Database & Digital Humanities

Youngmin Kim
Professor, Department of English
Director, Institute of Trans Media World Literature
Dongguk University, Seoul, Korea
Jack Ma Chair Professor of Ma Yun Education Fund,
Hangzhou Normal University, PRC
yk4147@daum.net



2017-20 "The Aesthetics of Trans Media, World Literature, and the Digital Humanities: Methodological Representations of Postmodern Sublime"

[Global Research Network Project endowed by the National Research Foundation of Korea]

Literature as Database & Digital Humanities

- I. Introduction: On the Wave of the 4th Industrial Revolution
- II. Database
- **III. Perspective**
- IV. Method
- V. Database and Narrative in DH and World Literature
- **IV. Conclusive Suggestions**
- VI. Digital Bibliography



I. Introduction: On the Wave of the 4th Industrial Revolution

"We stand on the brink of a technological revolution that will fundamentally alter the way we live, work, and relate to one another. In its scale, scope, and complexity, the transformation will be unlike anything humankind has experienced before. We do not yet know just how it will unfold, but one thing is clear: the response to it must be integrated and comprehensive, involving all stakeholders of the global polity, from the public and private sectors to academia and civil society."

Klaus Schwab, "The Fourth Industrial Revolution: what it means, how to respond"

- ✓ First Industrial Revolution: Steam Engine: 18th to 19th centuries in Europe and America
- ✓ Second Industrial Revolution: Electricity: 1870 and 1914, just before World War I:
- ✓ Third Industrial Revolution: Digital Technology: 1980s.
- ✓ Fourth Industrial Revolution: Communication and Connectivity: Emerging technology breakthroughs

And now on the wave of The 4th Industrial Revolution, this field of the Digital Humanities transforms itself into complex system of "connectivity."

Lauren F, Klein and Matthew K. Gold in their 2016 essay, "Digital Humanities: The Expanded Field," remind us of the three V's in the Fourth Wave of the Industrial Revolution: Volume, Velocity, and Variety (one can add more: Variability and Complexity now).

"Along with the digital archives, quantitative analyses, and tool-building projects that once characterized the field, DH now encompasses a wide range of methods and practices: visualizations of large image sets, 3D modeling of historical artifacts, "born digital" dissertations, hashtag activism and the analysis thereof, alternate reality games, mobile makerspaces, and more. In what has been called "big tent" DH, it can at times be difficult to determine with any specificity what, precisely, digital humanities work entails."

(http://dhdebates.gc.cuny.edu/debates/2)

- ✓ In fact, the data sets have transformed its complexity and variety into convergence, becoming bigger volume in scope and more rapid in speed.
- ✓ As a result, the spectrum of digital humanities covers enormous convergent fields of big data:

augmented reality (AR), coding, hypertext, crowdsourcing, gamification, hashtag, metadata, new aesthetic (NA), open access (OA), open source, QR Codes, Social Media, Text Encoding Initiative (TEI) visualization, data mining, text mining, GIS(Geographic Information Systems), and digital publishing, to name a few.

II. Database:

"Big data usually includes data sets with sizes beyond the ability of commonly used software tools to capture, curate, manage, and process data within a tolerable elapsed time. Big data philosophy encompasses unstructured, semi-structured and structured data, however the main focus is on unstructured data. Big data "size" is a constantly moving target, as of 2012 ranging from a few dozen terabytes to many exabytes of data. Big data requires a set of techniques and technologies with new forms of integration to reveal insights from datasets that are diverse, complex, and of a massive scale." (https://en.wikipedia.org/wiki/Big data).

Structured data:

- the field of "forms of integration" waiting for us to glean "insights."
- -organized pattern of texts, images, sound files, or other digitally encoded information.
- the explicit formal properties based upon the secondary organizing principle of encoding.

Unstructured data:

- the data without the secondary structure imposed upon it.
- All data is said to be structured, as far as it has the potential to be provided with useful information.

