DIGITAL DARK AGE: AN OVERVIEW FOR THE HUMANITIES AND SOCIAL SCIENCES

ANA LÚCIA MIGOWSKI

ana-lucia.migowski@gcsc.uni-giessen.de

Ana Lúcia Migowski received her Master's degree in Communication and Information from the Federal University of Rio Grande do Sul, Porto Alegre, Brazil, in 2013. Since then, she has taught as a lecturer in Social Communication faculties and postgraduate courses in Brazil. Currently she is a PhD candidate at Justus-Liebig-University, Giessen. Her doctoral research deals with the use of digital media, especially social networking sites, in the elaboration of cultural and communicative memories. Migowski is also a member of the *International Graduate Centre for the Study of Culture* (GCSC).

HOW TO CITE

Ana Lúcia Migowski. "Digital Dark Age: An overview for the Humanities and Social Sciences." *On_Culture: The Open Journal for the Study of Culture* 1. (2016). <http://geb.uni-giessen.de/geb/volltexte/2016/12071/>.

Permalink URL: <http://geb.uni-giessen.de/geb/volltexte/2016/12071/>

URN: <urn:nbn:de:hebis:26-opus-120710>

Digital Dark Age: An overview for the Humanities and Social Sciences

_Abstract

The possibility of a Digital Dark Age worries computer scientists, archivists, and librarians, but it also concerns humanists and social scientists. The absence of access to digital data and cultural products due to the obsolescence of technologies used for today's communication, entertainment, work, production and circulation of scientific knowledge is an imminent risk. This report is based on texts and interviews with experts. Here we provide an overview of this emergent and urgent problem and present suggestions for prevention.

1_Let's start with a Futurology exercise

The possibility of a Digital Dark Age worries computer scientists, archivists, and librarians, but it also concerns humanists and social scientists.

The absence of access to digital data and cultural products due to the obsolescence of technologies used for today's communication, entertainment, work, production and circulation of scientific knowledge is an imminent risk.

Preserving digital content is an emergent emergency, which becomes urgent as social appropriations of digital technologies increase. This complex phenomenon requires the special attention of scholars from humanities and social sciences, who may approach this problem by going beyond its technical dimension.

From the digitization of archives and the intensification of social practices across digital platforms emerges the need for studies and practical actions towards an understanding of the future of digital information.

This report is based on texts and interviews with experts in this and related topics. Here we provide an overview of this emergent and urgent problem and present suggestions for prevention.

Imagine two scholars working in the field of humanities in 2080. Nina and John have an interdisciplinary approach; this is appropriate for the academic period in which they live. They share a common interest in understanding social and cultural life in the first

part of the 21st century. The "Information Age," however, challenges and imposes constraints upon those willing to generate knowledge about it. Here starts the (fictional) history of our fellows from the future.

Welcome to the 2080's Scholarship in Humanities and Social Sciences

John deals with Creative Processes in Literature, which is the theme of his PhD research. This scholar proposes an analysis of books written by Nobel Prize winners between 2005 and 2015. Unfortunately, these writers are not able to narrate their reminiscences in 2080. Biographies and books can be easily found, but raw materials (i. e. manuscripts and messages sent through e-mail to editors and other people who provided the authors with ideas and suggestions) are gone. John faces a serious methodological problem. Some libraries and archives around the world still preserve letters, manuscripts and diaries of important writers, but not the digital messages that replaced traditional mail in the 21st century.

Nina is a German historian developing a research study on the Arab Spring, one of the most important social uprisings that happened at the beginning of this century. Her grandparents were young Syrian citizens when the revolution started in 2011. They used to tell her stories about the protests, as well as their journey, as refugees, to Germany during the civil war. Besides recounting the difficult moments of violence they experienced, they always emphasized how important smartphones and social networking sites were for them to communicate with friends and family those days. Nina's grandfather keeps his old iPhone 4S in a drawer. Unfortunately, he is not able to access pictures anymore, or videos and messages he exchanged with partners 69 years ago. Those contents were lost, together with Facebook and e-mail records and the smartphone's operating system. A study focused on secondary sources would still be a possibility for Nina, since newspapers and some academic papers written at that time are available for consultation. The messages exchanged during that period, however, are now missing or unreadable bits and bytes.

