

~~PRESERVATION OF~~ RESEARCH DATA: A CRITICAL PART OF FAIR!

Jonathan Crabtree

Director of Research Data Information Systems

Jonathan_Crabtree@unc.edu



UNC
THE ODUM INSTITUTE

THE ODUM INSTITUTE FOR RESEARCH IN SOCIAL SCIENCE
DAVIS LIBRARY, 2ND FLOOR, CB# 3355
UNIVERSITY OF NORTH CAROLINA AT CHAPEL HILL

ONLINE: www.odum.unc.edu
TWITTER: @Odum_Institute
YOUTUBE: The Odum Institute

BEM-VINDA

- Primeiro, eu gostaria de dizer meu coração está com todos lutando contra o Covid, especialmente todos os meus grandes amigos no Brasil
- Sinto falta do Brasil mais do que você imagina

OUTLINE

- What is FAIR?
- What are the FAIR challenges with research data?
- Why is FAIR not enough?
- What is TRUST?
- FAIR and TRUST compliment each other!
- Questions

FAIR PRINCIPLES



Herterich, Patricia, & Davidson, Joy. (2020). How repositories can contribute their FAIR share. Zenodo.
<http://doi.org/10.5281/zenodo.3872074>

The Magnifying glass, Tap, Gears set, Recycle sig, Storage, Infinity, Discussion, Shield, and Man User icons made by Freepik from www.flaticon.com are licensed by [CC 3.0 BY](https://creativecommons.org/licenses/by/3.0/). All other icons made by ARDC. Entire FAIR resources graphic is licensed under a [Creative Commons Attribution 4.0 International License](https://creativecommons.org/licenses/by/4.0/)



WHAT ARE FAIR CHALLENGES WITH RESEARCH DATA?

WHY IS RESEARCH DATA DIFFERENT?

- All the normal digital object preservation are required
- Many resources to guide digital preservation at this level
 - Cariniana
 - <http://cariniana.ibict.br/index.php/pre-dig>
 - Política de Preservação Digital Arquivo Nacional
 - http://www.arquivonacional.gov.br/images/conteudo/artigos/AND_Politica_Preservacao_Digital_v2.pdf
 - Educopia released Guide to Documenting Born-Digital Archival Workflows
 - https://educopia.org/wp-content/uploads/2020/06/OSSArcFlow_Guide_FINAL.pdf
 - U.S. National Archive (NARA) updated its Digital Preservation Framework
 - <https://github.com/usnationalarchives/digital-preservation>

Data Curation Profiles Toolkit



- “A Data Curation Profile is essentially an outline of the ‘story’ of a data set or collection, describing its origin and lifecycle within a research project.”
- “The Profile and its associated Toolkit grew out of an inquiry into the changing environment of scholarly communication, especially the possibility of researchers providing access to data much further upstream than previously imagined.”
- The DCP project explored the roles of researchers and librarians in data sharing.

<http://datacurationprofiles.org>



UNC
THE ODUM INSTITUTE



[Home](#)
[DMP Requirements](#)
[Public DMPs](#)
[News](#)
[Help](#)
[Contact Us](#)
[About ▾](#)
[Log In](#)

Data Management Planning Tool

Create, review, and share data management plans that meet institutional and funder requirements.

[Get Started](#)



PUBLIC DMPs

List of sample data management plans provided by DMPTool users.

- » CAREER: Parietal Cortex and the Transformation of Spatial Cognition into Action
- » : Biosignature Suites: Using Connections between Microbes and Minerals to understand Biogenic Carbonates
- » A unified approach to preserving cultural software objects and their development histories

[View All](#)



DMPTOOL NEWS

Latest information about data management and the DMPTool.

- » We need API use cases!
- » DMPTool downtime this Saturday 7/12/14
- » Misc Stats for DMPTool2
- » DMPTool workshop at DataONE Users Group...
- » New DMPTool Released Today

[More News](#)



DMPTOOL HELP

Overview of how to use the tool, plus resources and guidance on data management.