"The use of database technology among humanists has been invigorated by the realization--common, perhaps, to many other similar convergences--that a number of fascinating problems and intellectual opportunities lurk beneath these apparently practical matters.

The inclusion of certain data (and the attendant exclusion of others), the mapping of relationships among entities, the often collaborative nature of dataset creation, and the eventual visualization of information patterns, all imply a hermeneutics and a set of possible methodologies that are themselves worthy objects for study and reflection."

Stephen Ramsay, "Database":

(http://www.digitalhumanities.org/companion/)

- ✓ Since the 1960's, computer processing of textual data in textual and literary field has expanded enormously in intersection with digital technology in the forms of electronic texts:
 - 1) electronic transcription of a literary text in the form of a computer text file;
 - 2) encoded text in the form of the Standard Generalized Markup Language (SGML) or the Text Encoding Initiative (TEI);
 - 3) hypertext.
- ✓ Employing these electronic texts, digital humanities incorporates both digitized (remediated) and born-digital materials from traditional humanities disciplines including history, philosophy, linguistics, literature, art, archaeology, music, and cultural studies, and social sciences.

III. Perspective: Ezra Pound's Periplum, Franco Moretti's Distant Reading, & Gayatri Spivak's Supplementing, Pascale Casanova's Republic of Letters

1. Ezra Pound's Periplum a

"The point of Imagisme is that it does not use images as ornaments. The image is itself the speech. The image is the word beyond formulated language . . . In a poem of this sort, one is trying to record the precise instant when a thing outward and objective transforms itself, or darts into a thing inward and subjective." Pound's creative writing process manifests "the perpetual renewal of language," which develops later into an Ezra Poundian vortex, "from which, and through which, and into which, ideas are constantly rushing" (*Gaudier-Brzeska*, 1916, 92).

1. Ezra Pound's Periplum b

"Periplum, not as land looks on a map/ But as seabord seen by men sailing" (Ezra Pound, *Cantos* LII-LXXI lix. 83).

"A periplum is a map or drawing that that shows how land looks from a point at sea. That is to say that a cartographer often draws maps from a bird-eye view and not from the perspective as the land would actually appear from the crow's nest or deck of a ship. Therefore, a periplum would, theoretically, be drawn as if the cartographer were out to sea so that sailors could know which land or port they were approaching" (Carroll F. Terrell, "Chinese Characters as Traveling Metaphors in Canto LXXVII of Ezra Pound's Pisan Cantos," *A Companion to The Cantos of Ezra Pound*, Berkeley, U of California P, 1980).

In fact, this perspective provides an alternative way of close reading, that is, distant reading.

2. Franco Moretti's Distant Reading: "Less is more": a

Distant reading, except in its embryonic state, has never been a case of denying close reading in its entirety, but indeed it is an advancement of it, being designed to function in reverse direction.

"we know how to read texts, now let's learn how not to read them (Conjectures, 50).

"where distance is . . . not an obstacle, but a specific form of knowledge: fewer elements, hence a sharper sense of their overall interconnection" (Graphs, 2).

2. Franco Moretti's Distant Reading: "Less is more": b

"Literary history will become second hand: a patchwork of other people's research, without a single direct textual reading: the distance from the text . . . the trouble with close reading (from the new criticism to deconstruction) is that it depends upon an extremely small canon. It is a theological exercise—very solemn treatment of very few texts taken very seriously."

Moretti's distant reading is, in fact, "is a condition of knowledge" which "allows us to focus on units that are much smaller or much larger than the text, for example, devices, themes, tropes—or genres and systems. Between the very small and the very large, the text itself disappears: Less is more. If we want to understand the system in its entirety, we must accept losing something. We always pay a price for theoretical knowledge: realilty is infinitely rich: concepts are abstract, are poor."

(Franko Moretti, "Conjectures on World Literature," New Left Review 1 Jan-Feb 2000).

