

Trevor Croker, Dissertation Proposal
Summary of Key Issues and Dilemmas

Working Title — Grounding the Cloud: The Geography and History of Ubiquitous Computing

Summary — In this dissertation project, I plan to look historically at the development of ubiquitous computing services with an emphasis on the importance of physical geography in the construction of the notion of “cloud computing.” By tracing the historical trajectory of the term “cloud computing,” I hope to highlight the ways in which physical infrastructure remains a critical site of examination for those interested in Internet security, privacy, and social responsibility. A rhetorical decoupling of physical infrastructure from digital spaces, I argue, is a potentially dangerous move that ultimately makes shaping the direction of digital technologies more difficult for a broad range of actors. In order to interject, or to move geography back into the discussion, I plan to use a number of case studies that speak to the times and places in which the importance of space becomes apparent. As a whole, this dissertation hopes to add to the historiography of Internet historians by utilizing the interdisciplinary methods of Science and Technology studies, as well as the overlapping History of Technology scholarship.

Key Questions —

1. What are the historical origins of cloud computing? How did the term emerge? Who/What were the key actors that invented and popularized the term? How has the term changed over time?
2. How has the role of physical space changed over time? Has geography been less pronounced with the growth of certain technological systems? Where does the mainframe/terminal fit into the history of cloud computing?
3. For potential case studies or historical research, what are particular instances in the history of computing in which physical geography has become central to the story?

Methodological Approaches — This project is broken into two parts, one relies upon a more traditional historical approach of explaining the actors and events that led to the rise of cloud computing. The second is rooted in the methods of Science and Technology Studies (STS) in order to explain the geographic case studies.

Chapter Outline —

- I. Introduction: Cloud Computing and Literature
 - a. Literature Review [*In Progress*]
- II. Historical Origins of Cloud Computing
 - a. Creation/Early history of the term [*In Progress*]
 - b. Development of the term [*In Progress*]
- III. Cloud Meets The Ground: Geography Case Studies
 - a. Case Study 1: 2008 Submarine Cable Outage in the Middle East [*Completed*]
 - b. Case Study 2: Security of SSL Certificates [*In Progress*]
 - c. Case Study 3/4/5: TBD
- IV. Conclusions

Research Proposal: The Internet's Materiality, Clouds, and Technological Failures

Introduction

The following is a dissertation research proposal that looks historically at the changing definition and importance of materiality in cyberspace. The point of departure for this research is the assumption that the Internet and related information technologies have radically altered how individuals understand and interact with physical and digital objects in everyday life. For many individuals in the developed and developing world, the experience of work or play is mediated through computers and other digitized equipment. Documents, photos, videos, music recordings, which once were all stored on physical media in discrete locations, now have been digitized and are commonly spread across different digital ecosystems. The emergence of the term “cloud computing,” particularly in the last five years, has signaled a move towards computing that is seemingly not rooted in a particular physical space. The ability to access digital content regardless of your physical location, on a wide set of media devices, is a common image in technology communities that view the move towards the cloud computing as historically inevitable. Popular consumer products such as document repositories (DropBox), streaming music services (Pandora), webmail (Gmail), and streaming video (Netflix) promote this vision of distributed computing (not linked to a specific geographic location). The historical trajectory, at first glance, supports this vision of computing.

It is exactly this rhetorical framework of inevitable cloud computing that I wish to examine in my research project. I want to challenge this technological deterministic attitude towards modern computing through an examination of the historical roots of distributed computing and the contemporary concerns that certain communities have towards this new age of “non-materiality.” The idea of having an information technology system that does not have an underlying material reality is simply impossible. This is not to say that computing cannot be discussed as being distributed geographically. Instead, my research hopes to demonstrate that the physical copper, concrete, and fiber that moves data across the Internet is essential to cloud computing. To not acknowledge the material reality would be a major oversight.

