



LITERARY QUEST

An International, Peer-Reviewed, Open Access, Monthly, Online Journal of English Language and Literature

Digital Research in Literature

Mr. Showkat Hussain

Researcher, Department of English, Maulana Azad National Urdu University,
Hyderabad, Telangana, India.

Abstract

In the recent years there has been much deliberation about the use of digital technology in arts and humanities research. Keeping in view the limitations of traditional critical expression and interpretation of complex, multi-layered or multi-faceted works of human imagination the new advancements in digital technology can be used and exploited to widen the limits of literary and other humanities research. Digital technologies give people whether readers, teachers, scholars, and critics a better representation and access to literary works. Therefore, digital technologies are increasingly important to arts and humanities research, expanding the horizons of research methods in all aspects of data capture, investigation, analysis, modeling, presentation and dissemination. Digital research in humanities is a multi-disciplinary approach which crosses boundaries between computer science and humanities disciplines such as cultural anthropology, archaeology, classics, English, history, modern languages and literatures, library science, and the

arts. The emphasis is on humanities as a whole rather than specific discipline. This use of digital technology helps in the development of new and innovative scholarly tools. With the use of these innovative research tools we can explore literary and other texts from various perspectives which were not possible in the traditional modes of research. Largely, however, the emphasis is on the humanities, and teachers/scholars may be engaged in creating new approaches to understanding the humanities through technological means. Therefore, it is need of the hour to integrate research and digital technologies for the advancement of knowledge.

Keywords

Digital; Humanities; Data Mining; Hardware; Software; Data; Text Analysis; Quantification.



Digital technology over the years has taken an important place in every walk of life and its impact is inevitable for one and all. Various technological devices have cast a remarkable impact on education and research with various devices as tools are facilitating teaching, learning and research processes. Teachers and scholars whether born before or after the booming onslaught of the age of digital convergence can/shall make use of the various facets of this technology in their teaching and research pursuits. The innovations in digital technology are novel to a large number of teachers and scholars who are 'digital immigrants', brought up in a world of paper, pen and books, while the vast majority of learners today are 'digital natives', born into a world of technology and perfectly at home with all aspects of media.

Before we proceed further a basic question remains, 'what consists digital technology'? Digital technology is separated into hardware and software. The former includes computers, video and webcams, digital video recorders, digital cameras, mobile phones, overhead projectors and interactive

whiteboards. The latter are programmes and applications which run on and through the former. It includes operating systems, applications, database programs, multimedia programs, wikis, blogs, podcasts, power points, word processing documents, and websites. The effective use of digital technology in research in humanities needs some elaboration here. Digital research in humanities traverses the boundaries between computer science and the various disciplines of humanities like cultural anthropology, archaeology, history, classics, languages and literatures, library science, and the arts. The development of new scholarly tools using the multi-sensory domains of recent digital technology in humanities brings humanities closer to computer sciences. Through the use various digital devices and programmes, the emphasis is on the humanities, with a view to engage teachers and scholars to create new approaches to understanding the humanities through technological means.

The increased use of digital technologies has led, in almost all the spheres of knowledge, to fast changes in the form and conduct of scholarly activity. However most of the people who are pursuing more traditional forms of work are still unaware about the digital scholarship in humanities. Digital scholarship, often interdisciplinary in form and substance, possesses a technical component and is generally pursued through collaborative efforts. Digital research relies upon teamwork with research teams comprising scholars, technical experts, archivists, librarians, and statisticians who are more familiar with the scope, method, and contributions of digitally-based inquiry in the humanities. Developing new research tools/methods for scholarship and teaching in the digital age and also applying these tools/methods to humanities and other art forms is the essence of digital research in humanities.

Digital research, like other traditional forms of research, has three stages starting from techniques of data generation (such as multisensory fieldwork

and participant observation, video and photography, audio recording), followed by data analysis (including software and digital techniques for image and sound analysis, discourse and textual analysis, thematic coding of transcripts) and presentation and dissemination of research findings (such as film, video and photography). Most of the humanities scholarship consists of the interpretation of source material (texts). It is about analysing and explaining those texts with a view to find connections between them. In traditional modes of research the outcome of analyses and interpretations is published as journal articles and monographs and the subsequent scholars finding limitations with the earlier works would challenge these interpretations and come out with their own explanations with reference to the earlier ones. Humanities scholarship is thus firmly grounded in critical thinking and assessment. In making electronic resources work for humanities scholarship, ways must be found to facilitate this kind of critical thinking and to represent multiple views of the same material.