2_The problem

Despite being fictional, these two stories are likely to become reality. Interviews and public talks given by Vint Cerf, currently Google's vice-president, have alerted the world to the risks of the complete loss of digital information that shadows our society.

> As time goes on and as we become increasingly dependent on software to create, render and interact with digital content, we'll also become increasingly dependent on our ability to preserve access to and to correctly interpret digitized information of all kinds.[1]

Dr. Cerf, who was involved in the creation of the Internet back in the 1970's, co-designing the TCP/IP protocols and architecture of the Internet, has also presented this same futurology exercise, in a Lecture at Carnegie Mellon University in February 2015.

In his own words, in the 22nd century, Doris Kearns Goodwin wouldn't have been able to write the book "A Team of Rivals", a biography of former US president Abraham Lincoln, which was based on letters he exchanged with members of his cabinet. Apart from that, Vint Cerf addresses other topics involving digital preservation, as well as possible alternatives to the problem.

The discussion proposed by Cerf is highly focused on technological issues. But, is this just a matter of technological development and access to antique digital data? This is a decisive aspect that will allow future generations to know what content and information-sharing practices looked like in our present. Some scholars argue that other cultural and political issues are also involved in this problem.

Actually, this issue isn't new. Historic sources have always been the result of selective processes driven by the curatorial work of various social actors (archivists, historians, politicians, dictators, religions, and so on). Most of the content and ideas conveyed by a society within a certain period of time, for example, are not meant to be stored. They are recollected via oral historiography and other methods.

This will continue to be the case, but wouldn't it be interesting for social scientists, historians, linguists, anthropologists and all sorts of scientists dealing with culture to have access to that information? If this kind of digital preservation were technologically possible, wouldn't it open space for the emergence of new areas of study for the humanities?

This is a paradoxical topic. While some are worried about the preservation of digital information, many also fight for the right to forget on the Internet. Finding ways to

combine all the potentialities and challenges the information age has been posing seems an urgent task.

The paradox

> The "right to be forgotten" ruling allows EU residents to request the removal of search results that they feel link to outdated or irrelevant information about themselves on a country-by-country basis.[2]

> Digital preservation is the active management of digital content over time to ensure ongoing access.[3]

This is just one of the paradoxes we face today and it has to be framed and understood from various points of view.

With the collaboration of Desiree Butterfield-Nagy, archivist at the University of Maine (USA), and Gustavo Fischer, Professor in the Communication Sciences Department at The University of Vale do Rio dos Sinos (Brazil), we aim to present an overview of what the Digital Dark Age represents for the future and present of the Humanities and Social Sciences.

Desiree Butterfield-Nagy is an archivist in the Special Collections Department at the University of Maine's Raymond H. Fogler Library. She has coordinated several digitization projects through the DigitalCommons@UMaine institutional repository. She also teaches in the university's digital curation certificate program offered through the New Media Department.

Gustavo Fischer has a PhD in Communication Sciences from the University of Vale do Rio dos Sinos. Currently, he teaches in the same University, both graduate and undergraduate courses. His research topics, always from a techno-cultural approach towards objects with audiovisualities potency, intend to advance the studies of interfaces in online materialities, as well as the construction of methodologies influenced by many media archaeological perspectives.

3_Selection and preservation of digital content

Preserving all content produced by humankind is impossible, and in certain cases even undesirable. Technological aspects, power and cultural relations also affect the choices of whether something is worth storing and preserving or not. Carr, in his book *What is History?* illustrates this in a very meaningful way:

> Our picture has been preselected and predetermined for us, not so much by accident as by people who were consciously or unconsciously imbued with a particular view and thought the facts which supported that view worth preserving.[4]

Preservation policies, institutional or not, usually direct what cultural products ought to be selected and protected for long-term periods. These curatorial processes make it more likely materials of the past will be available for society in the future. In sum, values of the present guide the choice of what is worth saving for the future. A lack of historical distance with regard to these decision-making processes also has consequences.

> If we fall into the trap of saving only what is considered "important" by today's standards, we significantly limit possibilities for the future.[5]

Within this scenario, technological, cultural and economic issues are at stake. Digital preservation relies on many different and sometimes divergent aspects.

After selecting what ought to be preserved, properties such as authenticity, integrity, and chain of custody have to be considered.[6] Apart from these technological and political decisions, cultural aspects also have to be taken into account.

