Tiny Data: Building a Community of Practice around Humanities Datasets

Veronica Ikeshoji-Orlati  
@vikeshojiorlati  
21 February, 2018
What does it mean to produce FAIR data in the humanities?
Tiny Data: What’s in a Name?

“Data are big or little in terms of what can be done with them, what insights they can reveal, and the scale of analysis required relative to the phenomenon of interest.”

- Christine Borgman, 2015 (6)

‘Tiny Data’ are at the core of traditional and contemporary humanistic inquiry, reflecting scholars’ critical engagements with texts, images, sound, and performance.

Context: Digital Humanities at Vanderbilt University
Established with a grant from the Andrew W. Mellon Foundation, the Vanderbilt University Center for Digital Humanities is now open, as of October 2016.

Come to the DH Center to find learning experiences, project support, and intellectual community for your creative and scholarly explorations of the intersections of digital technology and humanistic inquiry.

Skeptics, novices, and those with developed interests in digital scholarship are all welcome.

See our Events schedule, our Working Groups, and our Calendar for ways to get involved - or just drop by during our Tuesday co-working hours, 3:00-4:00 pm.
Phase 1: Bringing Together a Community
Nathan asks: Does network analysis constitute another project or does it enhance a project? Is it too labor intensive? Is it a new iteration? Do the questions shift? What does it do behind the scenes? How is it (un)productive?

Suellen: Views it with approaching it from the little data perspective; feels that is where it can enhance the research/project.

Tiny Data — Using big data tools to engage with small data objects.
Digital Scholarship

Working Groups

The purpose of these working groups is to study and learn technologies in an ongoing collaborative environment. If you are interested in moving beyond familiarity with digital scholarship to mastery of a particular technique, these working groups are for you! Newcomers are always welcome. If you haven’t taken part until now, we’ll help you catch up. Questions? Contact the leaders of the various groups below or email Clifford Anderson, Associate University Librarian for Research and Learning: clifford.anderson@vanderbilt.edu

Semantic Web Working Group

Mondays, 12:10-1:00pm, 800FA, Central Library

Using the architecture of the World Wide Web, (such as URI's and http,) the Semantic Web supports a common framework allowing data to be shared and reused across boundaries both technical and social. The Semantic Web is built on the RDF model which allows for standardized data interchange on the web. RDF consists of triples comprised of a subject, predicate and object and with the SPARQL query language we can then ask questions, manipulate and transform RDF encoded data. For more information, please visit the working group website or contact suelten.stringer-lye@vanderbilt.edu
Tiny Data Working Group

Learn about Tiny Data in the company of Digital Humanists

We meet most Wednesdays from 12:10 to 1:00 p.m. in 800FA of the Central Library. Newcomers are always welcome! Check the schedule below for an outline of topics. Also see our Github repository for code samples. Contact Suellen Stringer-Hye with any questions or see the FAQs below.

What will we be working on?

Tiny Data
What is Tiny Data and what distinguishes it from Big Data.

Working Sessions
"Real World" examples using data from your projects.

Discussion
How can Tiny Data help solve your current research problems?
Phase 2: Establishing a Curriculum
Fall 2017 Topics

- What is Tiny Data?
- Scalability: Making Digital Humanities Data Modular
- Digital Project Methodology, Data vs. Metadata
- Collaborating on a DH Project: Who/What/When/Where
- Data vs. Interface
- Interactivity and Audience
- Sustainability & DH: Making Data Wills
- Bring Your Research (Data) Workshops

In this working group, we discuss how to collect, manipulate, analyze, and publish small- and medium-sized datasets for use in humanistic research. By examining the data behind diverse digital humanities projects, discussing readings about quantitative and qualitative data, and experimenting with data visualization tools such as Neo4j and Tableau Public, we aim to build a collaborative and multi-disciplinary cohort of data-literate humanists across campus.

September:

12 Introduction: What is Tiny Data?
What do you want to get out of this working group?
What do we mean by “Tiny Data”?
How can we evaluate and learn from established DH projects?

19 Scalability: Making Digital Humanities Data Modular
How do you work from a dissertation/article level project to something bigger and more sustainable?
What does making your current work long-lasting look like/mean?
How do you set priorities for a multi-part DH project?
How do you keep the “humanities” in ‘DH’?
Is ‘digital’ critical to your project?
Reading: C.L. Borgman. 2015. Big Data, Little Data, No Data: Scholarship in the Networked World (Chapter 7) - available online via the library
Project: Cass Gilbert’s Woolworth Building
Phase 3: Producing Tangible Deliverables
Collaborative Data Wrangling
Takeaways
Lessons Learned

- Need to define FAIR data for the humanities

- Success through focusing on research questions, not data curation methods

- Redefining librarianship as data consultancy
Next Steps
Towards a Community of Data Curation Practice for Humanists

- Make the community self-sustaining

- Tiny Data documentation for humanists

- Rethink what FAIR humanities data look like and the potential impact on humanities researchers’ ideas of reproducibility in their fields

- Teach liaison librarians how to find, and encourage engagement with, existing datasets
Thank you!

Veronica Ikeshoji-Orlati  v.ikeshoji-orlati@vanderbilt.edu  @vikeshojiorlati
Mary Anne Caton  mary.a.caton@vanderbilt.edu
Suellen Stringer-Hye  suellen.stringer-hye@vanderbilt.edu  @suellenstringerhye