Therefore, my research is motivated by the following preliminary questions: When did the concept of cloud computing emerge and what have been the major historical events that have led computing developments to move towards larger, more abstracted, technological systems? What are the political, social, and economic implications of moving a distributed computing environment? Is the

material base of the Internet being ignored by technological innovators, and if so, what are the implications of this oversight? Are there communities that reject this move, and hold onto a notion of geographically bounded computing? Are there instances of failure in cloud computing and, if so, what can these failures tell us about the long term impacts of these types of systems? Who are the individuals using these systems, why do they use them, and do questions of materiality matter to them? What are the technological specifications of the information infrastructure? Was the design of these technological systems built with the intention of cloud computing?

Prior to Case Study

In my search for research locations, I have attempted to select locations of study that pull apart the main tensions in the discussions about cloud computing. One of the primary methods through which I wish to uncover the social and political implications of network architecture is through an examination of technological systems. In the literature section of my proposal I discuss the work in the history of technology and STS that look at technological systems as holding certain political values. Using this scholarship, I am interested in doing research on how technological systems in the digital age are embedded with particular political values.

One way in which the politics of systems become exposed is in times of crisis or technological controversies. STS has produced much scholarship that looks at how black boxes become opened when controversies arise. Therefore, when looking at potential research locations, I have attempted to find controversies in digital technologies. Digital technologies, such as cloud computing, have created unique challenges because content is more “fluid” than physical media and is able to be easily lost or stolen from an individual user.

Prior to addressing the specific cases, I want to give a broad overview of my research plans and how I plan to structure my research. Instead of focusing on a single event, I have tried to select a few major cases to highlight the materiality of the Internet and the move to a cloud-based digital ecosystem. In an attempt to set up these events, I plan on looking at the origins of cloud computing. Although I will be unable to dedicate a lot of space to the historical origins of cloud computing, it will be necessary to talk about the ways in which cloud computing came into being and the important actors in pushing for particular types of systems. The history of cloud computing goes back to the 1960s, so the challenge will be identifying the most critical moment of cloud computing. Until I do further research on the origins, I will be unable to go into much depth on this issue.

One Case

In trying to study the materiality of the Internet, I have been interested in cloud computing services that demonstrate the links between geographic location, physical hardware, and end-users. Therefore, I plan on looking at data management practices and how they have changed as cloud computing has become more commonplace. Data management practices are simply the different ways in which individuals and organizations store, preserve, and use digital data. My research goal is to look at how these individuals and organizations have managed their data historically and how cloud stores have changed how actors manage their data today. It is difficult, at this point in my research, to narrow down a specific research location but I will present a rough outline of what I anticipate looking into.

One of the most compelling and relevant locations for study is the online data storage market. Prior to the invention of high speed Internet, for most people the option to store documents, photos, videos, or other data in the cloud would have been simply too expensive and impractical. Today, however, many users have some type of online backup service that they use to keep some of their digital property in the cloud. There are a number of online backup services and other data management services that operate to store information in the cloud. Whether it is an online email service like Gmail, or an online file backup service like DropBox, these services are an increasingly common way for everyday users to keep their information in the cloud.

I plan to pick a small set of these online services and examine how they operate. In my search for a specific service, I will be looking for technological systems that have failed in the past and how these failures have disrupted a particular vision of technological advancement. At the moment, my current research location is Amazon's Elastic Compute Cloud (Amazon EC2) which is an array of distributed computing power that is commonly used for hosting webpages and user content. Amazon is traditionally known for its online retail presence, but its EC2 service has quickly become one of the most used cloud computing services for webhosting and data backup. I will need to do further investigation as to whether this is a viable research location for a long term project.

Additionally, I would like to contrast the case of Amazon with another online backup service that has recently generated controversy due to the alleged involvement with copyright file distribution. The online file service, Megaupload.com, was shut down for hosting and distributing copyrighted files. This cloud computing system differs from Amazon because it is a more centralized system that does not have the same scaling ability as Amazon's EC2. Additionally, the differences in users and business practices put Megaupload under closer scrutiny. Using these two cases, I can look into some of the legal and political challenges of cloud computing. These two cases, however, could change depending on the

amount of depth each case offers. Once I have selected the additional cases, I will then try to link them up to the secondary case: the physical network geography of the undersea cables.

