Digital Scholarship in Humanities: The Agile Way

Gautam Gogoi
Designation: Librarian, Assam Don Bosco University, Tapesia Campus
Email: gautam.gogoi@dbuniversity.ac.in

Bhaigyashree Boro
Designation: Research Scholar, Dept. of Library and Information Science,
Mizoram University
Email: bhaigyashreeboro24@gmail.com

Mouchumi Das
Designation: Assistant Librarian, The Assam Royal Global University,Guwahati
Email: dasmouchumi84@gmail.com

Abstract:

The gradual transition to a digital society necessitates that all organizations and institutes adapt to the new environmental conditions created by their digital transformation. The status quo is already being challenged by digitization. Across the university the way in which we pursue research is changing, and digital technology is playing a significant part in that change. Indeed, it is becoming more and more evident that research is increasingly being mediated through technology. However, due to the fact that technological innovations necessitate knowledge that is not present or is not mature, some organizations and institutions are not fully prepared for the impending disruption. The knowledge gap affects many different types of organizations. This article examines the case of digital scholarship in the contexts of open access movements and digital humanities in order to achieve agility in the humanities and social sciences for the next generation in the digital era. It also highlights some of the difficulties encountered. These tasks are especially important for departments and libraries because digital technologies are altering core research activities such as information creation, storage, and dissemination.

Key Words: Open Access, Digital Humanities, Digitalization, Textual Analysis, Digital Divide

Digital Scholarship in Humanities: The Agile Way

1. Introduction

The hypertext visionaries saw the power of richly interconnected global information systems to advance human knowledge. The Internet provided the infrastructure to make those ideas a reality, and it quickly became a platform for collaborative research and data sharing. New ways of using the Web for eResearch have emerged as it has evolved, such as the social networking capabilities enabled by Web 2.0 technologies. The next generation of the Web, known as the Semantic Web, is now on the horizon, allowing for the emergence of new types of collaborative research. We need a discipline that studies the Web as a whole if we are to understand and anticipate these new modes of collaboration (Hall et al. 2009)

Researchers from all disciplines are using new technology to conduct new research. Much of this user-centered activity is based on the Web distributed application platform, which includes 'mashups' for integration, easy access to computational resources 'in the cloud,' and social networking to share digital science results and practice. The Semantic Web will enable advancements in this area, continuing the trend of technology empowering individuals.

The Digital Era is defined by technology that accelerates and broadens knowledge turnover in the economy and society. As an explanation for the system we live in, evolutionary theory states that sustainability is dependent on knowledge turnover. Knowledge turnover is low in parts of the system that are relatively stable, and new variation, when produced, is rarely retained. Faster knowledge turnover is advantageous in other, less stable parts of the system because new knowledge is produced more frequently, allowing for adaptation to the changing surrounding environment. The mixing and matching of knowledge turnover rates results in a dynamic but everlasting world. The Digital Era can be viewed as the evolution of an evolutionary system in which knowledge turnover is not only extremely high, but also extremely rapid (Shepherd, 2004).

Organizations operate in a world that is becoming increasingly dominated by digital technology. It is ingrained in the very core of many organizations' products, services, and operations. Everyday products such as televisions, watches, and automobiles now have embedded software-based digital capabilities, and organizations are routinely developing management systems of intelligent machines equipped with digital sensors, networks, and processors.The widespread adoption and innovation of digital technologies is radically altering the nature of products and services. Our culture and intellectual world is a shared experience, thanks to the reshaping and remediating of the computing world and the society in general.

Massive cultural digital objects include large-scale corpora such as the millions of books scanned by Google and others (Jacquesson 2010), millions of photos and micro-messages shared on social network services (Thusoo et al. 2010), massive geographical information systems such as Google Earth (Butler 2006), and ever-expanding networks of academic papers citing one another (Shibata et al. 2008). These interconnected objects – whether born digitally or reconstructed via digitization pipelines – are too large to read or watch. A single scholar confronted with a single document in the traditional 1:1 ratio cannot cope with such abundance.

2. Open Access

The open access (OA) publication movement aims to make research literature available to the public for free and without restrictions. While the movement's primary goal is to democratize access to scholarly literature, it is unclear whether OA has uniformly democratized the corpus of freely available research, or whether authors who choose to publish in OA venues represent a specific subset of scholars—those with access to resources that allow them to afford article processing charges (APCs).

