



Thoughts on “eResearch”: a Scientist’s Perspective

Amanda L. Whitmire

Oregon State University, Corvallis, OR, USA

Abstract

In response to the burgeoning practice of collaborative, networked, data-intensive research (known as eScience), university and research libraries are devoting significant consideration, effort and resources toward expanding their responsibilities to include research data services. The jargon that the librarianship community uses to discuss data-driven research is inconsistent and confusing, especially to non-librarians. This is problematic because when we attempt to engage research scientists in an effort to provide services, we risk alienating our

potential stakeholders by using language that they don’t understand. As a recent transplant to the library community, the difference between librarian and research scientist perceptions of data-driven research, and the vocabulary surrounding it, have been surprising. This paper summarizes the problem of “eResearch,” spoken from the perspective of a recent scientist-turned-data librarian. The main conclusions reached are that “eResearch” is a meaningless term that should be avoided, and that data support services needn’t be couched as an eScience issue.

A Stranger in a Strange Land

I recently moved into the library profession from a post-doctoral position in a data-intensive research field. I assumed that my confusion about the terms “eScience” and “eResearch” were cultural, and that the words must be standard library jargon (as all disciplines have), terminology with which I was trying to familiarize myself. While the jargon can be an impediment, I believe that my confusion about these terms is indicative of a disconnect between the perspectives of librarians and that of the researchers we are working to support. This disconnect is exemplified by the following observation: I was heavily engaged in eResearch for the last 15 years, but in all of that time, I had never

called it that. In fact, I had never even heard the term “eResearch” until I started my new job in a library.

“EResearch” vs. “eScience”

I have to admit that one of my first thoughts after starting my new position was, “What do they mean, ‘eResearch’ and ‘eScience’?” I searched for definitions in traditional resources, the Oxford English Dictionary for example, but came up empty handed because these terms are too new to have been formally defined. So, I took an etymological approach: the term “eScience” originated in 1999, and is attributed to John Taylor, the Director General of the United Kingdom’s Office of Science and Technology

Correspondence to Amanda L. Whitmire: amanda.whitmire@oregonstate.edu

Keywords: eResearch, eScience, jargon, research data services, data management

(Marcum and George 2010). Taylor's definition of eScience was, "large scale science that will increasingly be carried out through distributed global collaborations enabled by the Internet. Typically, a feature of such collaborative scientific enterprises is that they will require access to very large data collections, very large scale computing resources and high performance visualisation back to the individual user scientists" (University of Edinburgh 2011). A practical example of eScience is global climate change prediction modeling. Climate research on this scale involves assimilating huge datasets from many sources and running multiple iterations or scenarios on supercomputers or gridded machines. Wikipedia defines eScience as "computationally intensive science that is carried out in highly distributed network environments, or science that uses immense data sets that require grid computing; the term sometimes includes technologies that enable distributed collaboration" ("E-Science" 2012). In a general sense, we can think of eScience as research involving immense datasets that require large-scale computing resources to facilitate analyses. It's important to recognize that eScience is not a discipline in and of itself, and is not limited to pure science disciplines; it is a research methodology. As such, eScience is being practiced by researchers in fields ranging from particle physics to molecular genetics to holocaust research ("European Holocaust Research Infrastructure" 2013).

Finding a single definition of eResearch proved challenging. EResearch has been defined as being synonymous with cyber-infrastructure (Borgman 2006), that is, "technologies to facilitate distributed, collaborative, information-intensive forms of research and learning" (Borgman 2007). In the social sciences however, the term eResearch is used in a manner akin to eScience, in that it describes a research methodology. Meyer and Schroeder succinctly describe eResearch as "the use of digital tools and data ... for the distributed and collaborative production of knowledge" (2008). This defi-

nition places no limits on research discipline, and the volume of data involved has no bearing on whether or not the work qualifies as eResearch. In the context of the library community, the Association of Research Libraries (ARL) defines eResearch as "computationally intensive, large-scale, networked and collaborative forms of research and scholarship across all disciplines, including all of the natural and physical sciences, related applied and technological disciplines, biomedicine, social science and the digital humanities" ("E-Research, Association of Research Libraries" 2013). I wonder if ARL adopted the term eResearch instead of eScience as a practical means to broaden the perceived disciplinary scope of this focus-area topic (despite the fact that eScience is transdisciplinary). The definition is somewhat confusing because a 2010 ARL report defines eScience in a very similar way: "e-science is defined broadly not only as big computational science, but also team science and networked science. It includes all scientific domains, as well as biomedicine and social sciences that share research approaches with the sciences" (Soehner, Steeves, and Ward 2010). From the ARL perspective, what is the difference between eResearch and eScience? Soehner et al. acknowledge the difficult nature of these terms by saying, "The vocabulary of eScience is still evolving" (2010). Indeed, jargon is not a new problem for libraries (Naismith and Stein 1989).

