- Data transparency toolkit [Perkel, 2018].

Mandates for transparency or reproducibility:

- Nature: Transparency Upgrade [Nature, 2017].
- Nature: FAIR data in Earth science [Nature, 2019].
- Copernicus: FAIR data in atmospheric sciences [van Edig, 2018].
- Not just the natural sciences: AJPS requires data and code [Jacoby et al., 2017, AJPS, 2015].
- TOP Guidelines signatories include publishers representing 1000+ journals, as well as professional organisations and major private foundations that fund research [COS, 2019].
- Funders' data policies [DCC, 2019]



2. Authors using original data must:
 - a. make the data available at a trusted digital repository [...]
 - b. include all variables, treatment conditions, and observations described in the manuscript.
 - c. provide a full account of the procedures used to collect, preprocess, clean, or generate the data.
 - d. provide program code, scripts, codebooks, and other documentation sufficient to precisely reproduce all published results.
 - e. provide research materials and description of procedures necessary to conduct an independent replication of the research.

COMPARISON OF FOUR PUBLISHER DATA POLICIES TO THE TOP GUIDELINES

	NOT TOP COMPLIANT Encourages sharing	TOP LEVEL 1 Disclose	TOP LEVEL 2 Require	TOP LEVEL 3 Verify
ELSEVIER	Policy A Policy B	Policy C	Policy D Policy E*	
SPRINGER NATURE	Policy 1 Policy 2	Policy 3	Policy 4*	
TAYLOR & FRANCIS	Basic	Share upon reasonable request**	Publicly available Open data Open and fully FAIR	
WILEY	Encourages Data Sharing	Expects Data Sharing	Mandates Data Sharing	Mandates data sharing and peer reviews data***
MORE JOURNALS IMPLEMENTING POLICIES	Any journal that merely encourages data sharing	- Psychonomics Society Journals - Nature - Psychological Science - PNAS	- Science - PLOS - Royal Society Journals - Cognition	- AJPS - Biostatistics - JEPS - JPR - Meta-Psychology - QJPS

Figure 3: The Landscape of Open Data Policies [Mellor, 2018]

Private funders have endorsed via the Open Funders Research Group:

- Alfred P. Sloan Foundation
- American Heart Association
- Bill and Melinda Gates Foundation
- Howard Hughes medical Institute
- John Templeton Foundation
- Laura and John Arnold Foundation
- Open Society Foundations
- Robert Wood Johnson Foundation
- Wellcome Trust
- and six more [OFRG, 2019]

Key:

● Full Coverage ◐ Partial Coverage ○ No Coverage

Research Funders	Policy Coverage		Policy Stipulations					Support Provided			
	Published outputs	Data	Time limits	Data plan	Sharing/ access	Long-term curation	Monitoring	Guidance	Repository	Data centre	Costs
AHRC	●	●	●	●	●	◐	◐	●	○	◐	●
BBSRC	●	●	●	●	●	●	●	●	●	◐	●
EPSRC	●	●	●	◐	●	●	●	◐	○	○	●
ESRC	●	●	●	●	●	●	●	●	●	●	◐
MRC	●	●	●	●	●	●	○	◐	●	○	◐
NERC	●	●	●	●	●	●	●	●	●	●	◐
STFC	●	●	●	●	●	●	●	◐	●	◐	◐
Cancer Research	●	●	●	●	●	●	●	◐	●	○	●
European Commission	●	●	◐	●	◐	◐	◐	●	●	◐	●
Wellcome Trust	●	●	●	●	●	●	●	●	●	●	●

Figure 4: Overview of funders' data policies [DCC, 2019]

The NHMRC 'strongly encourages' data sharing in the National Statement on Ethical Conduct in Human Research and their Open Access Policy.
[NHMRC, 2018a, NHMRC, 2018b]

National Statement 3.1.50

In the absence of justifiable ethical reasons (such as respect for cultural ownership or unmanageable risks to the privacy of research participants) and to promote access to the benefits of research, researchers should collect and store data or information generated by research projects in such a way that they can be used in future research projects. Where a researcher believes there are valid reasons for not making data or information accessible, this must be justified.