4_Technological issues

In technological terms, Jeffrey explains the three main causes behind the claims of a possible Digital Dark Age.[7]

Main reasons for digital data loss:

Data corruption: Digital storage media, such as DVDs and magnetic tape degrade over time.

Media obsolescence: Hundreds of previously popular storage formats are now unreadable for practical purposes because the format is obsolete, e. g. 5.25-inch floppy discs, laser discs. The software used to create or access the data uses proprietary formats or becomes obsolete. Software and file formats change very frequently as technology changes; there is no guarantee that a document created in a word processing application, such as WordStar, will be readable in newer software. This also applies to proprietary (commercially confidential) formats, that is, formats that are created by a specific software package, but are not readable by any other without conversion.

Inadequate metadata: This problem should, in theory at least, affect only datasets that have been somehow abandoned before they could be prepared for archive. A data file may contain valuable and important information, but without the metadata, that is, the data explaining what it is and how to read and understand it, it may in fact be entirely unusable.

Comparing the stability and durability of digital content with its predecessors, i. e. paper and tape, we see that the life expectancies of contemporary formats are not better than the previous ones. There is no consensus regarding the lifespan of each media and its storage capacity. However, most of the reports and tutorials point to the necessity of:

- Migration of contents for newer media formats on a regular basis.
- Distribution of content to more than one media.
- Use of open source software.
- Introduction of metadata to raw data.

This scenario also points to the need for cultural, educational and political movements towards the preservation of digital content. What would have happened to Nina's research if only a digital and open source archive had been created during the Arab Spring? Wouldn't Nina's work become much easier and more insightful? How far could John have gone if the correspondences of the Nobel Winners had been saved as metadata to their books?

5_Cultural Issues

Technology alone won't solve the problems to come. Scholarship focused on the inclusion of digital technologies within the classroom, for example, has already stated*: without new teaching methods, nothing will change*. The same is true for social practices related to the way we produce, consume and store digital information.

> This is a very paradoxical time, especially concerning the notion of preservation, archiving and memory.
>
> On the one hand we have an idea — which is strongly defended by ICT companies — that one can have infinite storage space (cloud computing services and mobile operators). On the other hand, in a less explicit way, we see that the old idea of "planned obsolescence" still remains. A series of products and services based on digital technologies also have a life cycle (to keep the market system working, some would say).

However, there had perhaps never been such a trust in the capacity of digital devices to run and store data of all sorts that concern us. Thistensionbetweenpreservation/ discontinuity, or degeneration/re-generation is a sign of problems and challenges in the field of Humanities and Social Sciences.[8]

Both institutions and individuals might think of these paradoxical situations and envision not just new platforms with more and more storage space, but also reflect on how they are used and what their future is.

Today, e-mails are exchanged and pictures taken every second. Do we know what is worth saving? Or will we be worried about it only in the moment a hard drive crashes or the pictures we posted to a social net-working site are deleted when the company hosting them closes down.

Movies like "Her" help us think about the meaning of building ourselves as subjects crossed by the complexities of the current Software Culture (as Manovich would say).[9]

Trailer "Her": https://youtu.be/dJTU48_yghs

We live in a "software culture" that is, a culture where the production, distribution, and reception of most content is mediated by software.[10]

6_Preservation policies and initiatives

The cultural heritage community, which is composed of archivists, librarians, museologists, politicians, and activists, to cite some of its members, also plays a key role in regard to digital preservation.

They are responsible for setting rules and producing common ground that ensures selection and access to the representative information of our Era.

Certainly some institutions are developing specific projects to document topics particularly important within their communities, or those identified as likely to have enduring research value, including the University of California, Truth Tobacco Industry Documents, and the Occupy Archive by the Roy Rosenzweig Center for History and New Media.[11]

Truth Tobacco Industry Documents: An archive of 14 million documents created by tobacco companies about their advertising, manufacturing, marketing, scientific research and political activities, hosted by the UCSF Library and Center for Knowledge Management.[12]

#Occupy Archive is documenting and saving the digital evidence and stories from the Occupy protests worldwide that began in September 2011 in Lower Manhattan. Occupy Wall Street (OWS) inspired groups to form in small towns and large cities around the world. The #Occupy Archive seeks to represent each of those groups with individual collections.[13]

> Advocacy will continue to be important, as we share with legislators, institutions, and members of the general public the level of investment needed to preserve digital content, along with education and outreach to groups and individuals to offer suggestions for preserving information. Those in the humanities can continue to emphasize the breadth of research being conducted and the types of documents being drawn upon so that we can better understand how those needs translate into preservation efforts in a digital realm.[14]

7_Other initiatives involving Humanities and Social Sciences

Library of Congress: From manuscripts, papyrus, and historical documents to digital games and movie tapes, pre-digital and born-digital data are transformed to and preserved through zeros and ones. One of the most known programs is the National Digital Information Infrastructure and Preservation Program (NDIIPP), created within the Library of Congress (USA).