3. Gayatri Spivak's Supplementing

"[T]he exact shape of a place that is empty in what is to be supplemented, zooming out, but not in competition with zooming in, always crossing the border" in the newly resurrected form of "world literature."

"We might, then, as a globally dispersed and diversified collectivity, supplement that seemingly practical will to hold the world in a grid. I don't mean survey courses; I mean the presuppositions of world literature. Supplementing, remember, is to figure out the exact shape of a place that is empty in what is to be supplemented, zooming out, but not in competition with zooming in" (468).

Gayatri Spivak, in a debate with David Damrosch, delivered at The American Comparative Literature Association (ACLA) Convention in Vancouver, Canada, on April 2, 2011 [the transcription of the discussion was published in *Comparative Literature Studies* 48, no. 4 (2011): 481.]

4. Pascale Casanova. "The Republic of Letters"

- ✓ Henry James's "The Figure of the Carpet" from the right angle: the carpet will suddenly present the attentive observer with "the one right combination" of "superb intricacy"--an ordered set of motifs which can only be understood in relation to each other, and which only become visible when perceived in their totality, in their reciprocal dependence and mutual interaction.
- ✓ Concerning the Persian carpet metaphor: starting from a grasp of the overall pattern of the designs, it will be possible to understand each motif, each colour in its most minute detail; that is, each text, each individual author, on the basis of their relative position within this immense structure.
- ✓ [Casanova's] project is to restore the coherence of the global structure within which texts appear, and which can only be seen by taking the route seemingly farthest from them; through the vast, invisible territory which [Casanova] called the "World Republic of Letters." But only in order to return to the texts themselves, and to provide a new tool for reading them.

IV. Method: Nirvana Tanoukhi's Poetics of Distance

- "I want to consider why the comparative method, in the first instance, made a *cartographic* claim to scale. Why dedicate a discipline to the task of charting zones, paths, and crossroads obscured by strict adherence to "national traditions"—when logically, comparison depends for its existence on the entrenchment of nation-based geography?" (599)
- "From the perspective of a human geographer, the fact that the distance between two adjacent neighborhoods of distances dually: by differentiating places qualitatively and demarcating boundaries quantitatively." (603).
- "If we can indeed imagine a literary history that is entangled in the history of the production of space, it is time for comparative literature to develop both a critique of *scale*, which would examine the spatial premises of comparison—and, eventually, a phenomenology of scale which would help us grasp the actually existing landscapes of literature." (605)

IV. Database and Narrative in DH and World Literature: Distancing, Scale, Hermeutics and Representation

- ✓ Structured data beyond natural language processing (NLP) data is a crucial aspect of Digital Humanities work, providing the standards in data formats for the purpose of making it possible for digital humanists to extract, search, analyze, find, and style the data in files.
- ✓ As Jerome McGann perceives, the vast corpora of the traditional texts have been waiting to be "marked up" into digital forms, based upon "digital markup schemes" of "electronic encoding systems," such as SGML (Standard Generalized Markup), XML, and its derivatives, TEI (Text Encoding Initiative), and HTML.
- ✓ These markup schemes in terms of record keeping and simulating functions of traditional textual system on the web and elsewhere were designed to challenge against larger-scale electronic publishing.
- ✓ I will demonstrate how structured data have been materialized, including text encoding, Mark-Up, TEI, GIS mapping, visualization, network analysis, and other fields of digital humanities, by extracting and quoting directly from the examples of the platforms under consideration.

A. Structured Data displayed in a platform:

Follow links to Whitman's published work, material written in Whitman's hand (including poetry manuscripts and newly discovered scribal documents), biography and correspondence in our Life and Letters section, pictures and sound, and commentary (including contemporary reviews and recent scholarship).

