The Chicago School: Evolving Systems of Value

Primary researcher:
Dan Costa Baciu, orcid.org/0000-0002-0043-5616

Advanced Collaborative Support (ACS) Senior Advisors:
Asst. Prof. Jonathan Mekinda, Art History, University of Illinois in Chicago
Prof. David Van Zanten, Art History, Northwestern University
Prof. Michael Golec, Design History, the School of the Art Institute of Chicago
Prof. Dan Roth, Computer Sciences, University of Illinois Urbana Champaign
Dr. Irina Matveeva, Computer Sciences, Illinois Institute of Technology

HathiTrust Research Center (HTRC):
Boris Capitanu, Eleanor Dickson

Programmers Wikifier:
Chase Duncan, Mark Sammons

Programmers Illinois Institute of Technology (IIT):
Swagatam Guhathakurta, Chandra Kumar Basavaraju, Mayuri Kadam, Petter Castro Tigre, Shruti Naik, Harsha Keragodu Shivappa, Rima Sukadia, Khushboo Mrugender Shah, Arnab Mukhopadhyay

Results first presented:
Dan Costa Baciu, “The Chicago School: Wikification of 105’000 Volumes 1850-2015,” “Mining for Important Textual Content: A Novel Approach Mimicking Human Associative Memory,” “The Chicago School: Quantitative and Qualitative Data for a Complex Systems Analysis,” three papers presented at the Digital Humanities Summer Institute, Victoria CA, 13, 14, 15 June 2017.

Published by HTRC:
September 30, 2017

Grants:
Fulbright Program, Illinois Institute of Technology, Swiss American Society,
Swiss National Science Foundation project number 174883

Doctoral advisors:
Prof. Dr. Michelangelo Sabatino, Prof. Dr. Harry F. Mallgrave

Abstract:

Architects, sociologists, lawyers, and many other practitioners refer to the Chicago School as a historiographic construct, a school of thought, or a set of values and expectancies.[1] The term's precise definition varies with the speaker's mindset. At the HathiTrust, over 100,000 books and periodicals mention the Chicago School. Can this massive history of publication be seen as a written record of cultural change? This present work attempts the first overarching systematic and critical evaluation. Over the last two centuries, the values of the Chicago Schools have been disseminated and transformed, often resulting in a loss of historic reference. We disambiguate the many definitions that the term Chicago School may assume, and encounter a small number of very well known, frequently mentioned variants surrounded by a vast spectrum of rare ones. Enumerating the resulting numbers in decreasing order reveals a Pareto distribution, an empirical formula that illustrates a widely observed relationship between frequencies of occurrence and enumeration ranks. We sketch a theory that explains the mechanism that lead to this distribution; we will further elaborate along these lines at the Chicago Schools symposium that we organize in Fall 2017.[2] As part of this larger trajectory, the HTRC ACS project provided sufficient data to evaluate the complete history of publication of the Chicago Schools. We succeeded in implementing entity linking to Wikipedia on a scale previously unattempted, and it offered significant advantages over using the unstructured data alone. From our large Chicago School corpus, we also build three additional datasets together with a framework for non-consumptive research which allows us to filter and classify the data as well as cross-validate the results.