In 2000 the US government started investing in this project, which now counts on more than 130 partners (among universities and other national departments) in order to implement:

> a national strategy to collect, preserve and make available significant digital content, especially information that is created in digital form only, for current and future generations.[15]

The Internet Archive: The Internet Archive is a "non-profit company that was founded to build an Internet library. Its purposes include offering permanent access for researchers, historians, scholars, people with disabilities, and the general public to historical collections that exist in digital format". Through its Wayback Machine archive, users can have access to different versions of websites that are constantly scanned and saved by crawlers (a retrieval system used by search engines like Google to index websites on the web).

When asked what makes the information they collect useful, the Internet Archive answers:

> Most societies place importance on preserving artifacts of their culture and heritage. Without such artifacts, civilization has no memory and no mechanism to learn

from its successes and failures. Our culture now produces more and more artifacts in digital form. The Archive's mission is to help preserve those artifacts and create an Internet library for researchers, historians, and scholars.[16]

The Film Foundation: Martin Scorsese, film director of Taxi Driver (1979), The Departed (2006) and other famous movies, started his campaign for the sake of Film Preservation in 1980. After realizing that the media of thousands of films wouldn't survive for a long time, Scorsese wrote a letter to his friends stating: *"Everything we are doing now means nothing!"*[17]

The director's claim was that something urgent should be done if film directors wanted their analogue film shootings to last for future generations. A group of directors, editors and moving-images experts started working on the digitization, restoration and maintenance of old movies. Educational programs are also undertaken by this foundation. By now more than 700 have been preserved.

Gustavo Fischer draws our attention to Scorsese's statement over the importance of preserving movies.

> Film preservation is something absolutely essential. We know that movies older than 100 years can still be screened. Will we be able to screen a digital movie in 10 years? Unfortunately, we don't have an answer for that question. There is still lots of work to be done so that we can assure that the creative work of filmmakers will be preserved for the future.[18]

These initiatives live under the danger of losing their archives if they are not updated and migrated to newer formats and devices.

'Challenge' is the keyword evoked by most of these projects' descriptions, especially when it comes to preserving data. Their efforts selecting valuable information, implementing relevant metadata and understanding the needs of social scientists and scholars in the area of culture, however, cannot be done without representatives of those disciplines.

8_What can scholars do?

History, cultural studies, literature, archaeology, linguistics, sociology, anthropology, economy and political science scholars may also get involved in this debate. The maintenance of digital data, both those created in this environment or digitized afterwards, depends on an active work of selection and contextualization.

Papers and academic books also contribute to the documentation of contemporary social practices. Without it the interpretation of information about our present won't succeed in the future.

Digital humanities is generally defined as an interdisciplinary field that bridges gaps between the humanities and computational sciences. Scholars working in this field come from multiple disciplines — from media to studies to medieval history — and use digital technologies while dealing with source materials.

> Computer science has contributed greatly to archivists' efforts to develop methods for preserving digital files, but it is true that a broad, interdisciplinary approach could help make sure that we pre-serve a wide range of content using methods that will preserve its usefulness, even when dealing with complex digital objects like relational databases, layered image files, hyperlinked text and 3D renderings.[19]

There are various critics of this approach, especially concerning the fact that including new devices and techniques to the discipline, doesn't necessarily produce a new one. Despite those controversies, this field has been producing advanced techniques, combined with the capacity of critical interpretation of Social Sciences and Humanities scholars that can contribute to the preservation of digital data.

