

Introduction

Seth Cayley, VP Gale Primary Sources

“Founding myths” are an integral part of the technology sector. Every successful company has some version of a story about how it emerged by accident, or eureka moment, or how it was built from the ground up while the founder lived in a shipping container.

Examples of founding myths include Apple being started in Steve Jobs’s garage or Bill Gates dropping out of Harvard to start Microsoft. These are “myths,” not in the sense that they are false, but in that they simplify complex histories and downplay the important roles of other people and institutions.

Apple’s co-founder Steve Wozniak himself has confessed “the garage represents us better than anything else, but we did no designs there.... There were hardly ever more than two people in the garage and mostly they were sitting around kind of doing nothing productive.”¹ Similarly, the “drop-out” narrative that surrounds Bill Gates obscures his privileged background (along with his co-founder Paul Allen) attending one of the most prestigious private high schools in America, where he was lucky to have had access to computers.

The consequence of these myths is to amplify the ideal of the solo (usually white and male) genius.

Digital humanities has its own founding myth. Around 1949, the Jesuit Roberto Busa approached the CEO of IBM, Thomas J. Watson Sr., and convinced him to fund and resource his planned *Index Thomisticus*, a tool that Busa intended would allow the user to carry out textual analysis of the complete works of St. Thomas Aquinas. In Busa’s own telling of the story, it is presented as a David and Goliath scenario, with the humble Jesuit priest persuading a high priest of capitalism to support a project of pure scholarship.² From such unlikely beginnings a whole movement of “humanities computing” was born. As Julianne Nyhan has observed, this underdog story has helped bind the digital humanities community together as a group, differentiated from mainstream humanities.³

Without denigrating Busa’s unquestionable achievements, there are problems with the Busa founding myth. The scholar Tito Orlandi notes that there were other scholars working in the fields of translation and archaeology who should be considered amongst the pioneers of what is now known as digital humanities.⁴ Also, Busa did not work alone. As forthcoming work will highlight, he relied on a team of (female) punched card operators in the 1950s and 1960s to make his vision a reality.⁵ By focusing on Busa, we diminish both the importance of teamwork in his success, and we also ignore the wider international scholarly environment that contributed to the development of digital humanities.

Digital humanities in 2019 is much more diverse than the Busa founding myth implies. It is a community that emphasises collaboration, transcends narrow disciplines, and is truly international in outlook. At the time of writing, the project *Oceanic Exchanges*, which traces global information networks in historical newspapers, includes 39 researchers working across 6 national groups.⁶ Recent Alliance of Digital Humanities Organizations conferences have made conscious efforts to invite participants from the Global South, and in 2021 Tokyo will be hosting the first such conference to be held in Asia.⁷

While, as Nyhan says, the Busa myth may have been a “useful fiction,”⁸ digital humanities is moving beyond its confines and into the mainstream. It is notable that a significant amount of the community conversation is about how to *teach* digital humanities. In the last couple of years, a steady number of papers and books have been published discussing how to integrate digital humanities into the classroom and curriculum.⁹

My personal journey into digital humanities began in 2012, when I supported a project headed by Dallas Liddle to do distant reading of thousands of editorials from the pages of the *Times*. By analysing changing patterns of word counts, Dallas's research challenged long-held beliefs about the slow pace of change in Victorian newspapers ... all without specifically requiring him to read any of the actual editorials.¹⁰ The project challenged my own assumptions about what "history" could be. Text could be treated as data. I was hooked.

Since then, Gale has undergone a culture shift to make the underlying data of Gale Primary Sources archives more accessible for such projects around the world. We have provided support to a number of digital humanities projects, looking at ideas as disparate as Tudor networks of power (led by Ruth Ahnert¹¹), eighteenth-century printers' ornaments (Fleuron, led by Hazel Wilkinson¹²), and tracing the way news was reprinted across the Anglophone world in the nineteenth century (Scissors and Paste, led by Melodee Beals¹³). In 2018 we launched the Gale Digital Scholar Lab, a cloud-based text and data environment specifically designed to lower the barriers to entry for researchers of all levels to carry out digital humanities workflows.¹⁴

The time therefore felt right to bring together scholars, librarians, and students to learn about the latest research and trends in digital humanities from a diverse range of international speakers. *Digital Scholarship, Digital Classrooms* offers a selection of the eight papers that were presented at the resulting Gale Digital Humanities Day at the British Library, held on 2 May 2019.

The day was split into four distinct sessions: (1) Literature & Distant Reading; (2) Computers Reading the News; (3) Digital Humanities in the Classroom; and (4) Institutional Support and Infrastructure for digital humanities. Each session was designed to put a spotlight on different themes in digital humanities, with the first two focusing on the latest research, and the latter two exploring issues around teaching and the role of the library.

In this volume you will find four of the papers that were presented on that day.

In the first paper, "On Principles and Values: Mining for Conservative Rhetoric in the London *Times*, 1785–2010," **Joris van Eijnatten**, Professor of Cultural History at Utrecht University, provides a textbook example of how using n-grams and other textual analysis tools can lead to new and interesting discoveries. In his work on moral language and historical newspapers, Joris explores the development of conservative rhetoric in *The Times Digital Archive*. He charts the rise and fall of certain phrases in political reporting, such as the shift from "conservative principles" to "conservative values." Placing these findings in historical context, he notes that "political opinions began to be couched explicitly in moral terms after the 1960s," and that new forms of rhetoric "emerged from a tension with modernity" provoked especially by the permissive society of the 1960s.