1. What is the Chicago School?

A million tourists come to Chicago every year, many of whom take a boat tour to learn about the architecture of the Chicago School. Economists, philosophers, writers and many other practitioners also speak about the Chicago Schools that emerged in their own disciplines. Cataloging the different definitions oftentimes reveals a common ethos and shared historic roots as well as contradictions and disarray. This present research attempts an overarching evaluation that explains how the values of the Chicago Schools were disseminated and transformed. It might sound unthinkable to propose that hundreds of Chicago Schools coexisted and evolved over nearly two centuries. As preliminary results, scholars might therefore wish to hear about important milestones in the history of publication of the Chicago School. The present research has found that, in architecture, the term dates back to 1889 and 1893 when Henry Van Brunt wrote two magazine articles endorsing his Midwestern peers.[3] He propounded that a number of contemporary commercial buildings were first signs of a new synthesis that boldly unified art and engineering. With his contributions, Van Brunt triggered a long-lasting discussion.[4] Most prominently, his ideas were echoed in two books that dominated the architectural discourse for decades. Although each of these publications is unique in its scope, their authors discussed the Chicago School in similar terms. A.D.F. Hamlin, an early Columbia professor, published the first textbook of the history of architecture written in the United States and mentioned the Chicago School in five reprints, 1900 to 1907.[5] The term referred to a group of Midwestern practitioners that successfully integrated the engineer's work in the façades of their highrises, as opposed to the Eastern School that only focused on artistic expression. Three decades later, Sigfried Giedion lectured and wrote on the Chicago School claiming its relevance as the first synthesis of art and engineering on the scale of a whole city.[6] Giedion's book *Space, Time and Architecture* (1941, frequently reprinted) became a scripture of Modernism and his Chicago School essay was read, debated, and distorted. Colin Rowe and Manfredo Tafuri count among the most influential followers. Their work is an obvious case of late reinterpretation. However, generally speaking, in the history of publication, dissemination and transformation went hand in hand from the early beginnings. The years 1889 and 1893 brought the establishment of architectural education in the city of Chicago, and in this context, Van Brunt's Chicago School assumed a different format. Two institutes, one rooted in the arts, the other focused on engineering, found each other in their venture to unite the two disciplines into one joint program that was named the Chicago School of Architecture. Van Brunt's peers, including Daniel Burnham and Louis Sullivan, went in and out of the school's doors. The density of closely related Chicago Schools was certainly visible in 1908 when Thomas Tallmadge, a lecturer, redefined the Chicago School, slightly departing from all previous ideas.[7] He mentioned Sullivan as a key figure, but mostly referred to an architectural style developed for suburban mansions. No contradiction appears to have been felt by contemporaries. As the many Chicago Schools evolved side by side, they kept identifying themselves as countermovements to the East Coast, or repeated the claim of synthesizing theory and practice. In 1903, the Chicago School of Fiction was constructed around literary characters, that were "fine as frank," and whose pure thought flew in fountains of slang.[8] In 1904, the Chicago School of Thought was born as "a promising *via media* between the empiricist and transcendentalist ten-

dencies.” [9] In the ensuing decades however, the latter branched out into numerous other schools including sociology, economics and law, not all of which promoted the same values. What then is the Chicago School? At this point we can only say that it is a set of values that was disseminated, translated and transformed. In order to draw more conclusions, the present research sets out on a comprehensive critical and systematic evaluation, in which the rise and fall of major Chicago Schools is studied on the background of the complete history of publication. As part of this greater endeavor, the ACS project provided sufficient data to disambiguate the various definitions and quantify their development over time.

The Chicago School at the HathiTrust

Total number of volumes	105,000 volumes
Year range	1850 to 2015
Deduplicated number of volumes	91,000 volumes
Chicago School mentions	195,000 pages
Periodicals, Books	55%, 45%
Minimum volume length	2 pages
Maximum volume length	4000 pages
Median volume length	550 pages
Total collection length	50 million pages

2. Theoretical Synopsis

An overarching systematic of the Chicago Schools requires the development of a set of criteria that tell us what the generic, specific and accidental attributes of the narratives are. This present research develops a bottom up procedure by relying first on the positions taken by the authors. Aware of the coexistence of multiple Chicago Schools, 40% of the narratives specify their context. Our corpus contains a "Chicago School of Burnham," "of 1880," or "of 1930." These and other similar expressions can be used as a first baseline to disambiguate. In subsequent steps, the systematics are refined by further subdividing into additional categories as well as merging groups of too similar results. After this process of labeling, we classify the remaining 60% of the mentions by techniques described in part 3. Once a final label has been assigned to every record, the frequencies of occurrence can be counted. The result appears somehow familiar. If we enumerate the obtained numbers in decreasing order, we find a Pareto distribution. Frequently mentioned schools are surrounded by a vast spectrum of rare variants. Ultimately, most schools are mentioned only a couple of times. We can find among them the Chicago School of Tenors, the Chicago School of humorous basketball players, or the Chicago School of bone breakers. The Pareto distribution has only been formulated empirically; it can be found in a broad range of data, but it remains unclear why it emerges naturally in those different contexts.