The so-called digital humanities answers questions that otherwise wouldn't be answered with traditional humanities methods and technological apparatus. On the other hand, humanists and social scientists, with their disciplinary background, being familiar with the possibilities digital technologies, can ask even more complex questions and explore greater corpora and the relationships among them. Finally, digital humanities seem to be another space within the academy where the divide between making and interpreting might be bridged in productive ways.[20]
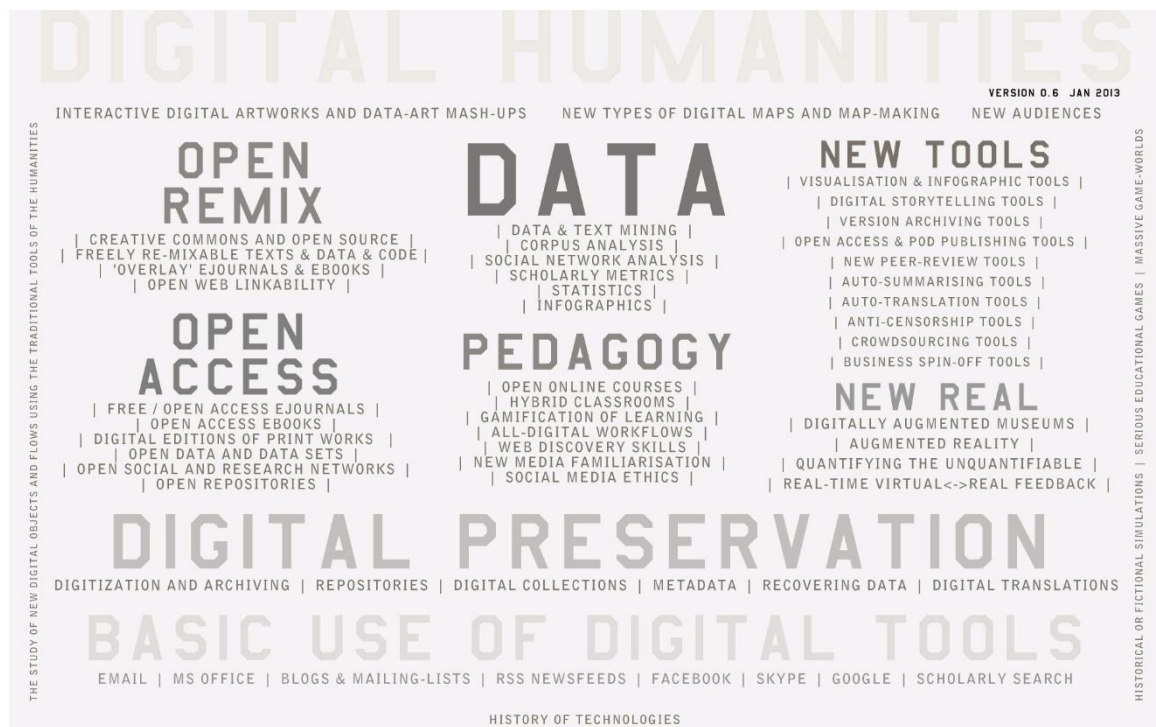
Fig. 1: Conceptual map of areas covered by the Digital Humanities[21]
Read more about this map in Haden's blog.

The following projects present academic work that developed according to the interdisciplinary approach claimed by digital humanities.

Infinite Ulysses: This project is part of Amanda Visconti's PhD research (defended in 2015). She developed a digital edition of the book Ulysses by James Joyce in order to allow users and collaborators to read the text "together," annotating and commenting in it. The text code is available online and annotations can also be accessed and downloaded. As an open access and open source project, other researchers can make use of the platform's code and content for their own projects. Infinite Ulysses contributes both to the field of literature and media studies. Contingency plans are described by Dr. Visconti and include constant backup and updates.

Creating new uses for digitalized materials seems an opportunity for the Humanities and Social Sciences.[22]

1914–1918 online: The project "1914–1918 online – International Encyclopedia of the First World War" is an English-language virtual reference work on the First World

War. The multi-perspective, open-access knowledge base is the result of an international collaborative project involving more than 1,000 authors, editors, and partners from over fifty countries.

More than 1,000 articles will be gradually published. Innovative navigation schemes based on Semantic Media Wiki technology provide nonlinear access to the encyclopedia's content". The project is based on the Open Access Paradigm, promoting free and unlimited dissemination of the content to individual users, search engines, and reference services.[23]

Signs@40: This project proposes visualization models of texts published by Signs Journal, focused on Feminism Studies. Signs@40 offers visualization models of co-citation patterns among authors, topics approached through the editions and other aspects. The project is a powerful tool for researchers dealing with the theme of feminism and willing to understand the academic production from a historical and epistemological perspective.[24]

The challenge of preserving the content of those projects remains, but the first step towards digitization and availability of data has been taken. With the regular preservation of digital-born data, updates, and migration of content to newer technologies, future scholars will still have access to our cultural heritage. Ethical, financial and political issues ought to be addressed in order to turn this into a reality.