- » Frequently Asked Questions
- » Create a DMP
- » Administer the DMPTool
- » Data management guidance
- » Community resources

[View All](#)



DMPTOOL is a service of the University of California Curation Center of the California Digital Library
Copyright © 2010-2014 The Regents of the University of California

[Privacy Policy](#) |
 [Accessibility Policy](#) |
 [Terms of Use](#) |
 [Contact Us](#) |
 [About](#)





<https://dmp.cdlib.org>



UNC
THE ODUM INSTITUTE

CONTEXT CRITICAL

- The research process has many actors
- The research process often has many instruments
- Many research methodologies
- Often very temporal
- Often hundreds of potential controls around experiments
- Data formats vary widely and are effected by context and methodology
- Unfortunately also impacted by political and social pressures!

SOFTWARE & CODE

- Often many components of research workflows
- Often specialized software and proprietary formats
- Often custom code
- Often novel analysis that few understand

SENSITIVITY OF RESEARCH DATA

- Human subject issues
- Health information
- Privacy issues
- Variation by country and culture
- Variation by research methodology



WHY IS FAIR NOT ENOUGH?

CHALLENGES WITH “AIR”

- Even if you preserve digital objects using industry standards they may not have “AIR”
- It may be dirty “AIR”
- It may be untrustworthy “AIR”

PRESERVATION REQUIRES SPECIAL SKILLS

- Archivists and Curators must have additional special skills
- Access to statisticians and researchers for assistance
- These skills vary across disciplinary content
- These skills vary across research methodologies
- These skills vary across data formats

TRUST IN “IR”

- Preservation package needs to contain information to ensure trust
- Archive or repository must ensure that the data can be trusted
- Or at the very least give users enough information that they can make the determination
- There are situations that researchers use marginal data to produce good results

TRUST IN REPRODUCIBILITY

- Highest level of TRUST is to ensure reproducibility
- Can anyone in the future use the data and code to repeat the process
- Will they have enough information and materials to reproduce the results
- Original researchers and curators could be unavailable

In addition to the C/C2 polymorphism in human transferrin, position 589 toggles exclusively between proline and serine across the primate lineage (Fig. 2B and fig. S10), a potential signature of antagonistic pleiotropy at a largely constrained position, as observed for other host-pathogen interfaces (7). Previous work has also implicated the C2 transferrin variant as a risk factor for disorders involving iron metabolism, including Alzheimer's disease; however, these associations remain controversial and appear dependent on the populations tested and interactions with other susceptibility loci (25, 26). Our findings provide a functional basis for human transferrin variation and establish an important role for nutritional immunity in recent human evolution.

Although canonical innate immunity factors have been appreciated as nodes of host-virus evolution, our work demonstrates that nutritional immunity has played a fundamental role in the survival of primate populations challenged by bacterial pathogens. *H. influenzae* and *N. meningitidis* remain a major source of morbidity and mortality in regions where vaccine coverage is poor (27, 28) and drug-resistant *N. gonorrhoeae* is developing into an urgent public health threat (29). By illuminating the battle for iron as a major driving force of host-pathogen evolution, from 40 million years of primate divergence to emerging human epidemics today, our studies reveal new reservoirs of genetic resistance to infectious diseases.

REFERENCES AND NOTES

1. E. D. Weisberg, *JAMA* 228, 39–41 (1975).
2. L. E. Grout, E. P. Shaw, *Cell Host Microbe* 13, 509–519 (2013).
3. M. I. Havel, E. P. Shaw, *Nat. Rev. Microbiol.* 10, 525–530 (2012).
4. E. Park et al., *Proc. Natl. Acad. Sci. USA* 109, 12631–12636 (2012).
5. M. D. Daugherty, H. S. Mudd, *Proc. Natl. Acad. Sci. USA* 109, 177–178 (2012).
6. A. Dominguez, J. Abramson, J. M. H. S. Mudd, *Proc. Natl. Acad. Sci. USA* 109, 177–178 (2012).
7. R. C. Loh, S. J. Choi, A. P. G. H. S. Mudd, *Proc. Natl. Acad. Sci. USA* 109, 177–178 (2012).
8. M. R. Patel, Y. M. Liu, S. M. Havel, E. P. Shaw, *Proc. Natl. Acad. Sci. USA* 109, 177–178 (2012).
9. J. T. Kauler et al., *PLoS Pathog.* 8, e1002666 (2012).
10. R. Hargreaves et al., *Nature* 483, 52–58 (2012).
11. A. B. Schreyers, L. J. Morris, *Nat. Microbiol.* 2, 289–288 (2012).
12. A. B. Schreyers, S. Gray Owen, *J. Infect. Dis.* 205 (Suppl. 1), S183–S184 (2012).
13. J. Abarca, R. H. Yu, A. B. Schreyers, *Mol. Microbiol.* 8, 125–143 (2012).
14. A. B. Schreyers, G. C. Gonzalez, *Infect. Immun.* 80, 2425–2424 (2012).
15. M. I. Zaslavskii et al., *Infect. Immun.* 75, 5600–5614 (2007).
16. C. Calzavara, J. Abarca, R. H. Yu, A. B. Schreyers, T. J. Morris, *Nat. Microbiol.* 2, 289–288 (2012).
17. C. N. Gonsky, G. D. Brown, J. J. Spaulding, *J. Infect. Dis.* 240, 2440–2440 (2012).
18. D. B. Schreyers, B. C. Lee, *Cell Microbiol.* 15, 409–415 (2013).
19. D. B. Schreyers, B. C. Lee, *Cell Microbiol.* 15, 409–415 (2013).
20. P. Hargreaves, S. Spillars, *Nat. Genet.* 43, 95–96 (2011).
21. P. J. Potts et al., *Evolutionary Biology of Infectious Agents* (2014).
22. M. E. Brown, H. S. Mudd, *PLoS Biol.* 8, e1000301 (2010).