2.1. From Pareto Distribution to Zipf's Law

To understand the implication of the Pareto distribution, historians might wish to know about its historic context and development. A civil engineer for Italian railroads, Vilfredo Pareto started his work in economics in his forties. In 1893, he was appointed as a professor at the University of Lausanne, and three years later he discovered what is today known as the Pareto distribution. Empirically observed in economics, the distribution was immediately seen to be valid in other fields.[10] Jean-Baptiste Estoup formulated an application in stenography, and the law also found its way to the German Avantgarde where it was applied by Felix Auerbach, a professor of physics and patron of the arts. The painter Edward Munch portrayed him in 1906, and Walter Gropius built his house, an icon of modern architecture, in 1924.[11] Henry van de Velde, Ernst Ludwig Kirchner, Max Bruch, Richard Dehmel were among his frequent guests. Wassily Kandinsky and Paul Klee, in particular, admired Auerbach's work explaining Albert Einstein's Theory of Relativity: *Raum und Zeit, Materie und Energie* (Space and Time, Matter and Energy).[12] *Space, Time and Architecture* Giedion's work mentioned in part 1 resembles Auerbach's title choice. Empirically studying the sizes of cities, Auerbach observed a compelling regularity: in any chosen area, there were few big cities surrounded by many smaller dwellings, villages, and hamlets. If the population numbers were sorted in decreasing order and multiplied by their enumeration rank, Auerbach received a constant value. Although Auerbach also rediscovered the validity of this power probability distribution for income data, his article of 1913 did not cite Pareto's work.[13] Only twenty years later, in February 1933, Auerbach suffered a second heart attack and commit-

ted suicide after the Nazi party seized power in Germany. During that period, the Pareto distribution was popularized as Zipf's law. After studying in Bonn and Berlin, George K. Zipf took the position of Lecturer in German at Harvard University, and in a series of articles starting with 1932, he further developed the regularity previously observed by Pareto, Estoup, and Auerbach. Zipf first counted word frequencies from concordances, but he followed Pareto's and Auerbach's line in recognizing that the distribution was to be found in many other collections of data, which led him to find an increasing number of socially relevant applications. $P_n \sim 1/n^a$ became a law that linked geography, society, literature, art, and engineering. Sigfried Giedion was among the readers. In his last published article, Zipf was calling for a scientific analysis of culture, building on his own work, as well as a longer tradition that had originated in the Vienna School. Similar to the Chicago School, the term Vienna School may assume multiple meanings. There exist a first and a second Vienna School of Music – Johann Strauss II's *The Blue Danube* might resound here. Zipf referred to the Vienna School of Art History, a line of scholars that promoted a scientific approach in assessing the arts, among them the preeminent art and architectural historian Alois Riegl, or Ernst Gombrich and Meyer Schapiro, known for the new Vienna School of the 1930s.[14] Overall, Zipf's essay brought together concepts from art history, semiotics, and psychology. The aim was to move historiography away from the concept of 'holymen' who dominate the discussion, to empirical and population studies. However, Zipf fell sick and died after three months of agony in February 1950.[15] Giedion did not grasp all of the implications of Zipf's law for the Chicago School. Ultimately, the law was applied to different words within collections of text. It was not applied to different definitions of the same term, as it is in this present research. Furthermore, Zipf's law is still an empirical formulation, which means that it does not make the step from dissemination and transformation to the actual shape of the distribution.

2.2. A Unifying Theory

In the process of writing, when an author develops his Chicago School narrative, he goes through a set of choices that others have been faced with before him. Most frequently, his decisions will adhere to his knowledge, often but not necessarily reflected by the tradition he stays in. Therefore most of his choices coincide with those taken by precursors. Yet, at a certain point, the author may search for alternatives and make a choice that might be artistic and not fully explainable but leads to a conceptually different and potentially interesting result. In the process of creation, new concepts move away from older ones by the degree to which new choices are accumulated. A similar picture can be drawn in retrospect, after the narratives have been written and published, by categorizing the Chicago Schools. Among the many unique narratives, recurrent definitions can be disambiguated by Genus and Differentia. A group of close definitions (in logic also called a Species) always share numerous common attributes which are referred to as their Genus. Differentia then denominates the choice that makes one variant depart from the others. Defining by Genus and Differentia dates back to Aristotle. Major systematics such as the Linnaean taxonomy are laid out by the logic of Species and Varieties, Genus and Differentia. In the epoch of Enlightenment, Hegel noticed that logic had remained almost unchanged since antiqui-