Additional Cases

A second case that I will be looking at is the history of Internet-infrastructure disruptions due to accidental or purposeful destruction of Internet-infrastructure. There are, broadly speaking, two major ways in which network traffic gets disrupted: either through a technical problem in the software or the hardware. It is impossible to discuss either of these disruptions as being separate from one another. I am interested in looking at these network disruptions and the way in which these failures are discussed. Therefore, for my research I plan on looking at one set of network infrastructures in particular — undersea data cables.

Underwater network cables are commonly referred to as “submarine communication cables” and are specially designed cables that are used to transmit data across long distances under the ocean or large bodies of water. In my research, I want to look at the history of these cables and the associated challenges that engineers faced when designing them. I will look specifically at the Middle East Internet crisis that occurred in 2008 where numerous submarine cables were broken. The existence of these cables is rarely discussed in terms of the cloud, but they are essential to the continual operation of distributed computing. There have been a few instances of these cables being accidentally cut (primarily by fishing boats or other marine equipment) and these breakages have caused massive Internet disruptions. Additionally, these underwater cables actually reach land at a limited number of locations. Therefore, any damage to these locations could cause major issues.

Using this case, I want to link the underwater cables to the online backup services that I discussed previously. I am interested in the implications of discussing online backup in terms of the rhetoric of “any-time access,” when the physical geography is ignored. What exactly are the implications of ignoring the physical reality that supports the Internet and cloud computing?

Literature and Proposed Theoretical Background

Although my project intends to be rooted in core concepts from STS, the nature of the research requires a wide variety of interdisciplinary literature research. In this section, I will describe the main theoretical traditions that I will be drawing upon. Although this list is far from comprehensive, it attempts to give a rough approximation of the relevant academic literature and the historical arguments that will inform my own research. The primary bodies of work that I will be looking at are texts from the

history of technology, digital law and humanities, the politics of space and representation in political science and architecture, and the politics of knowledge in STS.

Looking first at the history of technology, this discipline will be important for various reasons. The historians of technology, although the field is relatively small, have produced a great deal of work on technological systems. One theorist that is central to the field is David Nye and his work on technological momentum and the technological sublime. In terms of my project, both of these concepts are important because they help explain how visions of cloud computing appeal to a certain notion of the sublime, and, in moving towards that vision, produce a certain amount of technological momentum. The Internet, for as much as it is discussed as being flexible, is rooted in physical hardware that is not always easily modified (high momentum). Nye's work will likely be one of my main theoretical frameworks that I utilize in my paper.

Another historian of technology (and STS scholar) is Wiebe Bijker. Bijker is useful because he discusses both the embedded politics in artifacts, as well as offers a theory of technological change. I will look at his book *Of Bicycles, Bakelites, and Bulbs* to see if there are any other relevant theories that I should use in my work. In sum, I wish to use the history of technology to demonstrate the process by which Internet systems are constructed around different notions of progress and technological advancement.

One of the major bodies of literature that I will need to address, but have had little exposure to inside STS, is the core texts in digital law and humanities studies. Most of these texts have been produced in the past two decades due to the relative newness of the Internet and popular computing. These texts are particularly useful for my research because they are often concerned with many of the same problems that I am dealing with. As opposed to the scholars in my other literature categories, the writers interested in the digital humanities are not always from academia. Many of the texts are written by journalists or cultural commentators that have been steeped in the digital universe. For these reasons, I need to be cautious when I use these sources in my research. The polemical goals of some of these texts are useful because they raise major debates in the field, but they also might obscure less biased accounts of the story.

Although I have not read all of the relevant texts at this point, there are a few key texts in the digital studies realm that I plan to incorporate into my research. The first author that I will look at is a law professor named Lawrence Lessig. His books include, *Code and Other Laws of Cyberspace*, *The Future of Ideas*, *Free Culture*, *Code*, and *Remix*. These are essential texts to a project focused on the

Internet because they have been the foundation for many contemporary texts on the Internet. Lessig's concept of "code as law," suggests that computer code acts in a similar manner as other regulatory devices (such as the law). When looking at technological systems, this concept can help me immensely.