The open access movement's themes are research integrity, transparency, and accessibility. Indeed, research funding agencies are increasingly encouraging or mandating publication in

open access venues, as well as the dissemination of code, data, and methods in open repositories (see, for example, http://roarmap.eprints.org).

The most common forms of OA are referred to as a series of colors: Bronze (the article is free to read on the publisher's website but no explicit license is presented); Green (the article is available in a repository, self-archived by the author); Gold (all articles in the journal are OA); and Hybrid (all articles in the journal are OA) (individual articles are OA if the authors have paid a publication fee, but other articles in the journal are closed). While Bronze and Green account for the vast majority of OA publications (Piwowar, Priem, et al., 2018), Gold and Hybrid are distinct in that they reflect an author's deliberate decision to make their article immediately publicly available at the time of publication, often at the expense of an APC. We ask two specific questions in light of these various OA types: (a) What characteristics distinguish authors who intend to publish openly immediately (i.e., who chooses to publish OA articles), and (b) Which authors are ultimately represented in the OA literature, regardless of the means or type of OA? (Olejniczak, 2020)

Digital scholarship in open science practices is to be valued. This is something that a researcher cannot do on their own. They require adequate infrastructure, skills, funding, and even discipline-specific training to ensure that their data is published in a FAIR manner (findable, accessible, interoperable and reusable). Indian academia and researchers need that support. Open access is also a natural continuation of a trend where computer savvy researchers have used electronic networks to spread their work electronically.

3. Digital Humanities

The humanities are interested in the words and images of the past and present. Not only are primary texts, images, and archaeological finds represented in digital archives in a postdigital age, but so are scholarly texts describing and discussing these, as well as all other contemporary mediated communication. We are increasingly interacting with information technology in both our professional and personal lives, using computer interfaces to read and watch, research and compare, and write and communicate.

Defining the nature and boundaries of digital humanities is a long-debated and unresolved issue (Terras et al. 2013), not only because there is no agreement on this question, but also because digital humanities are undergoing a profound transformation that necessitates a reconsideration of its fundamental concepts (Gold 2012). For many years, the term "digital humanities" has been used to refer to computational approaches to humanities research problems as well as critical reflections on the effects of digital technologies on culture and knowledge (Schreibman et al. 2008). The field of digital humanities evolved from the field of humanities computing, which began in the 1940s and 1950s with the pioneering work of Jesuit scholar Roberto Busa, who began in 1946, and English professor Josephine Miles, who began in the early 1950s.

Literature has played an important role in the evolution of the digital humanities. With open access publishing, open access resources, digital archives, and other initiatives making scholarship and pedagogy so easily accessible to the public, digital humanities plays a critical role in the study and teaching of literature. Many digital humanities projects originated in English departments. Data analysis, data mining, databases, visualization, and text archiving are a few examples. Text analysis tools can be used to investigate recurring words, patterns, or themes in novels. Researchers can investigate changes in different editions of a work, examining what was added or changed in each edition to raise questions about why those changes occurred at the time. In his article titled "What is Digital Humanities and What's It Doing in English Departments?" which appeared in the ADE Bulletin (2010), Mathew Kirschenbaum cites the following reasons why literature and and digital humanities are so closely connected:

1. It is easier for computers to manage text data more than images, audio, video, or other data forms. Data mining makes it easier for literary scholars to study texts and generate new information.

2. Computers and composition have been associated together for a long time.

3. It is easier to create electronic archives and electronic editions of texts.

4. Literary projects involving hypertext and other electronic forms are more diverse and vibrant.

5. English Departments are more open to cultural studies where computers and other digital resources are used heavily for analysis.

6. With the proliferation of e-book devices and the interest in e-reading, it is easy to perform data mining and "distance readings" of millions of books at a time.

The identified gaps can be found in almost all infrastructure areas. Transparency is perceived to be lacking in monograph publishing, particularly in quality assurance. There should be more emphasis on best practices in open access book publishing. Authors must be better supported and given the option to publish their research outputs open access, while funders and institutions must improve infrastructure to support policy compliance and monitor research outputs. There are flaws in the use of metadata and standards for open access books that limit interoperability and discovery, affecting the OA book ecosystem as a whole. The recommendations are broadly applicable to the broader stakeholder community and can be considered in the context of policy development and measures to improve the infrastructure for open access books. Berry & Fagerjord, 2017 have suggested that a way to reconceptualize digital humanities could be through a "digital humanities stack"