1) Walt Whitman:

http://www.whitmanarchive.org/
.virginia.edu/edc/default.xqy



Here users will find the six American editions of Leaves of Grass published in Whitman's lifetime as well as the so-called deathbed edition of 1891–92. In addition to Leaves of Grass, we make available several other of Whitman's books, both poetry and prose, including Drum-Taps, Sequel to Drum-Taps, Memoranda During the War, and Complete Prose Works. Also available are two British editions of Whitman's poetry, produced by William Michael Rossetti and Ernest Rhys in collaboration with Whitman. Whitman's novel Franklin Evans was originally published as a supplement to the New World in 1842 and is therefore included under "Periodicals" below. Whitman's 1852 novel Life and Adventures of Jack Engle, first published serially in the Sunday Dispatch, is currently being edited for publication on the Whitman Archive.

Periodicals

This section of the Archive features page images and full-text transcriptions of Whitman's writings that first appeared in periodicals. These include Whitman's periodical poems; his short fiction and his novel Franklin Evans; and selections from his journalism.

Translations

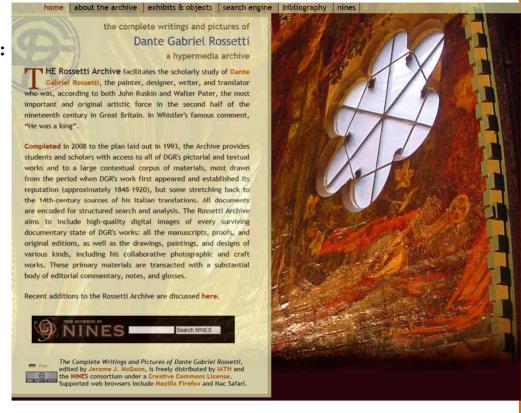
The Translations section of the Archive includes several book-length translations of Whitman's work as well as more than thirty different translations, in five languages, of Whitman's poem "Poets to Come."

A. Structured Data displayed in a platform:

All documents of The Rossetti Archive are encoded for structured search and analysis, including high-quality digital images of every surviving documentary state of DGR's works: all the manuscripts, proofs, and original editions, as well as the drawings, paintings, and designs of various kinds, including his collaborative photographic and craft works, with a substantial body of editorial commentary, notes

2) Dante Gabriel Rosetti:

http://www.rossettiarchive.org/index.html



A. Structured Data displayed in a platform:

Title: Emily Dickinson's Correspondences: A Born-Digital Textual Inquiry,: edited by Martha Nell Smith and Lara Vetter, with Ellen Louise Hart as consulting editor: The poems of Emily Dickinson: This XML-based archive includes seventy-four poems and letters from Emily's correspondence with her sister-in-law and primary confidante, Susan Dickinson with a digitized scan of the holograph manuscript. These images have zoom functionality as well as a special light-box feature that allows users to view and compare constellations of related documents.

3) Emily Dickinson:

https://rotunda.upress.virginia.edu/ed c/default.xqy

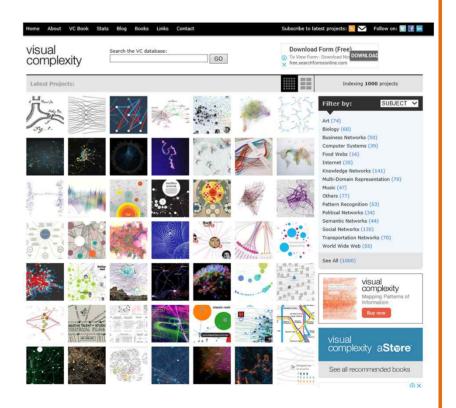


B. Visualization

- Graphic qualities of visualizations reveal rhetoric of the texts.
- Information visualizations are used to **make quantitative data legible**, useful for large amounts of information and for making patterns in the **data legible in a condensed form**. (Johanna Drucker, "Introduction to the Digital Humanities)
- VisualComplexity.com intends to be a unified resource space for anyone interested in the visualization of complex networks. The project's main goal is to leverage a critical understanding of different visualization methods, across a series of disciplines, as diverse as Biology, Social Networks or the World Wide Web.
- Not all projects shown here are genuine complex networks, although all projects have one trait in common: the whole is always more than the sum of its parts. Visualization of complex networks and mapping the dynamics of information diffusion in Blogspace. Scale-Free networks.