Other core texts include Jack Goldsmith and Tim Wu's *Who Controls the Internet* about Internet governance, Daniel Solove's *The Digital Person* about privacy, Siva Vaidhyanathan's *The Googlization of Everything* regarding cloud computing, and Julie Cohen's *The Networked Self* about digital rights management and privacy. All of these books share a concern with the challenges of managing data and people in the era of digital media. Certain texts, such as Vaidhyanathan's previously mentioned book or *The Net Delusion* by Evgeny Morozov, take a more cynical view towards the freedom of expression on the Internet. A challenge that I will face, because these books often deal with contemporary concerns, is how to put these texts in their historical place and not generalize from them. Lessig's texts generally have more staying power because they are not always rooted in the specifics of the Internet, but are more about the general problems that have arisen. These texts, however, often lack a more stable theoretical lens, so I will use the other categories to support these more contemporary texts.

Another set of literature that will be important for my study are some authors from political science, architecture, and geography. Because I intend to discuss the Internet in terms of both digital and physical space, I will need a vocabulary to discuss these issues. Geography is a discipline that is of immediate use when it comes to talking about space, but the history of geography has typically placed focus on physical space, rather than addressing the interplay between digital and "real" spaces. Still, certain geographers, such as David Harvey, are important figures because they have been successful in mapping power relations upon different spaces. I want to use geography texts because they can help me emphasize how material objects and physical spaces are hidden (or revealed) when digital technologies are discussed.

Of course, the discourse of space is not limited to geography. Political science works have also addressed the topic. People like Alexander Galloway and Jane Benitt look at the importance of materiality in systems. Other scholars who have looked at media ecologies (the Internet could be considered an elaborate media ecology) are Friedrich Kittler and Ron Deibert. The goal of using theorists such as these is to bring in literature that focuses on power as it relates to the use of space. Most of these academics are not talking about digital spaces, but their work will be useful nonetheless.

Finally, I will be looking at the literature in STS around the politics of knowledge. The Internet is a different space than the scientific laboratory (on which STS scholars have spent much time), but I

believe the ideas from classic STS texts can be adapted for my project. Harry Collin's and David Bloor's classic works on knowledge offer a general overview of sites where meaning is contested. I have not read Chandra Mukerji's "Towards a Sociology of Material Culture," but I imagine the focus on the material reality of the lab could help me when I try to stress the importance of looking at the physical Internet. Steven Epstein's work on lay expertise could be useful if my research begins to move towards a study of activists who resist cloud computing. I also need to do further research on "technoscientific governance" to see if that set of literature would help me (an example of this would be Alondra Nelson's "Bio Science").

My plan is to use these four categories (history of technology, digital law/humanities, political studies of space, and STS's use of the politics of knowledge) to frame and inform my research. I believe that there is sufficient theoretical depth in these categories to allow me to do research that is rooted in the traditions of materiality, but with enough flexibility to allow for a new take on materiality in cyberspace.

Design and Methods

The methodological approach that I will be using in my research is one that relies on primary source technical documents, news stories on cloud services, and an investigation of network architecture. I will use a variety of research methods to access the information that I will need for my project.

First, there is the issue of network infrastructure resources. There are a number of barriers that I might have to deal with when trying to access the specifics of how cloud computing services operate. The most reliable and easily accessed documents on the topic will likely come from academic journals and researchers who are designing or describing these systems for academic publication. These resources could be fairly easy to access due to their intended public nature. Private cloud systems, however, might be more difficult to access because of their business connections. Not being able to understand how the network systems are structured could be a major barrier in my project. It is unlikely that a company like Amazon would allow me to see their own network structure, but I may be able to discuss similar systems if I am unable to access the details of the specific proprietary network architecture. Instead, I can use similar models for cloud computing that are more accessible.

The biggest challenge that I think I will have address in my project is historical examination of cloud computing. This terms is central to my study because I am interested in the reasons behind the decision to employ the use of this term. Over time I will try to determine to what extent I can delve into

the history of this term. For now, I will continue doing research on the topic to see if I can find a repository of documents for this concept.