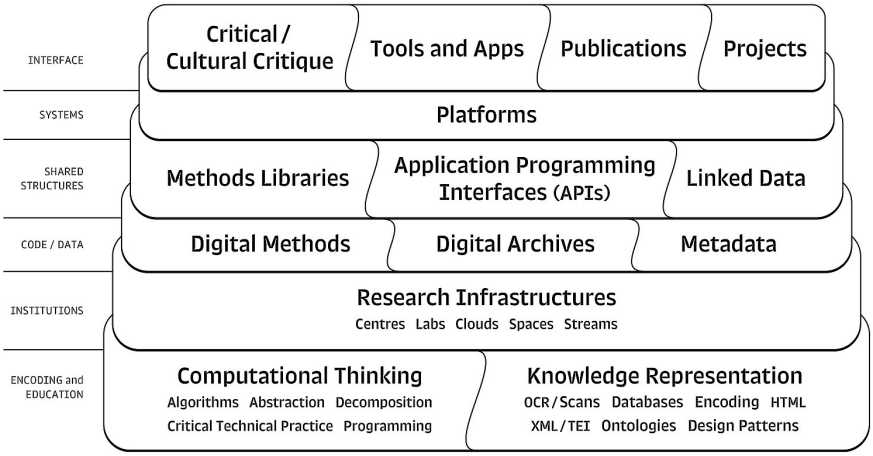

Image 1: Digital Humanities Stack (Berry & Fagerjord, 2017)

3.1 Text Analysis in DH Projects:

DH Projects support the humanities and cultural sciences working with digital resources and methods in research and teaching. For this purpose the association is developing a digital research infrastructure for tools and research data and is developing materials for teaching and further education in the field of digital humanities.

3.1.1 Computational Textual Analysis (CTA) is used widely in DH projects. CTA's primary value is that it enables the scale of traditional text analysis to expand. Whereas scholars were formerly limited to analyzing only one or a handful of texts at a time--a method often known as "close reading"--scholars via CTA can now also analyze thousands at once to identify large-scale patterns and trends--a method known as "distant reading." When close and distant reading are paired, scholars can make more informed generalizations.

Tool	**Features**
Voyant	Voyant is a web based set of tools for reading and analyzing digital text. You can plug corpora of websites or files into Voyant for textual analysis. Note that Voyant uses many tools to do different types of analysis, including topic modelling, concordance and even some network analysis.
R	R is a powerful and popular language and environment for statistical computing and graphics. R is particularly useful for technical analysis when coupled with the TM (text mining) package.
MALLET	MALLET is a tool for analyzing text through topic modelling.

stylo	A package for R to perform stylometric analysis. This package was used to generate the authorship results for the Hildegard of Bingen project in the Example Projects box on this libguide.
AntConc	AntConc is used to perform concordance studies of texts.
Google NGram Viewer	A tool by Google to study over 5 million of the books in google books up to the year 2008.
TEI (Text Encoding Initiative)	A standard for representing digital texts, TEI provides a way to mark up text for machine readability. This can improve the understanding of and things you can do with digital texts.

Table 1: CTA Tools

3.1.2 DH projects in India

Successful DH projects in India include for textual studies, the digital variorum of Tagore's works in "Bichitra" done by the School of Texts and Cultural Records at Jadavpur University; the Dutch and Scottish cemeteries project at Presidency University; Two Centuries of Indian Print with the British Library and Jadavpur; KSHIP at IIT Indore; Indiancine.ma and Pad.ma. It is important to note that the Indian definition of DH is far from derivative of a North American definition for these reasons, as well as the fact that much of Indian DH occurs in sites complementary to but outside of university spaces. Its meaning is still forming in India, so this work is offered as an exploration of what DH might be. In the end, access to digital tools and resources that have been used across fields to impact humanities work in a new way determines what we call DH in India, but the field as a whole is still emerging and dispersed across several sites (Todd, 2021).