Carrying out such research is impossible without the right infrastructure in place. In "Reflections on Infrastructures for Mining Nineteenth-Century Newspaper Data," **Julianne Nyhan**, Associate Professor of Digital Information Studies and Deputy Director of the University College London (UCL) Centre for Digital Humanities, discusses her team's experience of trying to get a large-scale text-mining project off the ground. She highlights how even if researchers have access to data, and are at well-resourced institutions such as UCL, there can be a succession of unexpected challenges to begin being able to analyse that data. Despite having support from the university's Research IT Services group, Julianne notes that "the computing infrastructure was not really set up to support Humanities research," and at a cost of £350 per day, each additional day it took to mount the data drained the project budget. She closes her piece with recommendations for how universities should address the "digital turn," and the potential role of public-private partnerships.

The topic of **Ryan Cordell's** paper is "Teaching Humanistic Data Analysis." Ryan teaches both graduate and undergraduate courses at Northeastern University and has reflected deeply on his own teaching philosophy.¹⁵ In his words: "One of my primary goals as a teacher is to work against notions of the 'digital native' and the ruinous pedagogical consequences following from that designation.... I believe the most pressing scholarly questions in the coming years will require true interdisciplinarity. By this I mean not individual scholars who dabble in many

disciplines, but groups of scholars who can contribute their various expertise to a sustained and substantial enterprise. It is this future that I hope to prepare my students to meet with creativity and verve. Doing so requires not a pedagogy founded on particular technologies or tools, but on capacious scholarly and technological imaginations.” Ryan’s paper expands on these ideas, arguing that it is not simply programming skills that we need to teach, but a mindset for thinking about data—in his words, “programmatically thinking.”

Continuing with the classroom theme, **Sarah Ketchley**, an Egyptologist and digital humanities specialist based at the University of Washington, discusses her own experiences of leading undergraduates through her Introduction to Digital Humanities course. Sarah talks through the learning goals and pedagogical challenges she faced, along with the methodological and technological solutions she used to overcome those challenges. Using her Autumn 2018 class as a case study, she explains how her students used the Gale Digital Scholar Lab to interrogate their data sets and use the resulting analysis to create impressive digital exhibits. Sarah concludes by reflecting on the learning outcomes and how teaching digital humanities helps students develop transferrable skills for life beyond university: “the group was able to clearly identify a range of specific skills they were developing and using in class as being relevant and ‘marketable’ to employers.”

Other speakers on the day included: **Mark Algee-Hewitt**, Director of the Stanford Literary Lab, who gave an overview of the Lab’s “Microgenres” project, which looks at the way in which novelists draw on disciplines such as philosophy, history, and natural science with their narratives; **Tomoji Tabata**, Associate Professor of Corpus Linguistics at the University of Osaka, who explained his use of stylometry to analyse eighteenth- and nineteenth-century literature, and revealed to the audience that Charles Dickens’s favourite phrase was “his hands in his pockets”; **Melodee Beals**, Lecturer in Digital History at the University of Loughborough, who discussed unexpected uses of heritage and historical data; and **Lisa McIntosh**, Director of Access Services at the University of Sydney, who explained how her university library is developing its infrastructure to support the latest research. Although their papers are not published in this volume, I encourage you to seek out their work.

I would like to extend my thanks to all the speakers for sharing their projects and ideas.

NOTES

1. <https://www.washingtonpost.com/news/morning-mix/wp/2014/12/05/steve-wozniak-refutes-apples-creation-story-the-garage-is-a-bit-of-a-myth/>.
2. R. Busa, “The Annals of Humanities Computing: The *Index Thomisticus*,” *Computers and the Humanities* 14, no. 2 (1980), 84.
3. Julianne Nyhan and Andrew Flinn, *Computation and the Humanities: Towards an Oral History of Digital Humanities* (Springer International Publishing, 2016), 270.
4. Julianne Nyhan and Andrew Flinn, 79.
5. Julianne Nyhan and Melissa Terras, “Uncovering ‘Hidden’ Contributions to the History of Digital Humanities: The *Index Thomisticus*’ Female Key punch Operators,” Digital Humanities 2017 (Montreal, QC, Canada). See also the authors’ paper presented at DH2017: https://discovery.ucl.ac.uk/id/eprint/10052279/9/Nyhan_DH2017.redacted.pdf.
6. Oceanic Exchanges Project Team. *Oceanic Exchanges: Tracing Global Information Networks in Historical Newspaper Repositories, 1840–1914*. 2017. DOI 10.17605/OSF.IO/WA94S. Available at osf.io/wa94s.
7. List of ADHO conference locations available at <http://adho.org/conference>.
8. Julianne Nyhan and Andrew Flinn, 80.
9. For example, Claire Battershill and Shawna Ross, *Using Digital Humanities in the Classroom: A Practical Introduction for Teachers, Lecturers and Students* (Bloomsbury, 2017), and Anna Wing-bo Tso, editor, *Digital Humanities and New Ways of Teaching* (Springer, 2019).
10. Dallas Liddle, “Reflections on 20,000 Victorian Newspapers: ‘Distant Reading’ the *Times* using *The Times Digital Archive*,” *Journal of Victorian Culture* 17, no. 2 (June 2012), 230–237.
11. <https://ahrc.ukri.org/research/case-study-archives/tudor-networks-of-power-1509-1603/>.
12. <https://fleuron.lib.cam.ac.uk/>.
13. <http://scissorsandpaste.net/>.
14. <https://www.gale.com/intl/primary-sources/digital-scholar-lab>.
15. <https://ryancordell.org/statements#teaching>.