5) Visual Complexity:

http://www.visualcomplexity.com/vc/



C. Network Analysis 7) Republic of Letters: http://republicofletters.stanford.edu/



The world of scholarship relied on its own networks: networks of correspondence that stretched across countries and continents; the social networks created by scientific academies; and the physical networks brought about by travel, facilitating the dissemination and the criticism of ideas, the spread of political news, as well as the circulation of people and objects.

Mapping the Republic of Letters through the development of sophisticated, interactive visualization tools aims to create a repository for metadata on early-modern scholarship, and guidelines for future data capture.

D. GIS (Geo-Spatial Information Systems)

- GIS (Geo-Spatial Information Systems) is built in digital environments. Many activities and visual formats that are integral to digital humanities: timelines, diagrams, tables and charts, maps.
- -Map shows homicide hotspots in medieval London Interactive map reveals the horror—and the patterns--of murder in 14th-century London.
- This map shows the 142 murders that were committed in London from 1300 to 1340.

8) Strange maps:

http://bigthink.com/blogs/strange-maps



In world literature, GIS Analysis plays an important role in integrating the many activities and visual formats such as timelines, diagrams, tables and charts, maps to build digital environments of literary texts of world literature, when one applies the following map of the world to the world literature studies:

- ✓ History of cartography on wikipedia
 http://www.press.uchicago.edu/books/HOC/index.html;
- ✓ Google Earth: https://www.google.co.kr/intl/ko/earth/)
- ✓ Sara McLafferty: Animal City. "re-humanize"
- ✓ http://www.stanford.edu/group/spatialhistory/cgi-bin/site/viz.php?id=397
- ✓ Strange maps: http://bigthink.com/blogs/strange-maps
- ✓ Weird Maps:
 http://www.guardian.co.uk/commentisfree/interactive/2012/sep/07/weird-maps-to-rival-apple-in-pictures
- ✓ Pleiades: http://pleiades.stoa.org/home O
- ✓ Orbis: http://arstechnica.com/business/2012/05/how-across-the-roman-empire-in-real-time-with-orbis/

V. Conclusive Suggestions

The Digital Humanities platforms provide us with the future orientation of our world literature studies which turns from close reading to distant viewing.

Technological innovations change quickly and the field of digital humanities expands, and more and more materials come online, through research projects. Other repositories or platforms combine technical, academic and cultural issues at a scale that is unprecedented. Figuring out how repositories can "talk" to each other or be integrated at the level of search and addressing the fundamental problems of intellectual property is what is at stake: modes of citation and linking that respect conventions of copyright while serving to support public access, education, and scholarship.

Perspective of Convergence

- ✓ Two of the most vibrant fields of emergent scholarship in the humanities and arts have been:
 - the examination of "world literature" in the circles of comparative literature
 - "new media arts" / "digital humanities" in the circles of visual studies and humanities.
- ✓ While scholars in global literature have emphasized the impact of international projects in "world literature" & "cultural translation," practitioners and scholars of the digital arts and humanities have experimented with "novel forms of multimedia communication" whose platforms are equally accessible to audiences across the electronic "archive" and "internet."



- ✓ What is at stake is the methodology of cultural representation and translation by way of the digital humanities which has to create the "proper distancing" based upon the balance between close reading and distant viewing.
- ✓ Benjaminia metaphor of "Fragments of a vessel which are to be glued together must match one another in the smallest details, although they need not be like one another" (78) is still persuasive, when digital humanist and literary critics attempt to build the vessel of the world literature via the methodology of the digital humanities.

Convergence of the Literary Database of World Literature and Digital Humanities!