Key changes to the Open Access Policy (15 January 2018)

Research data and metadata (2.2)

NHMRC now strongly encourages researchers to take reasonable steps to share research data and associated metadata arising from NHMRC supported research.

FAIR principles (2.7)

Reference to the Australian F.A.I.R. principles (Findable, Accessible, Interoperable, Reusable) when publishing research literature and sharing data has been made.

Metadata (4.1)

The metadata for the peer-reviewed publication must be made openly accessible via an institutional repository as soon as possible but no later than 3 months from the date of publication.

Data (4.2)

NHMRC acknowledges the importance of making research data publicly accessible and therefore strongly encourages researchers to consider the reuse value of their data and to take reasonable steps to share research data and associated metadata arising from NHMRC supported research.

- NSW General Retention and Disposal Authority GDA 23; data associated with 'significant' research or researchers must be kept forever (23.6.1) [NSW, 2015]
- NSW Privacy and Personal Information Protection Act 1998 No 133, esp. Part 2, Division 1, Section 19, which flags indicators of high sensitivity and establishes data sovereignty.[NSW, 1998] Compare the (Australian) Privacy Act 1988, esp. Part II, Division 1, Section 6 'Sensitive Information' and Schedule 1, and 'Australian Privacy Principles', Section 8, which covers some university-controlled entities. [AG, 2017, OAIC, 2019a]
- NSW Notifiable Data Breach guidance [IPC NSW, 2018]; see also the Australian Notifiable Data Breaches scheme [OAIC, 2019b]
- EU General Data Protection Regulation [GDPR, 2019]

What does this mean? Are we ready?



Emerging good practice - and publisher and funder policies - mean:

- Comprehensive, FAIR datasets will be deposited in domain-specific repositories. Data, and especially metadata, quality will be higher.
- Data will be captured digitally as early in research as possible, and provenance / version history maintained.
- Research approach, processes, and procedures will be documented.
- Data processing and analysis will use code (not Excel or ARCGIS!)
- Code will be documented and published for reuse.
- Further steps taken for analytical reproducibility (use of OSS, version control, automation, containerisation, etc.).

The same approaches that facilitate transparency and reproducibility support the kind of scalable and synthetic research that can address archaeological 'grand challenges'. [Kintigh et al., 2014]

- Paper data capture and manual digitisation and cleaning don't scale.
- Email and desktop software don't scale.

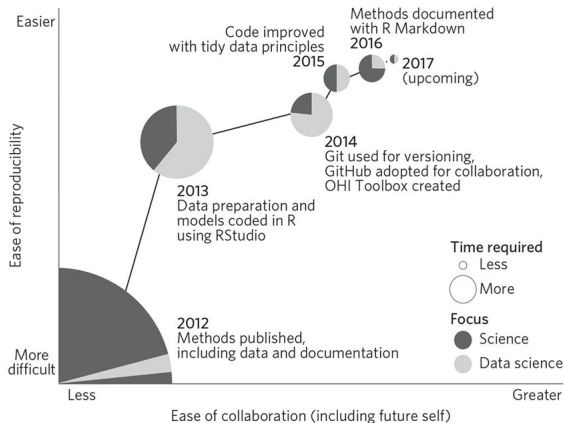


Figure 5: Better science in less time, illustrated by the Ocean Health Index project. [Stewart Lowndes et al., 2017]



MACQUARIE
University
SYDNEY • AUSTRALIA

‘Small data’ infrastructure across the data lifecycle



Figure 6: Archaeologists contemplate data standards (FAIMS Stocktaking, 2012)

'Long tail' research: most field data is small data [Borgman, 2015]

- Smaller scale; smaller communities; local control.
- Diverse questions, approaches, and methods.
- Heterogeneous data; variety of content, structure.
- Data and infrastructure emerge from fieldwork.
- Relative lack of standards.
- Limited infrastructure and funding.
- Challenges associated with big(ger) data from photogrammetry, SfM, video, geophysics, etc., will exacerbate these problems.