ty, although it lead to dramatic development in the scientific understanding of the world. Certainly, the process of defining by Genus and Differentia entered all disciplines, but the consequences of this logic have not been fully interpreted on the scale of whole populations. The merit of our present work lies in deriving a theory that unifies all fields in which Genus and Differentia apply and departure from existing variants may occur. Let us start by looking at the big picture. We assume that variants may coexist side by side, although some have higher utility and are disseminated faster than others. Furthermore, we may introduce the concept of an average probability of new choices being taken. The distance between any two variants may then be represented by the number of steps necessary to transform one definition into the other. From this set of premises already a great number of conclusions can be derived. In physical chemistry, a formalism known as "the Quasispecies equation for molecular evolution" was derived from similar premises.[16] The equation was developed from work that previously earned his researcher a Nobel Prize. Unexpectedly, this model proved even more useful in biology. The evolution of populations of virus such as HIV and HBV in the human body has been calculated using this approach and led to the adoption of new successful regimens.[17] Researchers especially valued the model's explanatory power on the level of genetics, and a major recent publication indicated that this was a consequence of the model's origins in chemistry.[18] In contrast to that verdict, this present work suggests that the applicability of the formula depends on the validity of the assumptions. The theory of variation that we derive from our four premises not only excels by interpreting Genus and Differentia on the scale of populations, or by the previous success of special formalisms applied in chemistry and biology, it also explains the evolution of the many definitions of Chicago School, and potentially many other terms such as the Vienna School, the Modern, the Beautiful, the Sublime, the Classical, the Romantic or the Vernacular. The results are easy to illustrate. At first, we encounter a small number of Chicago Schools that spread the fastest. If we allow the definitions to transform while being disseminated, a certain number of conceptually close neighbors immediately arise. With some additional transformation, an increasing number of more distant neighbors is generated. This already explains the mechanism behind the frequencies of occurrence for the various Chicago School definitions; the distant neighbors are more and more numerous, although less likely to have the same spreading ability, and in that case, they are increasingly rare. After ordering the results, we receive a Pareto distribution as described above, but it now emerges logically from premises that we might find hard to reject. The present work suggests that the Pareto distribution, or Zipf's law, result as a special case of the theory of variation. Mathematically formulating the theory of variation, the present work also explains the slight deviations from Zipf's law. Furthermore, there are a number of simple, practical conclusions: at the time of historic breaks, when the the utility of definitions is changing, some of them get obsolete while other gain relevance. At that point, a system based on constant transformation is prone to adapt easily as long as one of the many previously existing variants fits the new historic situation. Indeed, adaptation is expected to emerge more often from the spread of preexistent variants that fit the new context rather than the tailored design of wholly new definitions.[19] Another interesting conclusion, groups of mediocre, closely related definitions strive over unique very useful ones, which also matches our observations. However, variation also has limits given by the size of the population, the rate by which new variants emerge, and

the rate of historic change. Crossing the thresholds imposed by these factors might lead to chaos and disappearance of consensus. In every day life, the theory of variation might help institutions boost their creativity, but the theory might also be used or misused to push creativity beyond chaos thresholds at the point at which a population will need mechanisms of self regulation that do not depend on Genus, Differentia and transformation.

Distribution of Chicago Schools (logarithmic scale, preliminary results)

3. Employed Methods and Methodological Results

Quantitative evaluation has accompanied cultural theory for centuries, but the last decades have witnessed massive advances due to increased computational power.[20] Big Data was recently called the petroleum of the 21st century, because it is only useful once processed. Various Data Mining techniques are rapidly being developed. "Big Text" has been coined for the terabytes of written words that are generated on an every day basis. Researchers team up with companies to explore privately owned collections, while students develop their skillsets by deploying text mining techniques on open access data from online platforms such as news and twitter.[21] Developing, refining, and combining algorithms, this present work develops a framework for non-consumptive research.

Our innovations lie in: the implementation of a knowledge-based approach to structure text data on a large scale (3.1), setting a precedent for all future HTRC cooperations that require non-consumptive analysis of copyrighted text (3.2), the introduction of a process to automatically assess the accuracy of final results (3.3), the introduction of a filtering technique mimicking human associative memory (3.4).