Digital technology used in the preparation and presentation of new scholarly editions of works, whether literary or other, with novel interpretations can be an important aspect of a fruitful research in humanities. “Traditional critical editing, defined by the paper and print limitations of the codex format, is now considered by many to be inadequate for the expression and interpretation of complex, multi-layered or multi-text works of the human imagination” (Degan 1). So it becomes imperative, because of the significant achievements in the use of digital technology, to make more and more texts digitized. This possibility of online form of texts offers scholars, teachers, readers and critics a better access and representation of texts of their interest. Sutherland in “Being Critical: Paper-based Editing and the Digital Environment” (13-26) suggests that production of material in electronic form can hamper the kinds of engagement prompted by print versions. Besides this we have to think what purposes electronic editions may serve and what

assumptions about texts underlie their production. We urgently need an electronic equivalent of textual theory that will take into account the essential difference between print texts and digital texts. The change of texts from print to electronic variants may fail to retain an essential aspect – their ‘bookishness’ – in the electronic format. Traditionally, most often the reading experience looks inextricably related to the physicality of the book, of text as an object; and therefore, assumptions about the advantages of electronic editions have left such issues unexamined. Paul Eggert in an article “The Book, the E-text and the Work-site” (63-82) argues that there must be some clear way of authenticating electronic editions and preserving their integrity, just as in the print world the fixed and stable nature of the book preserves the integrity of the work. He proposes the use of what has come to be known as ‘just-in-time-markup’ (JITM) to ensure the accuracy and authenticity of the electronic text. This system runs counter to common practice in markup, where tag sets are inserted into the text and travel along with it when it is transmitted or transformed. JITM keeps markup and texts separate, and any corruptions or changes in the text are detected instantly using algorithmic methods (checksums) to keep track of even the slightest difference.

Digital imaging, currently in vogue, can also make unprecedented contribution to humanities research. It includes the procedures like photo manipulation, scanning, colour correcting and RIPing which not only helps in filling, storage, processing, preservation and retrieval of material, but also enables remote access to material. The recent developmental leaps in compression techniques have made it possible to undertake large-scale high resolution imaging projects. Manuscripts and other older texts, which were previously illegible, can be made readable by using image enhancement and morphing techniques. “Digital facsimiles of astonishing fidelity are now being produced, such as the ‘Turning the pages’ resources at the British Library, including the Lindisfarne Gospels and the Sherborne Missal. This means that

access to fragile or rare originals can be restricted, ensuring their preservation for the long term. It is also possible to carry out enhancement of digital images of originals, making lost readings visible again” (Griffin 215). However images need text in order to be useful. Text in the form of descriptors or metadata is necessary to help locate the image. Annotations to images must be in the form of text and text is also needed to explain why an image is linked to something else.

Digital Research in humanities is an amalgamation of various branches of knowledge, disciplines, techniques, methods and methodologies which are used to explore questions and issues of humanities by combining cohesively various multimedia formats. From digitizing texts, digital editing and translation, image processing, animation and novel publishing tools to techniques for extracting and analyzing, interpreting and disseminating data, the methods used in digital research in humanities are abundant and interconnected.

Data mining, a fast growing field, deals with problems of scalable data analysis by paying close attention to computing, storage, communication and human-factor-related issues. Data mining tools enable researchers to sift through and search for patterns across large amounts of data through tasks such as frequent pattern discovery, clustering and classification. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many different dimensions or angles, categorize it, and summarize the relationships identified. Technically, data mining is the process of finding correlations or patterns among dozens of fields in large relational databases.

Tools of citation management help record bibliographic information about research materials, manage indexes, create bibliographies, add footnotes, and manage research collections. It also makes sharing of resources between researchers an easy task.