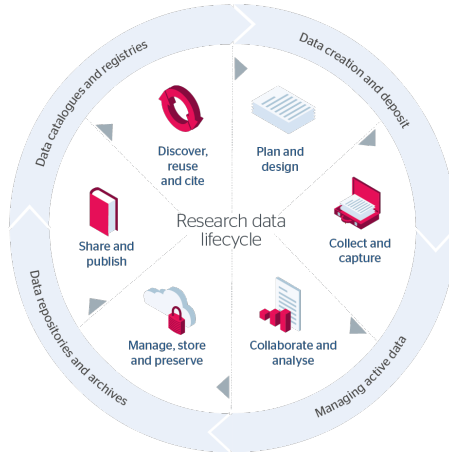


Figure 7: [JISC, 2018] Image CC-BY-ND

Consider the infrastructure needed to manage the three main phases of the data lifecycle

- Publication (most mature): domain-specific repositories.
- Processing and analysis (less mature): project-level code [Stewart Lowndes et al., 2017], then Virtual Labs / Science Gateways, like [Alveo, 2019] in language analysis.
- Capture (least mature): most varied, needs to work offline under difficult conditions. Commercial solutions insufficient [Bureau of Reclamation, 2017].



MACQUARIE
University
SYDNEY • AUSTRALIA

Lessons from FAIMS: summary

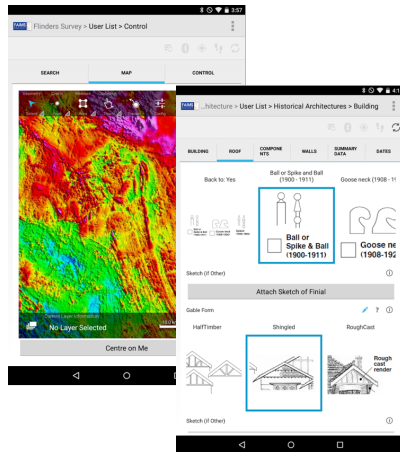


Figure 8: FAIMS Mobile: GIS and 'picture dictionaries'