One difficulty associated with jargon is inconsistent usage. In my experience so far, I have seen the term eResearch applied much more broadly than its ARL definition. Most often, I hear it used to indicate, "the use of information technology to support existing and new forms of scholarly research in all academic disciplines ... encompass[ing] computational and e-science, cyberinfrastructure and data curation ... usually data intensive, but the concept also *includes research performed digitally at any scale*" (emphasis added; E-Science Institute 2012a). Whether or not this definition con-

curs with the ARL definition of eResearch, it seems to reflect the more common usage of the term. Practically speaking, the term “eResearch” is regularly used to refer to anything on the wide spectrum from straightforward digital research all the way up to eScience. But what does eResearch actually mean?

Researching the etymology of eScience and eResearch helped me to understand something about these terms. First, the ARL definitions conflate eScience with eResearch, which can be seen as an error to some. Second, and more importantly, widespread usage of the term eResearch in the librarianship community is far broader than the scope of the ARL definition. It is commonly used to indicate any form of digital research. Despite the intent to define “eResearch” as something innovative and deserving of special consideration, the definition is effectively rendered meaningless by its lack of specificity in the community. If a new word is created and is then used to refer to the majority of modern-day research, does that word have any value or does it just obfuscate the conversation?

While my understanding of “eScience” vs. “eResearch” was clarified, I was still confused by the way the library community was talking about eResearch. As a part of my library’s effort to learn about the types of data being collected on campus, we had to develop interview questions for faculty regarding their research. As I was reading through some sample questions: “Do you generally see eScience as a benefit to your work, or another burden to worry about? Do you have eScience commitments imposed on you from your own broad research community?” (E-Science Institute 2012b) etc., I had a revelation: librarians call it “eResearch” and “eScience;” scientists call it “research.” The two groups interpret and label the same activity very differently. Having to imagine using those questions in faculty interviews clarified the issue: there was no way to communicate effectively with scientists about

their work if I used terminology that they didn’t understand. The term eResearch is foreign to most scientists, as it was to me before I worked in the library.

I wondered if perhaps it was just me who felt this disconnect between library and data-driven science perceptions of “eResearch,” so I contacted three colleagues in my previous research field and asked them if they considered their work to be eResearch. As you read their responses, keep in mind that all of these scientists are solely engaged in the ARL version of eResearch. Their answers included:

“Never heard of eResearch ... Sounds like something done solely over the Internet and perhaps about the Internet.”

“It was very tempting to look up what eResearch means ... but I didn’t. If I were to guess, I would think that eResearch means that some or part of one’s data comes from web-based resources.”

“EResearch sounds like data mining to me but I don’t really know what they mean. I found the term a bit off-putting.”

I realized that as librarians we need to be careful about how we engage research scientists in our efforts to support them in data management and curation. It is the case in most science disciplines, and has been for many years, that the research they do, data-driven or not, is just called “research.” If we frame all of our conversations with them around “eResearch” and “eScience,” we risk looking antiquated and out of touch while our intent is to be anything but. If we are creating a false dichotomy between “eResearch” and “research,” we are setting ourselves behind before we even get the conversation started.

Moving Away From “eResearch”

From the library perspective, the “e-” distinction makes sense. Librarians have been stewards of non-digital information for thousands of years, going all the way back to Al-

exandria. The expansion of our role to include digital resources and data is a substantial evolution, and one that requires significant and thoughtful effort. What we have to keep in mind is that research practices that seem state-of-the-art today will soon become common course. The question about a semantic transition from “eResearch” back to “research” is not “if,” but “when” (excluding, of course, the communities that never adopted the terminology in the first place).