3.1.3 DH Projects in Europe

Major DH projects in Europe given in the table below:

Origin	Project	Link
ERIC DARIAH and CLARIN ,French Node	Huma-Num	https://www.huma-num.fr/quest-ce-que-la-tgir-huma-num/
Institute for Computational Linguistics, National Research Council of Italy	CNR ILC	http://www.ilc.cnr.it/
DARIAH-DE (Digital Research Infrastructure for the Arts and Humanities), Germany	DARIAH-DE - Digital Research Infrastructure for the Arts and Humanities	https://www.dariah.eu/

Table 2: Major DH Projects in Europe

4. Research Infrastructures

At the most basic level, scholars and students require access to library facilities, which serve as a centrally provided set of research facilities, in order to conduct research. The provision of digital resources and digital research infrastructures has become increasingly important in the context of the digital humanities, as well as many other disciplines across the university. However, decisions regarding the provision of these digital services can be contentious – not only can these services be costly, but they also divert funds away from existing funding priorities and research infrastructures. They are also not risk-free, and if not carefully developed, they can result in costly failures, or white elephants. However, the scope of research infrastructures as a subject of study extends beyond what purely centralized research infrastructure provision might imply. The department, for example, is

an important research infrastructure that shapes and supports the creation of a shared environment for researchers working in the same field. Similarly, within the university, the research centre frequently serves as a specifically interdisciplinary structure that transcends both departmental and disciplinary boundaries. As a result, digital humanities is frequently located in various types of structures in various institutions, but it also has different research infrastructure needs than its cognate fields in the humanities.

The growing importance of infrastructure in the digital humanities, and its relationship to knowledge creation and explanation, broadly conceived, can be traced to the notion of digital infrastructure, in particular developed through the notion of research infrastructure or 'cyberinfrastructure' outlined by Atkins et al. (Berry & Fagerjord, 2017).

Mozhaevala & Renha (2016) find in their analysis- covering various subjects, from creation of databases before reconstruction of historical interiors, 3D - modeling of large cultural objects, visualization of existential data, etc., Digital Humanities include: use of digital technologies in humanitarian researches, first of all, for processing of big data files: from the analysis of ancient manuscripts ("digital paleography") before studying of literary works and documents of various historical eras ("distant reading" or "digital reading"); research of features of a new era, sociocultural consequences of digital technologies; work with "cultural heritage: "digital art", new media, creation of digital libraries, archives, databases of cultural heritage and museum collections, the digital reconstruction demanding joint efforts of humanists and experts in digital technologies.

5. Conclusion

Fundamental questions remain like "How can we fill the gap between qualitative and quantitative analysis by using digital networks resources?" "How can we fill the gap between individual and structure when analyzing a phenomenon through digital lenses?"
But we agree with Grusin that 'digital media can help to transform our understanding of the canon and history of the humanities by foregrounding and investigating the complex entanglements of humans and nonhumans, of humanities and technology, which have too

often been minimized or ignored in conventional narratives of the Western humanistic tradition' (Grusin 2014: 89). Digital humanities needs to be critical of the 'digital' in digital humanities as much as of the 'humanities'. Indeed, critical digital humanities could help to reposition our traditional humanistic practices of history, critique and interpretation, so these humanistic traditions can help to refine and shape the direction and critical focus of digital humanities and its place in the academy (Berry & Fagerjord, 2017).

A central priority in the Industrial Revolution 4.0 is avoiding being left behind by the progress of the communications technology industry. However, as members of the networked society, we have the right to question the referral and future direction of the goals of digital technology (Mahatma, 2021). With the idea network serving as a platform for collaboration, users must continue to consume the technology. In addition to obtaining information from web providers, interaction in the cyberworld has become a need. California University UCLA, Digital Humanities interprets the cultural and social impact of new media and information technologies, as well as creates and applies these technologies to interrogate cultural, social, historical, and philological questions. The logic of development of the humanities moves today in the direction of methodological interdisciplinarity, search of general scientific methodology which would allow not simply to unite tools of separate sciences, but to develop the general bases for humanitarian and natural-science researches. In India, where a clear picture of the 'field' as such has yet to emerge in the form of a theorisation of its key concerns, it is only through practice-mapping that one can locate what are, at best, certain discursive shifts in how we understand content, structures, and methods in the humanities within the context of the digital. These changes may be visible only in a few domains, particularly in India's multi-layered technological landscape, and there is no widespread agreement on whether they represent a larger epistemic shift or new direction of thought. In the words of Shah, 2015 - the digital humanities and digital scholarship for that matter must be viewed as an opportunity to question, contest, and remap the ways in which life, labor, and language conditions are rapidly changing in the country as a result of the country's widely varied and uneven transition to digital modernity. Digital humanities can serve a dual function by building relevant and sensitive infrastructure of production and access to knowledge, as well as a

critical voice that resists the dehumanizing principles of networked societies that reduce all human beings to actors and all human modes of engagement to actions and transactions.