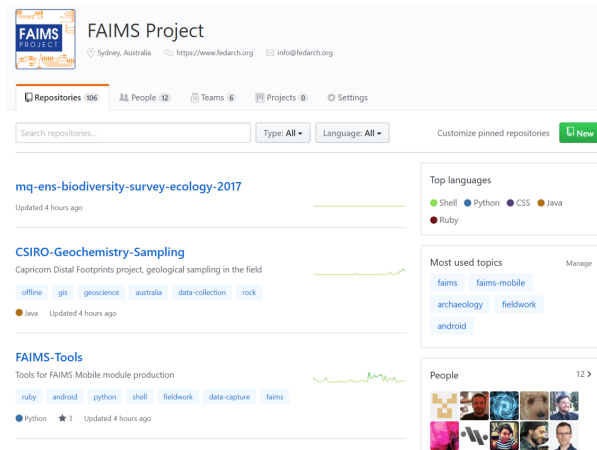


Figure 9: FAIMS Mobile customisations on GitHub

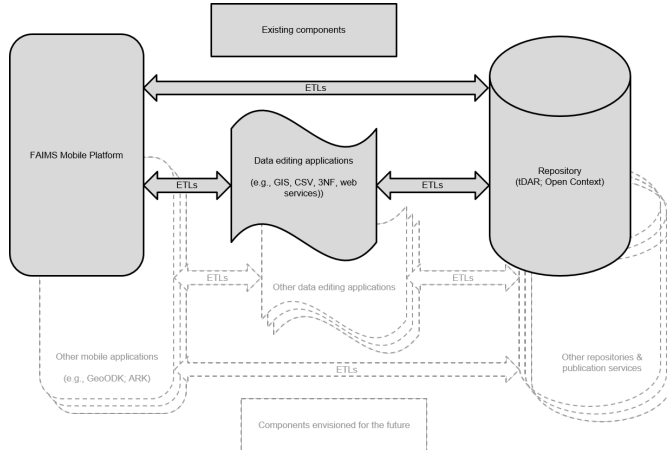


Figure 10: FAIMS Mobile federation

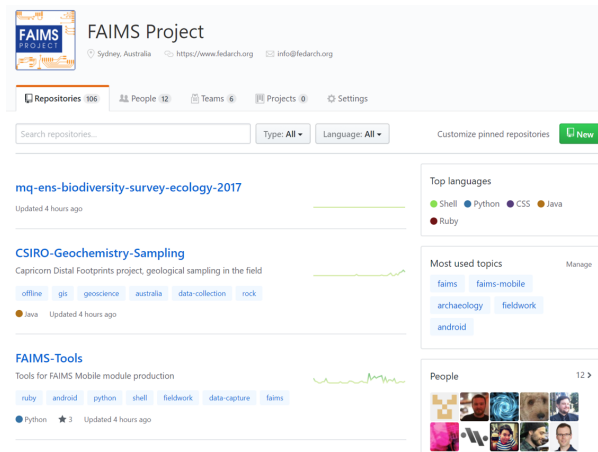


Figure 11: FAIMS Mobile 'core' code is GPLv3



- We deserve research-specific software.
- Diverse practices and limited resources require generalised software.
- Do one thing well with modular and federated software (but slice the pie thoughtfully).
- Open-source software has advantages (but is difficult to sustain).
- Scope requirements carefully.
- Invest in outreach and engagement.

How do we get from where we are now to where we want to be?

- Understand the evolving expectations of transparent research.
 - Look past desktop software (Excel, ARCGIS, Filemaker, Access, etc.).
 - Rally around emerging research- and domain-specific solutions (even if imperfect).
 - Overcome ‘not invented here’; you don’t need a bespoke solution.
 - Budget for ‘ground-up’ transparency (data and code). Up-front costs will be high but offer longer-term payoffs (in costs, time, and quality).
 - Implement (and budget for) fundamental good practice in data and code management before other technologies.
 - Improve research design (prioritise approach over methods)
- [Muthukrishna and Henrich, 2019, Hole, 1973]



MACQUARIE
University
SYDNEY • AUSTRALIA

Lessons from FAIMS: in-depth

- The Field Acquired Information Management Systems (FAIMS) Project began in 2012 as a national Australian information infrastructure project in archaeology.
- Developed FAIMS Mobile for field data capture [Ballsun-Stanton et al., 2018].
- Use expanded beyond archaeology to geoscience, ecology, ethnography, linguistics, oral history.
- Has been customised for over 50 workflows at more than 30 projects.
- Data and workflow modelling for these customisations provided deep insights into field data capture and the infrastructure needed to support it.

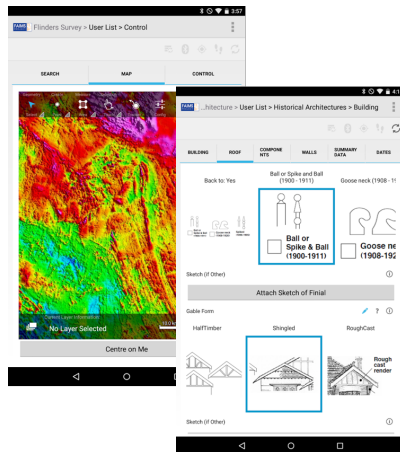


Figure 12: FAIMS Mobile: GIS and 'picture dictionaries'

- Fundamentally customisable.
- Tightly binds structured, geospatial, multimedia, and free text data.
- Works offline.
- Automated bi-directional synchronisation using local or online server
- Record history: append-only datastore, versioning, rollback.
- Mobile GIS.
- Connects to internal and external sensors, Bluetooth / USB devices.
- Multilingual.
- Granular help.
- Granular metadata / uncertainty.
- Generalised export.
- 'Hooks' for data interoperability, Open Linked Data approaches.



- We deserve research-specific software.
- Diverse practices and limited resources require generalised software.
- Do one thing well with modular and federated software (but slice the pie thoughtfully).
- Open-source software has advantages (but is difficult to sustain).
- Scope requirements carefully.
- Invest in outreach and engagement.

