eScience Software Development: The Beginnings of a Theoretical Understanding

Jacob Lopez, Devin Velasco, & Kerk Kee
Diffusion of Innovations (HCOM 597), School of Communication
Chapman University, Orange, CA

Introduction

Cyberinfrastructure, otherwise known as eScience, are considered "large-scale distributed scientific enterprises" (p. 1) that are composed of (and supported by) human and technological infrastructures (Lee, Dourish, & Mark, 2006). eScience projects are often characterized by cooperative efforts to conduct computer-mediated collaboration on scientific research and interdisciplinary scientific practice (Lee, Dourish, & Mark, 2006). They include computer hardware and software, virtual environments, virtual organizations, and virtual systems that have the capacity to hold vast amounts of information to increase productivity, innovation and collaborative research (Kee, Cradduck, Blodgett, & Olwan, 2011).

eScience is a modern extension to the modern phenomenon of "big science", yet its extension into scientific developments is still in its early stages, and many of these projects are primarily focused on efficient development (Lee, Dourish, & Mark, 2006). While practitioners may dwell on procedural limitations to development, such as funding, licensing, and organizational policies, the researchers—as communication scholars—would argue that interactions form the theoretical basis for the concept of development of new cyberinfrastructure. Specifically, we employ two lenses to guide our research: diffusion of innovations theory (Rogers, 2003) and Actor-Network theory (Latour, 2005).

Research Questions

1. How is new eScience software development shaped by person-person interactions?
2. How is new eScience software development shaped by person-object interactions?
3. How is new eScience software development shaped by object-object interactions?

Method

The researchers had access to the transcripts of 48 interviews of people involved in eScience. These original transcripts were conducted by the third author and his student team (www.octgroup.org). Interviews were conducted either in person, via telephone, or both. Participants included administrators, astrophysicists, biochemists, computational chemists, computational environmental scientists, computational molecular biologists, mathematicians, physical chemists, physicists, and technologists in various states throughout the United States. Using a grounded theory approach, the researchers first conducted an initial read through of the transcripts to form the research questions. The researchers found the most appropriate method of analysis would be to investigate instances of person-person, person-object, and object-object interactions as they relate to creating the idea of development. All reading, coding, and work with the data was conducted on two separate NVivo data files that were later merged to conduct the analysis.

Results

Table 1. Summary of theme analysis.

How is new eScience software development shaped by person-person interactions?

Theme	Description	Example
Face-to-face interactions for brainstorming	Developers have found that interpersonal interactions work better for launching new projects because of the allowance for the free flow of ideas. Computer-mediated modes of communication are effective for direct collaboration.	*"…often times, there's a lot of brainstorming that goes on at a conference like that, so when I sit in on workshops, tutorials, presentations, it stimulates creativity. And so it's a catalyst for interactions and exchanges."*
Spontaneous collaboration	Virtual organizing was described as impromptu attempts at fixing social/scientific problems	*"Nobody said in the beginning, I'm going to develop software, I'm going to create a virtual organization. It typically begins small and it depends on the adoption…"*
Management by influence	Many actors of cyberinfrastructure are highly-skilled individuals. Managing such a team requires influence more than other traditional management techniques	*"So what's important for me in my role is not that I can make somebody do something but I have to influence them such that they are convinced that it's the right thing to do and they do it. And that's a very different way of managing things- I can't say go do this. And to some degree I can, but an organization like this they can say whatever, I don't think that's the right thing to do."*

How is new eScience software development shaped by person-object interactions?

Theme	Description	Example
Iterative debugging	Developers should keep an open discussion with users in attempts to address their software issues or improve other software	*"…as you're using it and you're developing it, you're constantly making it better… debugging and finding what's wrong with it. And then essentially re-running it over and over until you get something that looks right. Using the tool, the software, whatever it may be until it breaks, until you find a new way that broke it, a use case that you can't think of but you finally hit in the field."*
Avoiding continuity errors when coding	Creating software involved multiple hands working on the same project, which can lead to development errors	*"…no one should be working on the same aspect of your software or whatever it may be… in a group, a lot of what you should do is in parallel but on different pieces. And communicate very often… I don't know, [if] it's going to crunch something."*
User indifference and needs	Users of e-science technologies expressed a need for cyberinfrastructure, yet were unconcerned about the prospects of future developments	*"The scientists, I feel like they really want something that's familiar and easy to use…They tend to not really want to use something different, unless you can really make it happen for them."*

How is new eScience software development shaped by object-object interactions?

Theme	Description	Example
Importance of interoperability	Developers believed software developments needed to be able to work with existing developments	*"You don't want to build something in horrible software that it lives in its own environment and whatever you build in there is just kind of stuck in there."*
Spontaneous collaboration	Interdisciplinary research can hinder developments in cyberinfrastructure	*"... if you bring some social science data set, and physicist brings ah some physics observational data set, they're going to be very different."*

Findings

As shown in Table1, the researchers came to an understanding of new eScience software as a continuous process of creation and implementation shaped by person-person, person-object, and object-object interactions. Not only is development of new eScience software shaped by these interactions, development itself is a series of these interactions continuously producing creation and implementation of an eScience software. Aspects of person-person interactions that directly shape development include face-to-face interactions for brainstorming, spontaneous collaboration, and management by influence. Aspects of person-object interactions that directly shape development include iterative debugging, avoiding continuity errors when coding, and user indifference and needs. Finally, aspects of object-object interactions that directly shape development include importance of interoperability and issues with sharing and merging data.

Theoretical Implications

Many participants described the development and implementation of an e-science innovation to be mutually dependent and simultaneous. Some participants even went so far as to say the two are indistinguishable. The researchers argue that the difficulty in delineating development from implementation (and vice-versa) of e-science innovations will have major implications for future diffusion research of cyberinfrastructure. More specifically, the innovation-decision process (as outlined by the diffusion of innovations theory) may not be sufficient in describing the diffusion of cyberinfrastructure. According to this process, an individual passes five stages when deciding to adopt or reject an innovation: knowledge, persuasion, decision, implementation, and finally confirmation (Rogers, 2003). The results of this study suggest that many of the e-science innovations that are released are done so even when the innovation itself is not yet complete. In fact, the idea of a finished product does not apply to e-science innovations because there is an unwritten agreement between users and developers to work together to continue improving and reinventing. Therefore, the innovative-decision process, or the time is takes to pass through all five stages of the process (Rogers, 2003), may continue on indefinitely as users and developers cycle through the decision, implementation, and confirmation stages. Therefore, the researchers contend that the findings of this study be considered in the diffusion of innovation theory. Investigations into cyberinfrastructure diffusion can be more productive when taking into account the instances where developers and users of an innovation are in a development-implementation cycle.

Acknowledgements

This material is based upon work supported by the National Science Foundation under Grant No. 1322305. This poster was presented on May 8 at the 2019 Chapman University Student Research Day, Orange, CA.