

Note taking template
Academic Planning Summit

January 16, 2015

Title of Session: Applications and Implications of Big Data

Faculty Facilitator(s) Joan Peckham and Julia Lovett

Note taker name Abigail Casavant

Check one:

Morning Breakout 1 **Morning breakout 2**

Afternoon breakout **Afternoon breakout 2**

Questions from the program: What does URI need in order to move forward with utilizing big data in the future? Discussion will focus on lessons and applications for the future of digital humanities, business and government arenas. Issues to consider are the acquisition of resources and encouraging collaboration across the disciplines and the steps to facilitate participation in big data projects.

Ideas proposed/Recommendations

What does URI need to move forward with big data?

Bid data in the humanities

Digital humanities – digital toolbox expands what we can do, how do we use it for big data and how will it help us teach in the future? It can expand research and teaching

Now it is very easy to find and use

Video Summary: Digitizing “Republic of Letters,” add life to lost world, map not the whole but selected case studies to represent whole

How do we use big data? A wild idea – Pell Archives, applied to RI Humanities Grant, suggestions for how to facilitate inter-disciplinary participation (Education, history, political science, library studies)

What tools are needed to apply big data?

- What do you do with all of this data
- Standardize a big data course and development program
- Staff, administrators need to be introduced and familiarized to big data too, not just students
- Business not listed in the slides – business and politics, could find interesting case studies, correlate with business law
- Documents and big data for historians – born digital, stay digital, disappear, need to harvest and store and categorize online data for Rhode Island
- We must have a broad view of what is worth saving
- Add biology to the list – everyone in the future will need to understand big data and high performance computing
- High performance computing can help when there is a lack of data points
- 3 Vs- Velocity variety and volume

- How important is metadata and organization online, data resource management needed on campus
- We must think about the entire life cycle of the data in order for it to be shared and used
- This center would also created opportunities for students
- Methodology courses needed to learn how to do this research, you have to know the tools
- Big data minor or certificate?
- Ways to use analysis around learning analytics – what do students learn here and what do they know before they come? Analyze bigger patterns of student knowledge in mathematics and writing
- What is the infrastructure that would help with research?
 - Genomic datasets – huge data volume, informatics highly sought after, need to hire
 - New hire should be interdisciplinary and broadly involved with campus that can serve any number of departments
 - 80% is digital but nobody knows about it the collection, even when it's interdisciplinary
 - Will this be promoted publicly? Or cross-referenced with other archival materials
 - Could this become more of a state repository rather than a university archive?
 - We need to have a relationship with the community
 - Create a database or website to run tools on top of what's there
 - Institution needs to help structure the questions regarding the data so that it becomes more available
 - Need for student training
 - Toby Pell and Clay Pell have URI connections
- What was this new center look like?
 - Guide for textual data
 - Monthly or bi-weekly consulting meetings as part of the center across disciplines, think-tank
 - Collect data on faculty and staff, race, gender, would this also be collected here?
 - Student who created database from URI Confessions learned more than in classroom time
 - Center for Digital Humanities?
 - Must remember regulations and sensitivity of certain data especially when analyzing data

Notable Issues, Questions, Challenges:

The most notable issue of this session involves the need for cross-disciplinary collaboration at the proposed center for high performance and computation.

How can big data be made accessible? The university must have a role in getting information out to the public.

Think about collection of data that isn't just in the form of a "hard copy" like the Pell Collection, but digital data that can disappear. We must have a broad view of what's worth saving.

Title of Session: Applications and Implications of Big Data

Faculty Facilitator(s) Joan Peckham and Julia Lovett

Note taker name Abigail Casavant

Check one:

Morning Breakout 1 Morning breakout 2

Afternoon breakout Afternoon breakout 2

Questions from the program: What does URI need in order to move forward with utilizing big data in the future? Discussion will focus on lessons and applications for the future of digital humanities, business and government arenas. Issues to consider are the acquisition of resources and encouraging collaboration across the disciplines and the steps to facilitate participation in big data projects.

Ideas proposed/Recommendations

Video Summary: Digitizing “Republic of Letters,” add life to lost world, map not the whole but selected case studies to represent whole, big data

- Data enabled research and education, research and computing center
- What sorts of resources would you need and what relationships could this create? How would this serve you and your department?
- Is everybody focusing on their own work and not working dynamically – go this office for this, and no one is talking to each other
- Big data can be centralized; do we need to spend our resources on budget, administration, HR when there is a need in the academics?
- Yale has 12 customizations in administration system, URI has over 2,000 – problem here
- Research office, Library, and IT are critical to the new center
- Can’t just be about software and hardware
- People to help faculty and researchers understand software, funding for re-training especially for technology when certain things become obsolete
- Education is critical here
- Do we depend on Brown’s facilities for data management, database, etc.?
- Text-mining and creation of intellectual property, problems with purchasing intellectual property
- What does faculty need to support big data?
 - Actual digitizers
 - Expertise to tailor software instead of creating own
 - University should offer support to realize this idea
 - Need to know what resources are available and accessible
 - Interdisciplinary support staff
 - A center for digital humanities within the center high performance computation center
 - How do we collaborate and get to know each other across departments?
 - Learning data and performance data into algorithm and create prediction models, is this compliant? IRB
 - Legal expertise, copyright restrictions and compliance for data management
 - Multidisciplinary across the university, the archives/library is the perfect place for this

Notable Issues, Questions, Challenges:

The most notable issue of this session involves the need for cross-disciplinary collaboration at the proposed center for high performance and computation.

A big challenge would be to make sure all disciplines are represented at the center and to make sure faculty, staff, and students are educated about big data and the goals of the center.

Must remember compliance and copyright issues when it comes to data collection and management especially at this scale.

Do we rely too much on other universities' models of databases and database management?