Archaeology needs (and deserves) research-specific software, contra [Roosevelt et al., 2015].

- Most commercial / mass-market software does not meet research needs.
- Risk of lock-in, unwelcome changes to features or business models, and product discontinuation.

Compare ecology in Australia: TERN, ALA, Biocollect, and associated research clouds [TERN, 2019, ALA, 2019a, ALA, 2019b].

Commercial software doesn't meet our needs, and bespoke development is too expensive and usually unsustainable.

- Generalised software can be deeply customised to accommodate our diverse data types, data models, workflows, etc.
- The code used to customise it describes the data model and workflow.
- Customisations can be published and re-deployed trivially.
- Can deliver research-grade software affordably.

FAIMS Mobile cost perhaps 3x a single bespoke application, but has been customised 50x. Customisation cost is 1/10th bespoke, and still $<1/2$ even if 'core' platform development costs are amortised across projects.

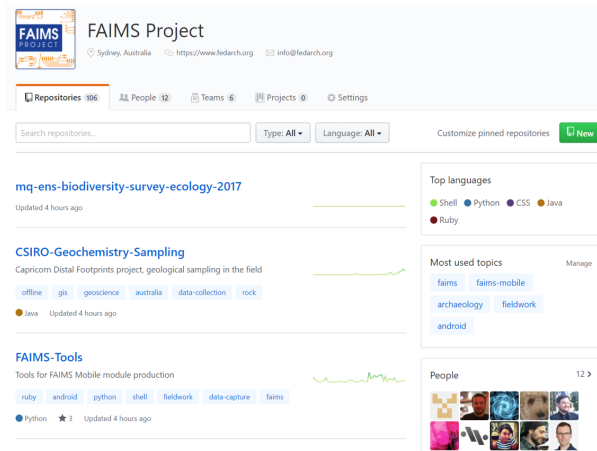


Figure 13: FAIMS Mobile customisations (XML files, mostly) on GitHub

Do one thing well.

- Identify other infrastructure in the domain and interoperate with it (via ETLs or APIs).
- It is better to divide by data-lifecycle phase rather than data type, since (1) our data is so integrated and (2) field data capture poses unique challenges.

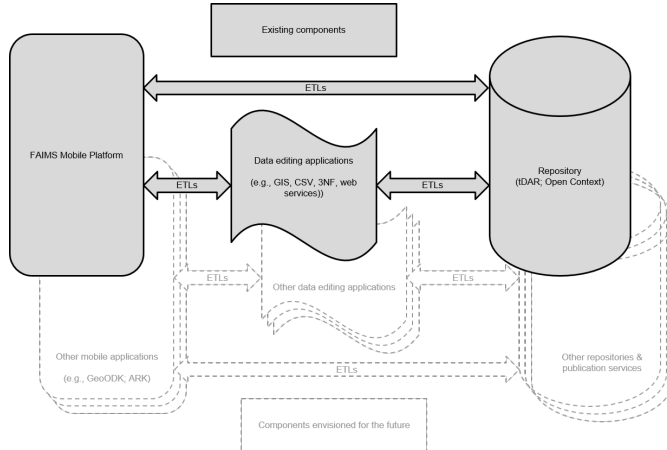


Figure 14: FAIMS Mobile federation strategy

Open source has advantages but is difficult to sustain.

- Emerging open research principles strongly prefer OSS as opposed to proprietary 'black boxes'.
- Transparency and reusability (esp. customisation code).
- Ability to hand off from one organisation to another (esp. 'core' platform code).
- Ability to fork code prevents lock-in and mitigates unwelcome decisions by software developers.
- BUT OSS business models are hard to scale and rely on occasional injections of grant or institutional funding.













Talk to a wide range of potential users, seeking facts not opinions.

- Don't ask researchers what they think, ask them what they have done
- what software they have adopted and why, and what problems they have expended resources to solve.
- 'Lean startup' methodology very useful, based around testing of ideas through interviews with potential users [Strategyzer AG, 2019].
- In our case, we over-invested in mobile GIS and under-invested in usability (especially a GUI for customisation).