3.1. A Digital Tradition Reaches a New Milestone

In architecture, early attempts in using computers to implement knowledge-based approaches date back to the 1990s. Researchers such as Roxane Kuter Williamson at University of Texas Austin, and Juan Pablo Bonta at MIT built databases with entries for several thousand architects to extract structured information from headlines.[22] Bonta's project was designed to fit on a computer with six megabytes disk space. Today, every single record is about this size. Compared to Bonta's publication, the size of the datasets generated by the current project increased by factor 100,000. At this scale, the algorithms were performed on a supercomputer, and instead of building a customized database, we relied on entity linking to Wikipedia, a type of algorithm developed since 2006.[23] The hyperlink structure and the broad coverage offer a great advantage in the task of disambiguation.[24] The Illinois Wikifier outperformed earlier tools by introducing global wikification.[25] Previous work developed a framework for the assessment of algorithmic accuracy that we continue using.[26] Although entity linking was only performed on relatively small datasets such as news groups, we made a step forward by implementing this computationally expensive algorithm on a much larger scale, a collection of much bigger size and broader rage. We were able to estimate a link accuracy of roughly 60%, although varying with OCR quality and writing style up to 80% accuracy is obtained. Not all assigned links were equally accurate; the Chicago School of Architecture, for example, was much better recognized than its author Thomas E. Tallmadge. Originally, the algorithm was developed to annotate texts for readers. In the present ACS project entity linking served a new purpose, our idea was to implement it as a

knowledge-based approach to collect structured data. In this context, entity linking made various implementations possible that work with this type of data, among them graph mining and analysis of concordances. More unexpectedly, the structured data proved highly beneficial in disambiguating the various Chicago School definitions. Although the accuracy of each assigned link lies greatly below 80%, the subsequent classification of Chicago Schools was performed with accuracies above that value. Classification performed on the data obtained from the Wikifier offered two to three times more accurate results compared to the other datasets. We conclude that Wikification is a potent approach for non-consumptive research, enabling a number of applications that range from filtering to accurate classification. The access to Wikification on a large scale might well contribute to revolutionizing non-consumptive research and the quantitative evaluation of culture.

3.2. Cooperation with HTRC, a Precedent for Future Work

In 2016, HTRC received access to the copyrighted data stored at the HathiTrust. The present ACS project developed a precedent for all future collaborations that will require access to this type of data. As part of the collaboration, large bodies of text were algorithmically processed. Our collection contained 105,000 volumes ranging from 1850 to 2015. Relevant book metadata is selected, and from 50 million pages, 190,000 text snippets mentioning the Chicago School are extracted as well as 0.5 terabytes of named entities and links to Wikipedia along with their volume IDs and page numbers.

3.3. Cross-validation of Results from Distinct Datasets and Algorithms

The assessment of classification accuracy requires labeled data. We developed a method to assess algorithmic performance on unlabeled data as well. The idea is to compare results that were obtained from at least three independent datasets. Mapping these results onto each other reveals the amount of mutual agreement among each combination of datasets. Based on this information, the accuracy of every set of results can be estimated mathematically. The range of potential fluctuations may then be simulated, if relevant. Thus we assess the final results without the need of a gold standard. Our method is crucial especially for big datasets that contain much unlabeled data. If the labeled data is not stratified enough, which is often the case, accuracy estimation becomes unreliable.[27] Our new type of cross-validation avoids this unwanted effect.[28]