Text analysis softwares are application softwares which can be used to ascertain the recurrence of words and phrases, idioms used in a text, create concordances, study words, contexts and patterns in texts. Antcon3.2.1(a concordance program), Attensity (text analytic and data extraction framework), CATMA (Computer Aided Textual Markup and Analysis), ClearForest (text tagging, extraction and analytic software), DICTION5.0 etc are some important tools of digital textual analysis. “Frequency and keyword analysis involves the construction of word lists, using Automatic computational techniques, which can then be analyzed in a number of ways, depending on one’s interest(s). For example, a researcher might focus on the most frequent lexical items of a number of generated word frequency lists to determine whether all the texts are written by the same author. Alternatively, they might wish to determine whether the most frequent words of a given text (captured by its word frequency list) are suggestive of potentially meaningful patterns that they might have missed had they read the text manually” (Qtd. in Archer 2). In digital research in humanities a researcher may find interest in keyword analysis, collocation and /or colligation where he/she will compare word frequency list of his/her chosen text (let’s call it text A) with the word frequency list of another reference text (let’s call it text B) with a purpose to identify words that are frequent and also words that are infrequent in text A, when compared to text B. For example in Zulfikar Ghose’s novels like *A New History of Torments*, *Figures of Enchantment*, *Don Bueno* and *The Triple Mirror of the Self* the recurring setting in the interiors of South America, and the repeated use of words and phrases like east, west, north, south, carcasses, bodies, Indian tribes, buzzing engines, labyrinthine rivers, incest, torment, flight into the interiors of jungles, search for something, civilization, Amazon, unpredictability of events, unknown fears, savage state, uninhabited wilderness, sea etc give an idea about the larger thematic concerns in his corpus of writing. This has the advantage of removing words that are common to all the texts, and so allows

the researcher to focus on those words that make text A distinctive from text B (and vice versa).

Text analysis and text mining processes help in the gathering desired data and information from texts such as poems, novels, treatises, monographs, articles, speeches, historical documents, web pages, etc. Text analysis tools very easily help in searching through a large body of texts, to develop word clouds or find word frequency, to more complex tasks like identifying patterns in parts of speech or identifying sentiments, moods, and emotions in a corpus. Some available software supports morphological analysis, natural language processing, tagging, transcription, building concordances, speech analysis, and other linguistic research.

Networking sites allow researchers to discover and connect with other people working in their field, as well as publicize their own research interests and accomplishments via available downloads, applications, and plug-ins that allow us to take notes, share them with other researchers, attach them to digital resources, and more.

One of the most debated issues pertaining to the possibility/impossibility of quantitative research in literature can be answered by the work in the context of computers and literary studies and the recent developments in digitizing techniques; and proliferation in the availability of electronic texts, increasingly sophisticated statistical techniques, and the advanced powerful computers have produced much more accurate and persuasive analyses.

“Quantitative approaches to literature represent elements or characteristics of literary texts, numerically applying the powerful, accurate, and widely accepted methods of mathematics to measurement, classification, and analysis” (Hoover, N.P.). With the current developments in research methodologies and the availability in large number of electronic literary texts, the traditional modes of reading and analysis are no more sufficient as tools of research. With an increased demand for objective and unbiased research, the

scientificness of quantitative approaches offers innovative ways of textual analysis. Other than addressing the problems of authorship and style, quantitative approaches “can be used to investigate larger interpretative issues like plot, theme, genre, period, tone and modality” (Hoover, N. P.).

Reliable identification and quantification of items, features or characteristics of a text like frequency of punctuation marks, repetition of certain letter sequences, specific sentence structures can be used in authorship attribution. David L. Hoover in “Quantitative Analysis and Literary Studies” while discussing about the methods of modern quantitative studies in literature elaborates:

Words themselves, as the smallest clearly meaningful units, are the most frequently counted items, and syntactic categories (noun, verb, infinitive, superlative) are also often of interest, as are word n-grams (sequences) and collocations (words that occur near each other). Thematic or semantic categories (love words, words related to time), while more difficult to count, have the advantage of being clearly relevant to interpretation, and automated semantic analysis may reduce the effort involved. Phrases, clauses, syntactic patterns, and sentences are often been counted, as are sequences or subcategories of them (prepositional phrases, subordinate clauses, passive sentences). Many of the items listed above are also used as measures of the lengths of other items: word length in characters, sentence or clause length in letters or words, text length in words, sentences, paragraphs, and so forth. Nonlinguistic textual divisions ranging from small units like lines and couplets to larger structural units like paragraphs, stanzas, scenes, acts, and chapters can also sometimes be usefully counted, as can literary categories like narrators and characters (including subcategories like first-person and third-person narrators, and characters

divided by age, ethnicity, nationality, class, and gender), and plot elements (marriages, deaths, journeys, subplots (N.P)).

With the help of modern digital computing techniques we can count various items and features of single literary text as an object of study or make comparative analyses of texts, dialogues, narrators, genres to interpret and show how writers differ in their vocabularies by finding the raw counts of the numbers of different words (word types).

Many kinds of studies of literary texts use quantitative methods. Quantitative thematic analysis can trace the growth, decay, or development of vocabulary within a thematic domain, or study how authors differ in their expressions of a theme. Many empirical studies of literature translate readers' judgments into numerical scales to study literary response using techniques borrowed from the social sciences. Metrical analysis, because of the inherent reliance of meter on pattern, is a natural area for quantitative study, though there has been less research in this area than one might have expected.

Attributing authorship to anonymous texts by the use of statistical stylistics can be explored as one of the areas of digital analysis of literature. Forensic authorship attribution which also uses certain scientific digital techniques to ascertain the facts is somewhat different from the literary authorship attribution which bases on linguistic and other stylistic features of a text. "Determining who wrote a text generally requires much the same methodology whether the text is ransom note, a threatening letter, a legal opinion, the federalist papers, a contemporary political novel, an anonymous eighteenth-century verse satire, or play by Shakespeare" (Hoover, N.P).

Therefore, the integration of research and digital technologies, although a challenging enterprise for many researchers, teachers and learners, is to be explored and used for the advancement of knowledge. When mobile phones and other digital devices have sneaked into classrooms in the students' backpacks,

it is only wise to adopt and use them as teaching-learning and research tools rather than banning and avoiding them.

References:

Archer, Dawn. Ed. *What's in a Word-List?: Investigating Work Frequency and Keyword Extraction*. Surrey: Ashgate Publishing Limited, 2009. Print.

Chambers, Ellie. "Computers in Humanities Teaching and Research: Dispatches from the Disciplines." *Computers and the Humanities*, 34.3 (2000): 245-254. JSTOR. Web. 20 Nov. 2013.

Deegan, Marilyn and Kathryn Sutherland, Ed. *Text Editing, Print and Digital World*. Surrey: Ashgate Publishing Limited, 2009. Print.

Eggert, Paul. "The Book, the e-Text and the Work-Site." *Text Editing, Print and the Digital World*. Eds. Marilyn Deegan and Kathryn Sutherland. Surrey: Ashgate Publishing Company, 2009. Print.

Hoover, David L. "Quantitative Analysis and Literary Studies". *A Companion to Digital Literary Studies*. Ed. Susan Schreibman & Ray Siemens. Oxford: Blackwell, 2008. Web. 20 Nov. 2013.

Scholes, Robert and Clifford Wulfman Source. "The Changing University and the Humanities" *South Atlantic Review*, 73.4 (2008): 50-66. JSTOR. Web. 20 Nov. 2013.

Short, Harold and Marilyn Deegan. "ICT as a Research Method". *Research Methods for English Studies*. Ed. Gabriele Griffin. New Delhi: Rawat Publications, 2007. 213-235. Print.

Siemens, Raymond G. "A New Computer-Assisted Literary Criticism?" *Computers and the Humanities*, 36.3 (2002): 259-267. JSTOR. Web. 20 Nov. 2013.

MLA (7th Edition) Citation:

Hussain, Showkat. "Digital Research in Literature." *Literary Quest* 2.1 (2015):
116-127. Web. DoA.

DoA – Date of Access

Eg. 23 Aug. 2015. ; 05 April 2017.