References

ACRL Research Planning and Review Committee. (2018). Top Trends in Academic Libraries A Review of the Trends and Issues Affecting Academic Libraries in Higher Education, *C&RL News* 79, no.6 (2018): 286–300. https://doi.org/10.5860/crln.79.6.286

Berry, D. M., & Fagerjord, A. (2019). *Digital humanities: Knowledge and critique in a digital age.* Wiley.

Bogosian, R. Creating and Sustaining an Agile Organization: The Impact of the Digital Economy (2018). *Rutgers Business Review*, Vol. 3, No. 1, 2018, https://ssrn.com/abstract=3176037

Dodd, M. & Nidhi, K. (Eds.) (2021). *Exploring Digital Humanities in India: Pedagogies, Practices and Institutional Possibilities*. Routledge.

Gitelman, L. (2013). *"Raw Data" Is an Oxymoron*. Cambridge, MA: MIT Press.

Gold, M. K. (2012). *Debates in the Digital Humanities*. Minneapolis, MN: University of Minnesota Press.

Grusin, R. (2013). The Dark Side of the Digital Humanities – part 2. *THINKING C21*. https://www.c21uwm.com/2013/01/09/dark-side-of-the-digital-humanities-part-2/

Hall, W., de Roure, D., & Shadbolt, N. (2009). The Evolution of the Web and Implications for eResearch. *Philosophical Transactions: Mathematical, Physical and Engineering Sciences*, 367(1890), 991–1001. http://www.jstor.org/stable/40485755

Kaplan, F. (2015). A map for big data research in digital humanities. *Frontiers in Digital Humanities*. https://doi.org/10.3389/fdigh.2015.00001

Kenney, M., & Zysman, J. (2016). The Rise of the Platform Economy. *Issues in Science and Technology*,61-69. http://www.jstor.org/stable/24727063

Kirschenbaum, M. (2010). What is digital humanities and what's it doing in English departments?, reprinted in Terras, M, Nyhan, J & Vanhoutte, E 2013, *Defining Digital Humanities: A Reader*, Taylor & Francis Group, http://ebookcentral.proquest.com/lib/deakin/detail.action?docID=1426876>

Mozhaeva, G. & Renha, P. M. (2016). *Digital Humanities: to a question of the directions and prospects of development of interdisciplinarity in humanitarian researches and education* [Paper Presentation] https://www.shs-conferences.org/articles/shsconf/pdf/2016/04/shsconf_erpa2016_01017.pdf

Olejniczak, A. J., & Wilson, M. J. (2020). Who's writing open access (OA) articles? Characteristics of OA authors at Ph.D.-granting institutions in the United States. *Quantitative Science Studies*, 1(4), 1429–1450. https://doi.org/10.1162/qss_a_00091

Shah, N. (2015). "Beyond Infrastructure: Re-Humanizing Digital Humanities in India", in Between Humanities and the Digital, edited by Svensson, P. & Goldberg, D. T. Cambridge: MIT Press.

Schreibman, S., Siemens, R., and Unsworth, J. (2008). *A Companion to Digital Humanities*. Malden, MA: Wiley-Blackwell.

Shepherd, J. (2004). What is the Digital Era? In Doukidis, G., Mylonopoulos, N., & Pouloudi, N. (Eds.), *Social and Economic Transformation in the Digital Era*. IRMA International. https://www.irma-international.org/chapter/digital-era/29024/

Snee, H., Hine, C, Morey, Y, Roberts, S, & Watson, H (2015). Digital methods as mainstream methodology: an introduction in Roberts, S, Snee, H, Hine, C, Morey, Y & Watson, H 2015, *Digital Methods for Social Science: An Interdisciplinary Guide to Research Innovation*, Palgrave Macmillan UK, London, UNITED KINGDOM. http://ebookcentral.proquest.com/lib/deakin/detail.action?docID=4082283>

Yudhi Mahatma, I (2019, August 13-15). Digital Capitalism: Logic of Capitalism Practice in the Context of Communication Technology in Industry 4.0 [Paper presentation]. Proceedings of the Asia-Pacific Research in Social Sciences and Humanities Universitas Indonesia Conference http://conference.ui.ac.id/aprish/aprish_2019

Wong, S. (2016). Digital humanities: What can libraries offer?. portal: *Libraries and the Academy*, 16 (4), 669-690. https://doi.org/ 10.1353/pla.2016.0045