23. B. J. Shapiro, L. A. David, J. Friedman, E. J. Alm, *J. Infect. Microbiol.* 12, 195–204 (2010).
24. Y. Nishikawa et al., *Hum. Genet.* 120, 129–130 (2007).
25. Y. Wang et al., *Can. J. Neurol. Sci.* 40, 468–467 (2013).
26. K. L. O'Brien et al., *Lancet* 376, 859–862 (2010).
27. S. A. Johnson et al., *Nature* 30 (Suppl. 2), 105–106 (2012).
28. J. Friedman, *Antibiotic Resistance Threats in the United States 2013* (Centers for Disease Control and Prevention, U.S. Department of Health and Human Services, 2013).

ACKNOWLEDGMENTS

We thank M. Mulvey and members of the Mulvey laboratory for assistance with bacterial genetics and helpful discussions; V. Chandra and W. S. Garg for assistance with transferrin protein purification; J. Kaplan and D. Ward for helpful advice and discussions; J. St. Geme II and F. Porisch for *Haemophilus* strains and helpful discussions; M. So for providing *Neisseria* DNA and strains; S. Wang for providing mouse mucosa tissue samples; Z. Krimberg for assistance with human population genetics; and C. Condit for the gift of the TfpA antibody as well as helpful discussions.

SUPPLEMENTARY MATERIALS

www.science.org/cgi/content/full/10.1126/science.1256151. Materials and Methods (fig. S1 to S13), Tables S1 to S10, Movies S1. References (30–34). 29 July 2014; accepted 14 November 2014. 10.1126/science.1256151

POLITICAL SCIENCE

When contact changes minds: An experiment on transmission of support for gay equality

Michael J. LaCour¹ and Donald P. Green²

Can a single conversation change minds on divisive social issues, such as same-sex marriage? A field experiment tested whether gay (n = 22) or straight (n = 72) canvassers were effective at encouraging voters (n = 972) to support same-sex marriage and whether attitude change persisted to spread to others in voters' social networks. The results measured by an unrelated panel survey, show that both gay and straight canvassers produced large effects initially, but only gay canvassers' effects persisted in 3-week, 6-week, and 9-month follow-ups. We also find strong evidence of within-household transmission of opinion change, but only in the wake of conversations with gay canvassers. Contact with gay canvassers further caused substantial change in the attitudes of gay men and lesbians more generally. These effects, however, and contagious effects were confirmed by a follow-up experiment in which canvassers coupled with discussion of issues pertinent to them, rather than of issues unrelated to them, produced a causal effect of opinion change.

Forced among themselves to have robust conversations (1) is the hypothesis that, when people from different groups interact with one another, enough contact eventually leads to reduced prejudice toward the other group (2). If it is often difficult to facilitate intergroup contact of sufficient duration to dispel negative stereotypes and build empathy. For this reason, research attention has recently focused on alternative interventions that may be deployed in a more compressed time frame. Examples include brief personal contact with outgroup members during the course of a conversation (4) and the "extended contact" that occurs when one learns that a close friend has experienced positive contact with an outgroup (5). The question is whether brief or indirect con-

tact is sufficient to produce meaningful and enduring attitude change. At this point, noting the lack of randomised experiments that track attitudes over time after the intervention (6).