3.4. A Filtering Technique Mimicking Human Associative Memory

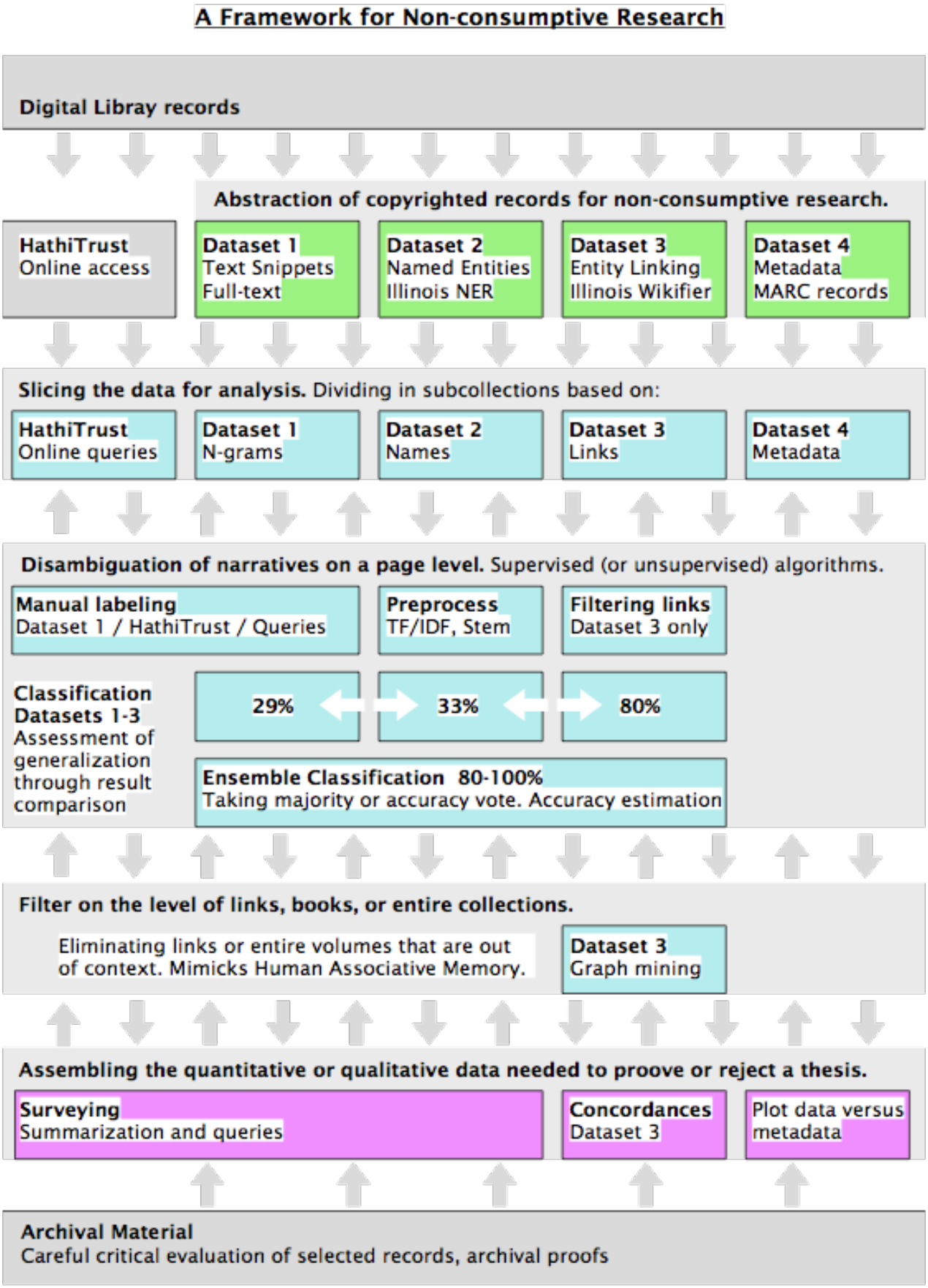
Applying the basic concepts of associative and sequential memory to the very act of reading a book, we develop a simple algorithm that simulates the reader's mind-map. Based on two seminal publications,[29] we decide to record associative memories by linking together pairs of close named entities, and we keep track of the episodic trace by taking into account the reading direction. The result is a directed graph with weighted edges. Each node represents a named entity encountered in the reading. The weight of the edges records the relationship between two named entities. The algorithm then rates the nodes by a recursive definition: memorable nodes are those

that are connected to many other memorable nodes. Technically, the computer solves this iterative process by running power iteration.[30] To add weight to the beginning of the paragraphs, the power iteration is run with restart.[31] This procedure selects clouds of important named entities, a feature reminiscent of the unifying theory described above.

In any given text, most named entities are mentioned only once. In comparison to counting frequencies of occurrence, the associative memory algorithm increases the ranking of rare named entities that play a key role in an important context within the book. Furthermore, the algorithm also tends to lower the ranking of terms in portions of text, in which the entity linking worked badly.[32] A generally occurring observation in our corpus, the algorithm increases the weight of big terms such as the names of cities. In most books, city names such as Chicago or New York move up in the hierarchy of named entities. As performed on the Chicago School corpus, the tool allowed to filter for important information and perform content-based searches. We also developed a similar algorithm to eliminate books that are out of context in a given collection. As an outlook, the ability to simulate human memory could be used to reconstruct collective memories of entire social groups.