If you build it they will not come; people can't use technologies they don't know about.

- As per industry standards, dedicate at least 30% of any information infrastructure budget to outreach and engagement (sales and marketing).
- Typical academic outreach (journal articles, conference presentations, workshops, even booths at major conferences) are not enough.

- [Sobotkova, 2018]
- [Ballsun-Stanton et al., 2018]
- [VanValkenburgh et al., 2018]
- [Sobotkova et al., 2016]
- [Ross et al., 2015]
- [Sobotkova et al., 2015]
- [Ross et al., 2013]

-  AG (2017).
Privacy act 1988.
Technical Report 119, Attorney-General's Department.
-  AJPS (2015).
AJPS replication and verification policy.
<https://ajps.org/ajps-replication-policy/>.
Accessed: 2019-4-15.
-  ALA (2019a).
Atlas of living australia.
<https://www.ala.org.au/>.
Accessed: 2019-4-24.
-  ALA (2019b).
BioCollect: Advanced data collection for biodiversity projects.
<https://www.ala.org.au/biocollect/>.
Accessed: 2019-4-24.
-  Alveo (2019).
Alveo: A virtual lab for human communication science.
<http://alveo.edu.au/>.
Accessed: 2019-4-24.
-  Baker, M. (2016).
1,500 scientists lift the lid on reproducibility.
Nature, 533(7604):452–454.
-  Ballsun-Stanton, B., Ross, S. A., Sobotkova, A., and Crook, P. (2018).
FAIMS mobile: Flexible, open-source software for field research.
SoftwareX, 7C:47–52.
-  Begley, C. G. and Ellis, L. M. (2012).
Drug development: Raise standards for preclinical cancer research.
Nature, 483(7391):531–533.
-  Borgman, C. L. (2015).
Big data, little data, no data: scholarship in the networked world.
MIT press.
-  Bureau of Reclamation (2017).
DataApp: A mobile app framework for field data capture.
<https://www.usbr.gov/research/challenges/dataapp.html>.
Accessed: 2018-3-27.
-  COS (2019).
TOP guidelines.
<https://cos.io/our-services/top-guidelines/>.
Accessed: 2019-4-25.
-  DCC (2019).
Overview of funders' data policies | digital curation centre.

<http://www.dcc.ac.uk/resources/policy-and-legal/overview-funders-data-policies>.

Accessed: 2019-3-29.



GDPR (2019).

EUGDPR - information portal.

<https://eugdpr.org/>.

Accessed: 2019-6-3.



GO-FAIR (2017).

FAIR principles - GO FAIR.

<https://www.go-fair.org/fair-principles/>.

Accessed: 2019-3-29.



Hole, F. (1973).

Questions of theory in the explanation of culture change in prehistory.

In Renfrew, C., editor, *The Explanation of Culture Change: Models in Prehistory*, pages 19–34. Duckworth.



IPC NSW (2018).

IPC NSW data breach guidelines.

Technical report, Information and Privacy Commission NSW.



Jacoby, W. G., Lafferty-Hess, S., and Christian, T.-M. (2017).

Should journals be responsible for reproducibility?

<https://www.insidehighered.com/blogs/rethinking-research/should-journals-be-responsible-reproducibility>.

Accessed: 2019-4-15.



Jasny, B. R., Chin, G., Chong, L., and Vignieri, S. (2011).

Data replication & reproducibility. again, and again, and again introduction.

Science, 334(6060):1225.



JISC (2018).

Research data management toolkit.

<https://rdmtoolkit.jisc.ac.uk/research-data-lifecycle/>.

Accessed: 2018-11-6.



Kintigh, K. W., Altschul, J. H., Beaudry, M. C., Drennan, R. D., Kinzig, A. P., Kohler, T. A., Limp, W. F., Maschner, H. D. G., Michener, W. K., Pauketat, T. R., Peregrine, P., Sabloff, J. A., Wilkinson, T. J., Wright, H. T., and Zeder, M. A. (2014).