Our theoretical contribution is to introduce the distinction between active and passive contact, which are posited to produce different effects in the context of a brief intergroup encounter. Whereas passive contact involves personal exposure to an outgroup member (e.g., through collaborative activity), active contact involves, in addition, communication about an issue that divides the two groups (e.g., discussion of recent communal violence). The effects of active contact doubtless depend on whether the conversation is respectful or accusatory, but in principle, active contact has the potential to both reduce hostility toward outgroups and to change attitudes on divisive issues. Our empirical contribution is the first field-based experimental demonstration of persistent attitude change in the wake of active

Downloaded from http://science.sciencemag.org/ on February 23, 2017



...or it didn't happen.

LaCour, M. J., & Green, D. P. (2014). When contact changes minds: An experiment on transmission of support for gay equality. *Science*, 346(6215), 1366–1369. <https://doi.org/10.1126/science.1256151>

based on scans acquired at rest, our model is blind to different strategies that are chosen by the participants in performing a given task. We refer to these features as "inherent," but we acknowledge that these can be "structurally inherent" (related to brain organization and connectivity) or "functionally inherent" (related to the cognitive state of subjects during the resting-state scan).

The idea that brain connectivity can predict activation has previously been reported for different modalities (22), where diffusion MRI tractography was used (23) to measure connectivity. This study was limited to a specific cognitive task and a predefined anatomical region. More recently, resting-state connectivity has been shown to be predictive of subjects' identity, in a way similar to a fingerprint (24). Rather than simply identifying subjects, our goal was to predict the entire layout of brain activity for each subject. Moreover, we also aim to predict such layout of activity in a number of different cognitive domains, from a single task-free scan, including in subjects that show patterns of activation that are different from the group average (perhaps most strikingly in right lateralized subjects when the majority of training subjects are left-lateralized).

There are important practical implications of the proposed framework in basic research and translational neuroscience. It provides a method for inferring multiple individualized functional localizers based on a single resting-state scan. Such a tool could be used to investigate in detail the response profiles of localized brain regions without the need to acquire often time-consuming task localizers. Such a tool, if generalizable beyond the young, healthy population that makes up the HCP database, could be used to investigate functional regions in subjects who cannot perform tasks, such as infants or paralyzed patients.

REFERENCES AND NOTES

1. M. Jonides, *J. South. Med. Assoc.* **5**, 143-156 (2002).
2. J. B. Lee, M. W. Shacter-Perschke, H. Iwaki, S. Fennell, D. Schlegel, *J. Neurophysiol.* **100**, 2351-2365 (2008).
3. F. M. N. Le, *J. Neurosci.* **22**, 1031-1037 (2002).
4. A. M. Lee, *J. Neurosci.* **27**, 1410-1413 (2007).
5. S. K. Lee, C. H. Lee, C. T. Lee, R. A. Poldrack, *Science* **305**, 585-588 (2004).
6. S. K. Lee, *J. Proc. Natl. Acad. Sci. U.S.A.* **101**, 10510-10514 (2004).
7. R. K. Lee, *J. Neurosci.* **27**, 221-242 (2007).
8. T. E. Lee, M. W. Shacter-Perschke, H. Iwaki, *J. Soc. Neurosci.* **24**, 1031-1037 (2002).
9. R. E. Passafium, R. E. Shapira, R. K. Lee, *Neurosci. Lett.* **349**, 103-107 (2003).
10. R. E. Lee, *J. Neurophysiol.* **106**, 1325-1335 (2002).
11. S. K. Lee, *J. Proc. Natl. Acad. Sci. U.S.A.* **101**, 10510-10514 (2004).
12. D. C. Van Essen et al., *Neuroimage* **88**, 62-79 (2015).
13. D. C. Van Essen et al., *Neuroimage* **88**, 309-319 (2015).
14. R. F. F. et al., *Cereb. Cortex* **28**, 1573-1580 (2018).
15. T. J. Andrews, S. D. Nijm, D. Parvizi, *J. Neurosci.* **37**, 2850-2861 (2017).
16. S. K. Lee et al., *Brain* **123**, 74-81 (2000).
17. M. Thiel, in *Statistical Models of Brain Activity*, 205-228 (2005).
18. W. W. Lee, C. T. Lee, J. D. Lee, T. E. Lee, S. K. Lee, *Neurosci. Lett.* **378**, 238-241 (2004).
19. T. E. Lee, *Neurosci. Lett.* **378**, 238-241 (2004).
20. R. K. Lee, *Neurosci. Lett.* **378**, 238-241 (2004).
21. R. K. Lee, *Neurosci. Lett.* **378**, 238-241 (2004).
22. J. M. Lee, *Brain Connect.* **1**, 13-36 (2015).
23. J. M. Lee, *Brain Connect.* **1**, 35-57 (2015).
24. P. J. Lee, S. K. Lee, C. T. Lee, J. D. Lee, A. K. Lee, *Brain Connect.* **1**, 35-57 (2015).
25. E. S. Lee et al., *Nat. Neurosci.* **18**, 1654-1671 (2015).