Datasets & Algorithms

	Snippets	**Named Entities**	**Links to Wikipedia**	**Metadata**
Type of data	Unstructured	Partly structured	Structured	Structured
Contents per page	42 words	around 70	around 70	29 keys
Accuracies of data	rare typos	rare typos	60%	fuzzy
Classification accuracy easily reachable on data	29%	33%	80%	
Slice data based on	ngrams	names	links	metadata keys
Filter inside volumes			graphs	
Filter whole collections			graphs	


A Framework for Non-consumptive Research
Digital Libray records
Abstraction of copyrighted records for non-consumptive research.
HathiTrust
Online access
Dataset 1
Text Snippets
Full-text
Dataset 2
Named Entities
Illinois NER
Dataset 3
Entity Linking
Illinois Wikifier
Dataset 4
Metadata
MARC records
Slicing the data for analysis. Dividing in subcollections based on:
HathiTrust
Online queries
Dataset 1
N-grams
Dataset 2
Names
Dataset 3
Links
Dataset 4
Metadata
Disambiguation of narratives on a page level. Supervised (or unsupervised) algorithms.
Manual labeling
Dataset 1 / HathiTrust / Queries
Preprocess
TF/IDF, Stem
Filtering links
Dataset 3 only
Classification
Datasets 1-3
Assessment of
generalization
through result
comparison
29%
33%
80%
Ensemble Classification 80-100%
Taking majority or accuracy vote. Accuracy estimation
Filter on the level of links, books, or entire collections.
Eliminating links or entire volumes that are out
of context. Mimicks Human Associative Memory.
Dataset 3
Graph mining
Assembling the quantitative or qualitative data needed to proove or reject a thesis.
Surveying
Summarization and queries
Concordances
Dataset 3
Plot data versus
metadata
Archival Material
Careful critical evaluation of selected records, archival proofs

Endnotes

[1] The term's definition could be laid out as a set of values and expectancies as in Niklas Luhmann's system theory.

[2] Dan Costa Baciu, "Chicago Schools: 150 years of publication history," symposium lecture, *Chicago Schools: Authors, Audiences and History*, Chicago: Illinois Institute of Technology in partnership with Chicago Architecture Biennial, fall 2017.

[3] Dan Costa Baciu, "Sigfried Giedion: Historiography and History of Reception on a Global Stage," editors Iris Aravot and Dana Margalith, *Ar(t)chitecture,* Haifa: Technion Faculty of Architecture and Urban Planning 2016, 40-52.

[4] The discussion in the press can be documented back to 1893 with: John Willis Abbot, "The Makers of the Fair: A Family Paper," *The Outlook* 48, 18. November 1893. The text contains a full mention of the term "Chicago School of Architecture," at the content is similar to Van Brunt's article the same year. These early Chicago School mentions were never included in scholarly bibliographies. They were re-discovered with data mining tools as part of this present research.

[5] Alfred Dwight Foster Hamlin, *A Text-book of the History of Architecture,* New York: Longmans Green and Co, 1895, reprints Oct. 1900, Oct. 1902, Sept. 1904, June 1906, Nov. 1907.

[6] Sigfried Giedion, "The Danger and Advantages of Luxury," *Focus* 3 (1939), GTA Archive 43-T-15-1939-1.
—— , lecturer. "America influences Europe: The Chicago School and Frank Lloyd Wright." Charles Eliot Norton Lectures, Harvard University, 1939, GTA Archive, 43-T-13-7-1-8-2.
—— . *Space Time and Architecture.* Boston: Harvard University Press, 1941.
—— . Lectures. GTA Archive, 43-T-13-3-1; 43-T-13-1-5-5; 43-T-13-1-19-4; 43-T-13-7-1-8-2.

[7] Tallmadge, Thomas Eddy. "The 'Chicago School.'" *The Architectural Review* XV, IV (1908), 69-74.

[8] W.D. Howells, "Certain of the Chicago School of Fiction," *The North American Review* 1903, 740.

[9] William James, "The Chicago School," *Psychological Bulletin,* Vol 1, (1904), 1-5.

[10] Micheline Petrusyewycz, "L'histoire de la loi d'Estoup-Zipf: documents," Mathématiques et sciences humaines 44 (1973), 41-56.

[11] Barbara Happe and Martin S. Fischer, *Haus Auerbach: Von Walter Gropius mit Adolf Meyer/ Of Walter Gropius with Adolf Meyer*, Tübingen: E. Wasmuth, 2003.

[12] Felix Auerbach, *Raum und Zeit, Materie und Energie,* Leipzig: Dürr'sche Buchhandlung, 1921. Quoted from: Ulrich Müller, *Raum, Bewegung und Zeit im Werk von Walter Gropius und Ludwig Mies van der Rohe,* Berlin: Akademie Verlag, 2004.