❖ Digital Humanities: Basic Concepts

- ✓ Analysis of DH Projects, Platforms, and Tools
- ✓ HTML: Structured Data, Content Modelling,
- ✓ Ontologies and Metadata Standards
- ✓ Data and Data Bases: Critical and Practical Issues
- ✓ Database and Narrative
- ✓ Information and Visualization Concepts
- ✓ Critical and Practical Issues In Information
- ✓ Data Mining and Text Analysis
- ✓ Text Encoding, Mark-Up, and TEI.
- ✓ Distant Reading and Cultural Analytics
- ✓ Network Analysis
- ✓ GIS Mapping Conventions
- ✓ Interface Basics.
- ✓ Interface, Narrative, Navigation, and Other Considerations
- ✓ Virtual Space and Modelling 3-D Representations
- ✓ Critical Issues, Other Topics, and Digital Humanities





(* INTRODUCTION TO DIGITAL HUMANITIES: Concepts, Methods, and Tutorials for Students and Instructors COURSE BOOK, JOHANNA DRUCKER, et al., DAVID KIM, IMAN SALEHIAN, ANTHONY BUSHONG)

New Reality: Digitization, Trans, Mobile

❖ Digitization: Algorithm

Film: <Matrix I,II,III>

- **❖** The Age of Trans: Automatization, Ubiquitous → Post/Trans-humanism
 - ✓ Transnationalism
 - ✓ Transculturalism
 - ✓ Transhumanism
 - ✓ Transmedia
 - ✓ Transdiscursivity
 - ✓ Translation
 - ✓ Transgender, Transexual
- ***** Mobile Environment



Digital Bibliography 1

*Local Model Websites for the Future Global Platforms for the Convergence of Digital Humanities and World Literature

•18th Connect http://www.18thconnect.org/

Australia Network: http://ww35.australianetwork.com/nexus/stories/s2160521.htm

- •Dante Gabriel Rosetti: http://www.rossettiarchive.org/index.html
- •Digital Public Library of America: http://dp.la/
- •Emily Dickinson: https://rotunda.upress.virginia.edu/edc/default.xqy
- •Europeana: http://www.europeana.eu/portal/en
- •Kindred Britain: http://kindred.stanford.edu/
- •National Library of Australia: http://www.nla.gov.au/
- •NINES http://www.nines.org
- •Pelagios http://pelagios-project.blogspot.kr/p/about-pelagios.html
- •Republic of Letters: http://republicofletters.stanford.edu/
- •Strange maps: http://bigthink.com/blogs/strange-maps
- •ubuweb http://ubu.com/
- •Walt Whitman: http://www.whitmanarchive.org/
- •Websites for the digital humanities:
- •Valley of the Shadow http://valley.lib.virginia.edu/VoS/choosepart.html
- •Visual Complexity: http://www.visualcomplexity.com/vc/
- •Women Writer's Project http://www.wwp.northeastern.edu/

Digital Bibliography 2

❖ Digital Humanities: Connectivity-Network

Selected Digital Humanities Centers and Initiatives

The Alliance of Digital Humanities Organizations (ADHO)

Digital Research Infrastructure for the Arts and Humanities (DARIAH)

Yale Digital Humanities Laboratory (DHLab),

Price Lab at the University of Pennsylvania.

Princeton Center for Digital Humanities

Columbia, The Center for Digital Humanities

Oxford Research Centre in the Humanities (TORCH),

The DXARTS Center for Digital Arts, Experimental Media at the Univ

of Washington

Stanford CESTA Labs

Scholars Lab at the University of Virginia

Digital Bibliography 3

Digital Humanities Summer Institute: University of Victoria, Canada

http://dhsi.org/

✓ The **Digital Humanities Summer Institute:** A time of intensive coursework, seminars, and lectures, participants at DHSI share ideas and methods, and develop expertise in using advanced technologies: workshop, international conference, and summer camp.