SUPPLEMENTARY MATERIALS

www.sciencemag.org/content/352/6282/1031/suppl/DC1
Materials and Methods
Fig. S1 to S23
Tables S1 to S7
Caption for Table S1 to S7
References (25-30)
Tables S1 to S7
1 November 2016; accepted 29 February 2017
10.1126/science.1257777

POLITICAL SCIENCE

Durably reducing transphobia: A field experiment on door-to-door canvassing

David Broockman¹* and Joshua Kalla²*

Existing research depicts intergroup prejudices as deeply ingrained, requiring intense intervention to lastingly reduce. Here, we show that a single approximately 10-minute conversation encouraging actively taking the perspective of others can markedly reduce prejudice for at least 3 months. We illustrate this potential with a door-to-door canvassing intervention in South Florida targeting anti-transgender prejudice. Despite declines in homophobia, transphobia remains pervasive. For the intervention, 56 canvassers went door to door encouraging active perspective-taking with 501 voters at voters' doorsteps. A randomized trial found that these conversations substantially reduced transphobia, with decreases greater than Americans' average decrease in homophobia from 1998 to 2012. These effects persisted for 3 months, and both transgender and nontransgender canvassers were effective. The intervention also increased support for a nondiscrimination law, even after exposing voters to counterarguments.

Intergroup prejudice, defined broadly as negative attitudes about an outgroup, is a root cause of numerous adverse social, political, and health outcomes (1-3). Influential theories depict intergroup prejudices as deeply ingrained during childhood and highly resistant to change thereafter (4-6). Consistent with these theories, empirical research has found that durably reducing prejudice is challenging. Mass media interventions and other brief stimuli may fail to reduce prejudiced attitudes (7) or have only temporary effects (8); lasting change ap-

pears to require intense intervention over months (9, 10). Rare are studies demonstrating prejudice-reduction interventions relatively brief in duration yet proven to have lasting effects (6).

Theories of active processing, however, suggest a method for even brief interventions to durably change attitudes. A recurring finding of laboratory studies is that brief messages can durably change individuals' attitudes when individuals engage in active, effortful, processing (known as "system 2" processing) of these messages (11). These studies, conducted on other topics, raise the possibility that brief interventions encouraging active consideration of counterprejudicial thoughts could produce lasting changes in attitudes toward an outgroup. Perspective-taking, "imagining the world from another's vantage point," is one

¹University School of Business, Stanford University, Stanford, CA, USA. ²Department of Political Science, University of California, Berkeley, CA, USA.
*Corresponding author. E-mail: dbroockman@stanford.edu

DATA
A

<http://dx.doi.org/10.7910/DVN/WKR39N>

"The data are solid and the analysis convincing," says Gabriel Lenz, a political scientist at UC Berkeley who was asked by the funders of the study to verify that the data for this new study were truly collected.