[13] Felix Auerbach. "Das Gesetz der Bevölkerungskonzentration." *Petermanns Geographische Mitteilungen*, 59, 1913, p. 73-76.

[14] Martin James, Harry Holtzman, letter to Sigfried Giedion, gta archives ETH Zurich, Sigfried Giedion Estate, 43-K-1949-12-30. Interrelations Art, Architecture, Engineering, scheme of the Article of George K. Zipf attached to the letter.

[15] "Zipf Dies After 3-Month Illness," *Harvard Crimson*, September 27, 1950. [http://www.thecrimson.com/article/1950/9/27/zipf-dies-after-3-month/]

[16] Manfred Eigen, "Selforganization of matter and the evolution of biological macromolecules," *Die Naturwissenschaften* 58, 10 (October 1971), 465–523.

[17] Martin A. Nowak and Robert M. May, Virus Dynamics: The Mathematical Principles of Immunology and Virology, Oxford: Oxford University Press, 2000 [first edition].

Esteban Domingo and Peter Schuster, "What is a Quasispecies? Historical Origins and Current Scope," in: Esteban Domingo, ed., *Quasispecies: From Theory to Experimental Systems,* Berlin: Springer 2016.

[18] Esteban Domingo, ed., *Quasispecies: From Theory to Experimental Systems,* Berlin: Springer 2016.

[19] In biology referred to as Bonhoeffer's Law. Martin A. Nowak and Robert M. May, Virus Dynamics: Mathematical Principles of Immunology and Virology, Oxford: Oxford University Press, 2000, 99-100.

[20] Thinking for example of Jean-Baptiste Estoup's work. Today inverted indices have become a frequently used data structure that allows for fast results.

[21] Jean-Baptiste Michel Yuan Kui Shen, Aviva Presser Aiden, Adrian Veres, Matthew K. Gray, The Google Books Team, Joseph P. Pickett, Dale Hoiberg, Dan Clancy, Peter Norvig, Jon Orwant, Steven Pinker, Martin A. Nowak, and Erez Lieberman Aiden, "Quantitative Analysis of Culture Using Millions of Digitized Books," *Science* 331 (2011) [Published online ahead of print 12/16/2010]

[22] Roxane Kuter Williamson, *American Architects and the Mechanics of Fame,* Austin: University of Texas Press, 1991.

Juan Pablo Bonta, *American Architects and Texts A Computer-Aided Analysis of Literature; Electronic Companion to American Architects and Texts,* Cambridge: MIT Press, 1996.

[23] R. Bunescu and M. Pasca, "Using encyclopedic knowledge for named entity disambiguation," *EACL-06*, Trento, Italy, April. 2006, 9–16.

Entity Linking to Wikipedia is also known as Wikification, disambiguation to Wikipedia, D2W

[24] Rada Mihalcea and Andras Csomai, "Wikify! Linking Documents to Encyclopedic Knowledge," *Proceedings of the Sixteenth ACM Conference on Information and Knowledge Management*, (CIKM '07), New York, 2007, 233–242.

[25] Lev Ratinov and Dan Roth, D. Downey and M. Anderson, "Local and Global Algorithms for Disambiguation to Wikipedia," *ACL* (2011).

[26] Marco Cornolti, Paolo Ferragina, Massimiliano Ciaramita, "A Framework for benchmarking Entity-annotation Systems," *WWW 2013.*

[27] For example obtained through n-fold cross-validation, because it is always based on labeled results.

[28] Ensemble methods typically work by comparing the results of the same type of algorithm performed on the same type of dataset. Therefore, the accuracies of each set of results is similar. In our case, the sets of results have very different accuracies, assessing the generalization accuracy of each algorithm individually helps improving the performance of the ensemble method as well.

[29] John R. Anderson and Gordon H. Bower, *Human Associative Memory,* Hillsdale New Jersey: Lawrence Elbaum Associates Publishers, 2014. [First edition 1973]

Endel Tulving, *Elements of Episodic Memory,* Oxford: Oxford University Press, 2016. [First edition 1985]

[30] Richard von Mises and H. Pollaczek-Geiringer, "Praktische Verfahren der Gleichungsauflösung," *ZAMM - Zeitschrift für Angewandte Mathematik und Mechanik* 9, 1929,152-164. The algorithm is most notably used by google page rank.

[31] Instead of the usual teleportation, we restart to the named entities that are first mentioned in their paragraphs.

[32] Because those portions of text result in long chains of words that occur only once without repetition.