Grand challenges for archaeology.

Proceedings of the National Academy of Sciences of the United States of America, 111(3):879–880.



Mellor, D. (2018).

The landscape of open data policies.

<https://cos.io/blog/landscape-open-data-policies/>.

Accessed: 2019-4-15.



Munafò, M. R., Nosek, B. A., Bishop, D. V. M., Button, K. S., Chambers, C. D., du Sert, N. P., Simonsohn, U., Wagenmakers, E.-J., Ware, J. J., and Ioannidis, J. P. A. (2017).

A manifesto for reproducible science.

Nature Human Behaviour, 1(1):s41562–016–0021.



Muthukrishna, M. and Henrich, J. (2019).

A problem in theory.

Nature human behaviour, 3(3):221–229.



Nature (2017).

Announcement: Transparency upgrade for nature journals.

Nature, 543(7645):288.



Nature (2019).

Announcement: FAIR data in earth science.

Nature, 565(7738):134.



NHMRC (2018a).

National statement on ethical conduct in human research.

Technical Report E72, National Health and Medical Research Council.



NHMRC (2018b).

Open access policy - january 2018.

Technical report, National Health and Medical Research Council.



Nosek, B. A., Alter, G., Banks, G. C., Borsboom, D., Bowman, S. D., Breckler, S. J., Buck, S., Chambers, C. D., Chin, G., Christensen, G., Contestabile, M., Dafoe, A., Eich, E., Freese, J.,

Glennerster, R., Goroff, D., Green, D. P., Hesse, B., Humphreys, M., Ishiyama, J., Karlan, D., Kraut, A., Lupia, A., Mabry, P., Madon, T. A., Malhotra, N., Mayo-Wilson, E., McNutt, M., Miguel, E., Paluck, E. L., Simonsohn, U., Soderberg, C., Spellman, B. A., Turitto, J., VandenBos, G., Vazire, S., Wagenmakers, E. J., Wilson, R., and Yarkoni, T. (2015).

Scientific standards: Promoting an open research culture.

Science, 348(6242):1422–1425.



NSW (1998).

Privacy and personal information protection act 1998.

Technical Report 133.



NSW (2015).

GDA-23 general retention and disposal authority: University records.

<https://arp.nsw.gov.au/gda-23-general-retention-and-disposal-authority-univers>

Accessed: 2019-6-3.



OAIC (2019a).

Australian privacy principles.

<https://www.oaic.gov.au/privacy-law/privacy-act/australian-privacy-principles>.

Accessed: 2019-6-3.



OAIC (2019b).

Notifiable data breaches scheme.

<https://www.oaic.gov.au/privacy-law/privacy-act/notifiable-data-breaches-scheme>.

Accessed: 2019-6-3.



OFRG (2019).

Open research funders group.

<http://www.orfg.org/>.

Accessed: 2019-6-3.



Open Science Collaboration (2015).

PSYCHOLOGY. estimating the reproducibility of psychological science.

Science, 349(6251):aac4716.



OSF (2014).

Sample implementation of guidelines for transparency and openness promotion (TOP) in journal policies and practices (version 1.0.2).

<https://osf.io/edtxm/>.

Accessed: 2019-4-20.



Perkel, J. M. (2018).

A toolkit for data transparency takes shape.

Nature, 560(7719):513–515.



Prinz, F., Schlange, T., and Asadullah, K. (2011).

Believe it or not: how much can we rely on published data on potential drug targets?

Nature reviews. Drug discovery, 10(9):712.



Roosevelt, C. H., Cobb, P., Moss, E., Olson, B. R., and Ünlüsoy, S. (2015).

Excavation is destruction digitization: Advances in archaeological practice.

Journal of Field Archaeology, 40(3):325–346.



Ross, S. A., Ballsun-Stanton, B., Sobotkova, A., and Crook, P. (2015).

Building the bazaar: Enhancing archaeological field recording through an open source approach.

In Wilson, A. T. and Edwards, B., editors, *Open Source Archaeology: Ethics and Practice*, pages 111–129. De Gruyter Open, Warsaw, Poland.