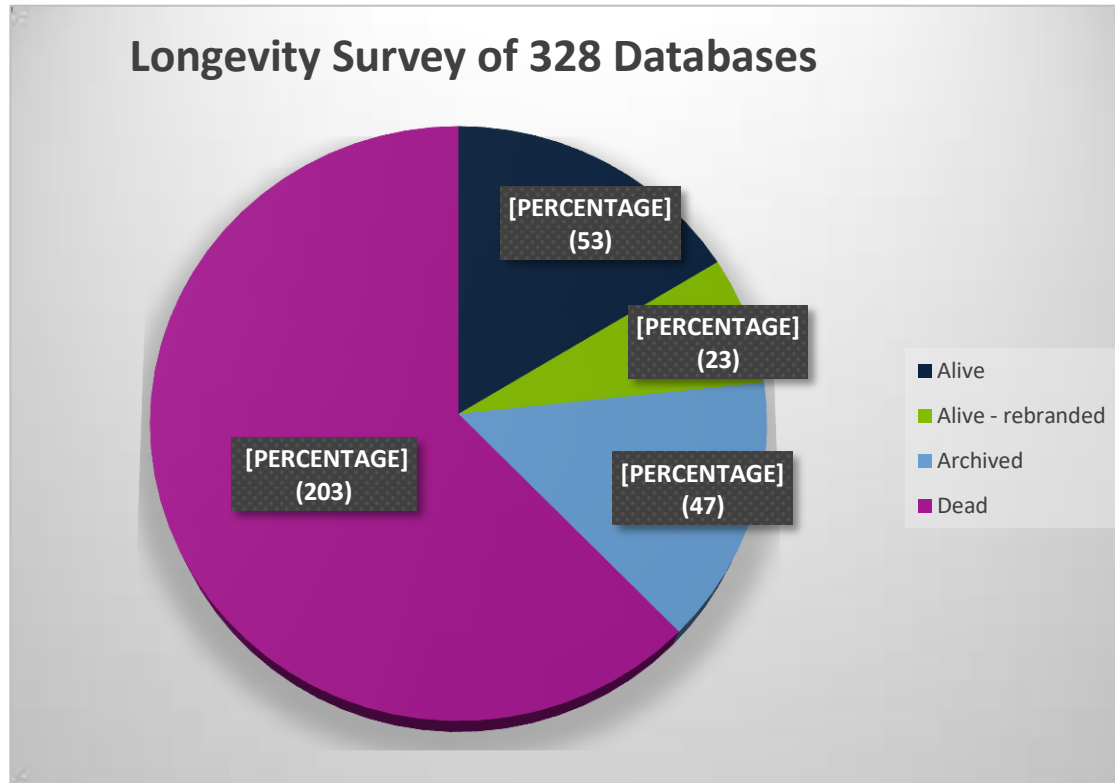
Bohannon, J. (2016). Ironic coda to fraudulent study of bias. *Science*, 352(6282), 131-132. <https://doi.org/10.1126/science.352.6282.131>

Broockman, D., & Kalla, J. (2016). Durably reducing transphobia: A field experiment on door-to-door canvassing. *Science*, 352(6282), 220-224. <https://doi.org/10.1126/science.aad9713>



WHAT ARE TRUST PRINCIPLES?