In looking at current capabilities and future trends in research data services (RDS) at academic libraries, Tenopir et al. (2012) found that up to a third of all academic libraries (within the Association of College and Research Libraries) are planning to offer research data services in the next two years, mostly in the realm of providing guides on how to locate and cite datasets, and in providing help with data management plans. Neither of these support activities is strictly related to eScience or eResearch. Considering the Tenopir et al. findings, the community emphasis on the role of libraries and librarians in eResearch and eScience seems misplaced; research data services benefit patrons who produce digital data of any size or type (and in reality, practitioners of “small science” are in the large majority (Staff 2011) and need us the most). Libraries are positioning themselves to be able to provide support for, and training in, data management (planning, storage and backup, metadata, sharing and reuse, ethics, and preservation). From my perspective, these services aren’t strictly for, or necessitated by, eScience or eResearch; they are the result of broadening the scope of what libraries do to support the evolving nature of academic research and scholarly communication. In the same way that we work with students to improve their information literacy, we can engage scientists by helping them better navigate the research process and integrate data management best practices where appropriate. If we approach helping scientists with data as an “eResearch” issue, we risk

losing our audience because of jargon that has no meaning to them.

As information literacy experts, librarians are better positioned than most to understand the power of words. That being the case, I urge the librarianship community, of which I am proud to now be a part, to think critically about the emphasis being placed on “eResearch,” and what that means. The differences between eScience and eResearch are clear and important. The term eScience defines a new method of research that is data-driven (as opposed to being hypothesis driven), collaborative and very computationally intensive. eResearch is a word with a definition that is so general that it’s difficult to know when to use it appropriately. In cases where I want to distinguish a field of study or specific research topic as using a lot of data, I refer to the field or the research as being data-intensive. Thus this term isn’t perfect, but it seems less confusing.

The scientist in me is always inclined to define, categorize, and sort. In the case of eResearch however, the attempt to differentiate it from other types of digital research has been a distraction when I’m working with researchers to determine how I can help them better manage, preserve, and share their hard-earned body of work. While the headline-grabbing “data deluge” and subsequent evolution of a “fourth paradigm” of research may be responsible for getting data onto our radar, the scope and impact of burgeoning library research data services far exceed the realm of eScience. Framing what I do as support for eResearch or eScience makes it difficult for me to effectively communicate my role to the research community that I am trying to support.

References

"About EHRI: Connecting Collections." *European Holocaust Research Infrastructure*. Accessed February 11, 2013. <http://www.ehri-project.eu/drupal/about-ehri>.

Association of Research Libraries. "E-Research." Accessed April 26, 2013. <http://www.arl.org/focus-areas/e-research>.

Borgman, Christine L. "What Can Studies of e-Learning Teach Us About Collaboration in e-Research? Some Findings from Digital Library Studies." *Computer Supported Cooperative Work* 15, no. 4 (2006): 359–383. <http://www.escholarship.org/uc/item/6cj1756r>.

———. *Scholarship in the Digital Age: Information, Infrastructure, and the Internet*. Cambridge, Mass: MIT Press, 2007.

"E-Science." *Wikipedia, the Free Encyclopedia*. <http://en.wikipedia.org/w/index.php?title=E-Science&oldid=490989467>.

E-Science Institute. "Key Concepts and Terms." Duraspace, 2012.

———. "E-Science Institute Overview." *Duraspace*, 2012. <http://duraspace.org/e-science-institute>.

Marcum, Deanna B. and Gerald George, eds. *The Data Deluge: Can Libraries Cope with e-Science?* Santa Barbara: Libraries Unlimited, 2010.

Meyer, Eric T. and Ralph Schroeder. "The World Wide Web of Research and Access to Knowledge." *Journal of Knowledge Management Research and Practice* 7, no. 3 (2009): 218-233. <http://papers.ssrn.com/abstract=1153922>.

Naismith, Rachael and Joan Stein. "Library Jargon: Student Comprehension of Technical Language Used by Librarians." *College & Research Libraries* 50, no. 5 (1989): 543–

552. <http://crl.acrl.org/content/50/5/543>.

Soehner, Catherine, Catherine Steeves, and Jennifer Ward. "E-Science and Data Support Services: A Study of ARL Member Institutions." Association of Research Libraries, 2010. http://www.arl.org/bm~doc/escience_report2010.pdf.

Science Staff. "Challenges and Opportunities." *Science* 331, no. 6018 (2011): 692–693, <http://dx.doi.org/10.1126/science.331.6018.692>

National e-Science Centre. "Defining e-Science." 2011. <http://www.nesc.ac.uk/nesc/define.html>.

Disclosure: The author reports no conflicts of interest.

All content in *Journal of eScience Librarianship*, unless otherwise noted, is licensed under a Creative Commons Attribution 4.0 International License.

<http://creativecommons.org/licenses/by/4.0>

ISSN 2161-3974