I have attempted to pick a small number of cases, starting with underwater cables and cloud storage to limit the scope of my research. Even with a smaller number of cases, the scale of the project will be difficult to manage. One issue is that the topic of cloud computing is linked to many different social issues. Economic, social, and political concerns are all relevant categories and it will be a challenge for me to narrow my focus to something manageable. Therefore, to address this problem, I intend to focus on looking at the implications of digital materiality by looking at the problems of not addressing the physical dimensions of information technologies.

I will have to utilize various collection methods in my research. The first, and likely easiest place for data collection, will be the use of the Internet in searching for news stories and other material on cloud services. These texts should help me estimate the general feelings that individuals have about cloud services and the problems that have arisen from cloud service failings. The second source of data will be the use of academic journal searches to find documents on the technical details of network infrastructure. These, like webpages, are easily accessed and should not require any extra level of work.

The more difficult portion of my collection methods will be gaining access to individuals or documents that talk about specific cloud services or data equipment (particularly those used for monetary or safety purposes). For instance, I am interested in the major Internet cables that land on the East coast of the United States and I would like to visit one of these locations to understand how they are built and whether they are adequately protected. As with any critical infrastructure project, access to the internals of a system is often closely guarded. Therefore, I plan on contacting individuals who have studied these systems or individuals that are responsible for managing these systems. If I can find an individual who will give me access, then I will have a much easier time doing research. Part of this process may require doing an interview with one of these primary actors.

Once I have collected the necessary materials I will analyze the data using a qualitative approach. I do not anticipate that I will use a quantitative approach. Different sections of the paper will lend themselves to different methods. In my network analysis section I will rely heavily on textual sources and databases that contain critical network infrastructure information. I may incorporate an interview into this section depending on how well I can gain access to a particular network center. In the second section on an online backup service, I will rely on a mix of popular sources and technical information to assess the structure of the online service.

Preparation

Due to the relative accessibility of digital resources, the preparation time prior to conducting primary source research should be minimal. The same should hold true for any material that I should need to access.

In terms of courses to take prior to undertaking research, it would be useful to enroll in a class focused around the use of space (either in architecture or geography) to gain access to a particular vocabulary that I can utilize in my written paper. Additionally, taking courses centered around the history of technology and the Internet can enrich my theoretical background that I have previously described. I already have personal knowledge of digital technologies so I do not anticipate needing to take more technically-based courses. However, if I do encounter technologies that I need an explanation of, I hope to draw upon the expertise of my committee and the Virginia Tech community.

The topics of my research (network architecture and cloud data management) will require that I make connections with people knowledgeable on network architecture. I will need to connect to some of these people so that I can talk to them informally and use them to help me with the technical information in my final paper. I have already met with a few knowledgeable people about network infrastructure at Virginia Tech. I need to find some people with more specialized knowledge.

In terms of institutional approval, I will have to create a consent form for the institutional review board. Additionally, any field work that I may do at certain network locations will require not only IRB approval but may also entail applying for grants for travel funds. Without much knowledge at the moment about specific locations, I am unable to anticipate specific costs for travel to the specific location(s).

Significance

The research I plan to conduct has a great deal of relevance for both the STS community the broader scholarly community interested in the impact of digital technologies and the Internet.

As I previously mentioned, my research topic is not entirely new but it does look at new sets of actors that present possible new avenues for other researchers. The most compelling aspect of the research topic is not necessarily the topic itself, but the specific ways in which I draw upon an interdisciplinary approach to address one of the emerging phenomena on the Internet.

It is in this methodological approach that I hope to contribute and uphold the primary spirit of STS scholarship. Although I do not anticipate my research to be an activist piece of scholarship, I do hope that my dissertation and its byproducts can help inform other actors who are interested in political

work around this topic. Information system builders, additionally, could potentially develop computing practices that are historically informed by the cases of cloud computing failures and successes.

Schedule/Timeline

In sum, my hope is that this research project will be a total of a two to three year project. As of right now, I am entering my third year and will continue to do research locally on the topic. For the following two years, I will work on my project by reading through the textual sources and begin reaching out to actors who are knowledgeable about the history of cloud computing. While doing this research, I will continue to write and work on different sections of the dissertation. I have included a chapter outline of the different sections that I hope to write on (on summary page).