THE LONGEVITY OF 328 BIOMEDICAL DATABASES OVER 18 YEARS

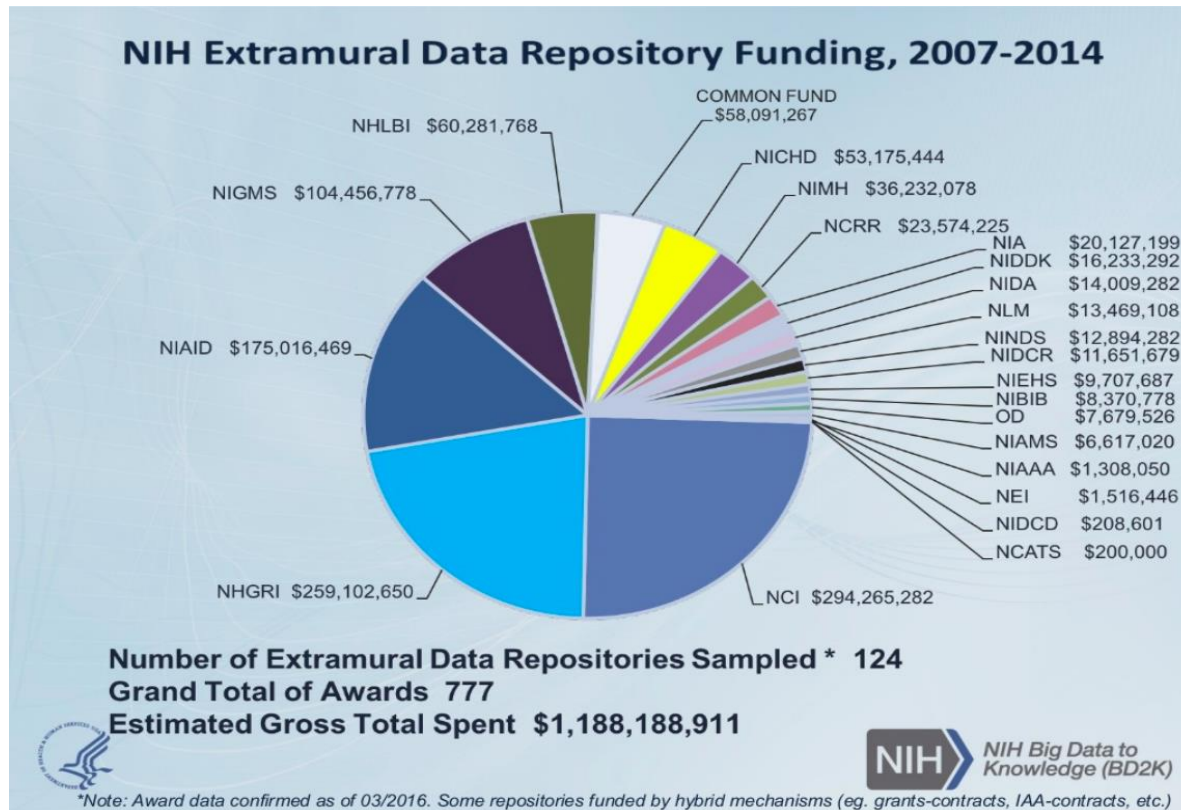


UNC
THE ODUM INSTITUTE



<http://journal.embnet.org/index.php/embnetjournal/article/view/803/1209>
(Attwood et al, 2015)

Money not always the solution



FAIR VS TRUST

- **FAIR** defines the properties of data and metadata.
- **TRUST** describes the characteristics of data repositories that are responsible for managing and disseminating the data over a long period of time.



FAIR DATA IN REPOSITORIES WE TRUST

T



Transparency

R



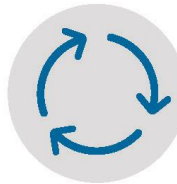
Responsibility

U



User focus

S



Sustainability

T



Technology



TRUST is about providing a trusted repository for archiving and distributing data.

TRUST is about having transparent policies, organizational capabilities and people behind the websites, infrastructure, and databases, who understand deeply what FAIR means to the users of their designated community.

TRUST is about maintaining reliable and secure operations through technology and data stewardship procedures.

TRUST is about sustaining infrastructures that are needed to support sustainable operations and long-term data and knowledge preservation.

TRUST also represents a commitment to transparently fulfill the services promised to support the continuing use of data.

TRUST PRINCIPLES

- **FAIR** defines the properties of data and metadata
- **TRUST** describes the characteristics of data repositories that are responsible for managing and disseminating the data over a long period of time
- **FAIR** data in repositories we **TRUST**

T - Transparency is achieved by providing publicly accessible evidence of the services that a repository can and can not offer.

R - Responsibility is a commitment to provide high (technical) quality data services.

U - User focus is the focus on the uses and potential uses of the data and services offered.

S - Sustainability is the capability to support long-term data preservation and use.

T - Technology is the infrastructure and capabilities to support the repository operations.

FAIR AND TRUST ARE COMPLEMENTARY

“Research data will not become nor stay FAIR by magic. We need skilled people, transparent processes, interoperable technologies and collaboration to build, operate and maintain research data infrastructures.”

Mari Kleemola, Finnish Social Science Data Archive/CoreTrustSeal Board, Secretary

<https://tietoarkistoblogi.blogspot.com/2018/11/being-trustworthy-and-fair.html>

TRUST PRINCIPLES ACQUIRING COMMUNITY SUPPORT

- Nature Publication
 - Lin, D., Crabtree, J., Dillo, I. *et al.* The TRUST Principles for digital repositories. *Sci Data* **7**, 144 (2020). <https://doi.org/10.1038/s41597-020-0486-7>
- NIAID Support
 - <https://www.niaid.nih.gov/grants-contracts/trust-digital-repositories>
- RDA Adoption
 - <https://www.rd-alliance.org/trust-principles-rda-community-effort>

DISCUSSION & QUESTIONS

■ Obrigado

CONNECT WITH THE ODUM INSTITUTE

The Odum Institute

<http://www.odum.unc.edu>



@Odum_Institute



The Odum Institute