Ross, S. A., Sobotkova, A., Ballsun-Stanton, B., and Crook, P. (2013).

Creating eresearch tools for archaeologists: The federated archaeological information management systems project.

Australian Archaeology, 77(1):107–119.



Sobotkova, A. (2018).

Sociotechnical obstacles to archaeological data reuse.

Advances in Archaeological Practice, 6(2):117–124.



Sobotkova, A., Ballsun-Stanton, B., Ross, S., and Crook, P. (2015).

Arbitrary offline data capture on all of your androids: The FAIMS mobile platform.

In Traviglia, A., editor, *Across Space and Time. Papers from the 41st Annual Conference of Computer Applications and Quantitative Methods in Archaeology (CAA)*, pages 80–88. Amsterdam University Press.



Sobotkova, A., Ross, S. A., Ballsun-Stanton, B., Fairbairn, A., Thompson, J., and VanValkenburgh, P. (2016).

Measure twice, cut once: Cooperative deployment of a generalized, Archaeology-Specific field data collection system.

In Averett, E. W., Gordon, J. M., and Counts, D. B., editors, *Mobilizing the Past for a Digital Future: The Potential of Digital Archaeology*, pages 337–371. The Digital Press @ University of North Dakota, Grand Forks, ND.



Stewart Lowndes, J. S., Best, B. D., Scarborough, C., Afflerbach, J. C., Frazier, M. R., O'Hara, C. C., Jiang, N., and Halpern, B. S. (2017).

Our path to better science in less time using open data science tools.

Nature Ecology & Evolution, 1(6):s41559–017–0160.



Strategyzer AG (2019).

Strategyzer | trusted by over 5 million business practitioners.

<https://www.strategyzer.com/>.

Accessed: 2019-4-24.



TERN (2019).

TERN - terrestrial ecosystem research network.

<https://www.tern.org.au/>.

Accessed: 2019-4-24.



van Edig, X. (2018).

Copernicus publications - data policy.

https://publications.copernicus.org/services/data_policy.html.

Accessed: 2019-3-29.



VanValkenburgh, P., Silva, L. O. G., Repetti-Ludlow, C., Gardner, J., Crook, J., and Ballsun-Stanton, B. (2018).

Mobilization as mediation: Implementing a Tablet-Based recording system for ceramic classification.

Advances in Archaeological Practice, pages 1–15.



Vines, T. H., Albert, A. Y. K., Andrew, R. L., Débarre, F., Bock, D. G., Franklin, M. T., Gilbert, K. J., Moore, J.-S., Renaut, S., and Rennison, D. J. (2014).

The availability of research data declines rapidly with article age.

Current biology: CB, 24(1):94–97.



Wilkinson, M. D., Dumontier, M., Aalbersberg, I. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J.-W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., Hoen, P. A. C. t., Hooft, R.,

Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S.-A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., van der Lei, J., van

Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Mons, B. (2016).

The FAIR guiding principles for scientific data management and stewardship.

Scientific Data, 3:160018.



MACQUARIE
University
SYDNEY · AUSTRALIA

From current practice to better practice

How do we get from where we are now to where we want to be?

- Understand the evolving expectations of transparent research.
 - Look past desktop software (Excel, ARCGIS, Filemaker, Access, etc.).
 - Rally around emerging research- and domain-specific solutions (even if imperfect).
 - Overcome ‘not invented here’; you don’t need a bespoke solution.
 - Budget for ‘ground-up’ transparency (data and code). Up-front costs will be high but offer longer-term payoffs (in costs, time, and quality).
 - Implement (and budget for) fundamental good practice in data and code management before other technologies.
 - Improve research design (prioritise approach over methods)
- [Muthukrishna and Henrich, 2019, Hole, 1973]



MACQUARIE
University
SYDNEY • AUSTRALIA

References

Source code for this presentation is available at: <https://github.com/saross/Ross-FAIMS-current>.

FAIMS Project software and documentation can be found at:
<https://github.com/faims>.

This work is licensed under a Creative Commons Attribution 4.0 International License.