> I suspect that intriguing questions will still be applied to content that manages to survive, but there may be many questions that researchers feel compelled to explore, but find surprisingly little to draw from, content that we take for granted as being available today.[25]

If the work on the construction, preservation and maintenance of digital data is done today, in 2080 our researchers will have access to richer datasets and there will be fewer limitations to the questions they want to address.

_Endnotes

1   Vint Cerf, "Lest we Forget," in *IEEE Computer Society* (2015), 80–81, accessed May 6, 2016, <http://ieeexplore.ieee.org/stamp/stamp.jsp?reload=true&tp=&arnumber=7111912>.

2   Samuel Gibbs, "Google to extend 'right to be forgotten' to all its domains accessed in EU," in *The Guardian*, February 11, 2016, accessed February 12, 2016, <http://www.theguardian.com/technology/2016/feb/11/google-extend-right-to-be-forgotten-googlecom>.

3   Library of Congress, "About," accessed 2 February, 2016, <http://www.loc.gov/library/libarch-dig-ital.html>.

4   Edward Carr, *What is History*? (London: Penguin, 1987), 13.

5   Desiree Butterfield-Nagy, "Digital Dark Age and the Future of Humanities and Social Sciences," e-mail interview by author, January 11, 2016.

6   Cf. Reagan Moore, "Towards a theory of digital preservation," in *The International Journal of Digital Curation* 3/1 (2008), 63–75.

7   Stuart Jeffrey, "New Digital Dark Age? Collaborative Web Tools, social media and long-term preservation," in *World Archaeology* 44/4 (2012), 553–570, here: 556.

8   Gustavo Fischer, "Digital Dark Age and the Future of Humanities and Social Sciences," e-mail interview by author, February 19, 2016.

9   Ibid. (cf. note 8).

10   Lev Manovich, *Software takes command* (New York: Bloomsbury, 2013).

11   Butterfield-Nagy, e-mail interview by author, January 11, 2016 (cf. note 5).

12   Truth Tobacco Industry Documents, "Home," accessed February 14, 2016, <https://industrydocu-ments.library.ucsf.edu/tobacco/>.

13   Occupy Archive, "Home," accessed February 6, 2016, <https://occupyarchive.org>.

14   Butterfield-Nagy, e-mail interview by author, January 11, 2016 (cf. note 5).

15   Library of Congress, "About," (cf. note 3).

16   Internet Archive, "About," accessed February 17, 2016, <https://archive.org/about/.

17   The Film Foundation, "The Foundation," accessed February 6, 2016, <http://www.film-founda-tion.org>.

18   "Diretor Não Quer Só a Cópia Digital, Diz Martin Scorsese," interview by Martin Scorsese, *Folha De São Paulo* (São Paulo), September 29, 2015, accessed February 3, 2016, <http://folha.com/no1687724>.

19   Butterfield-Nagy, e-mail interview by author, January 11, 2016 (cf. note 5).

20   Kathleen Fitzpatrick, "The Humanities, Done Digitally," in *Debates in the Digital Humanities*, eds. Matthew K. Gold and Lauren K. Klein (Minnesota: University of Minnesota Press, 2012), n. p, accessed May 6, 2016, <http://dhdebates.gc.cuny.edu/debates/text/30>.

21   David Haden, January 13, 2013, "Digital Humanities Map," in *My general observations*, accessed May 11, 2016, <https://jurnsearch.wordpress.com/2013/01/13/digital-humanities-map/> (2013).

22   Cf. Infinite Ulysses, "The project," accessed February 6, 2016, <http://www.infiniteulysses.com>.

23   Cf. 1914–1918 Online, "About," accessed February 6, 2016, <http://encyclopedia.1914-1918-online.net>.

24   Cf. Signs@40, "About," accessed February 2, 2016, <http://signsat40.signsjournal.org/>.

25   Butterfield-Nagy, e-mail interview by author, January 11, 2016 (cf. note 5).