# Erica Folk, PhD – citation analysis specialist

Erica holds an MLIS (Master of Library and Information Studies) and an interdisciplinary PhD in Science, Technology and Society. Her research is focused on studying the development and evolution of academic networks of research, from both historical and contemporary perspectives. She is familiar with the field of citation analysis, which she finds significant from the point of view of her research. She has published some articles describing emerging nuclei of scholarly research in interdisciplinary fields and has supported her findings with citation data extracted from ISI Web of Science. She is also familiar with large scale visualizations of citation patterns like HistCite or CiteViz, which she finds interesting and useful for her own research.

More recently, she is increasingly interested in annotation as a mean of identifying trails of scholarly thought and the social aspects of scholarly research, with a particular view on humanist research where she believes that due to the customary ample critical apparatuses the genome of a scholarly work might be easier to trace.

Erica is aware of the shortfalls of traditional citation analysis when studying humanist writings with non-parenthetical citations and hopes that our tool will help her overcome these difficulties by allowing her a more nuanced and complex view of the way knowledge and argument are built in the humanities.

She is also:

* teaching an undergraduate class in Science, Technology and History about the development of 19th century scientific thought and its impact on other areas of knowledge, and
* facilitating a graduate seminar on theoretical aspects of scholarly research.

## Goals:

* Have a visual, interactive interface to track the references included in a lengthy monograph
* Have continuous access to the text of the monograph while the visualization runs
* Have the possibility to analyse and visualize both
  + the bibliographic characteristics of the works cited and
  + the semantic characteristics of the citations (i.e. context in which the citation occurs, function of the citation, and the relationship with other works cited in the location)
* Be able to add additional monographs to the database of the tool, using an XML editor preloaded with the tool’s schema.
* Be able to modify the semantic portion of the mark-up of the visualized monograph, (i.e. tags that relate to the context and function of the citation; this could be done if users would login an work on their own copies of the deposited document?)

## Scenarios:

* Erica uses the tool for her research on “annotation as a primary scholarly instrument in the age of scientific humanism”. She is studying the role of citations within the larger context of a monograph’s argument. She is interested in the construction of argument in large-scale works and wants to see if:
  + she can identify citation patterns in humanist monographs (that can be predicted), and if
  + she can trace the development of the argument and it’s supporting annotations in large-scale humanist works.

She performs various inquiries using the visualization tool in order to compare different citation features (both syntactic and semantic) at different locations in the document.

Details:

Erica analyzes a literary history monograph on the sources of the Victorian novel. After loading the document, the tool defaults to the comparison view. Here Erica has the list of cited works in a format similar to that at the end of the book (author, title, date).

She first explores this list visually, by experimenting with the sorting options: she sorts the list alphabetically by author names, looking for the ones that were previously identified as the essential names in the particular subfield of the monograph (maybe in a previous traditional citation analysis study). She then sorts it by date of publication to see what’s the most recent study cited and to get a feel of whether a singular period is better represented than the rest. If she can identify such a period, she checks if she was right by filtering the list to show one or more periods of publication.

She might also be interested in getting a feel of how much of the list of references (or the list after she applies the “filter by period”) consists of monographs vs. articles, so she might want to compare the two categories. She might notice when doing this comparison that the citations from articles are mostly grouped in the first part of the monograph, while the citations from monographs are predominant in the second part. A quick look at the text of the monograph loaded into the visualization tells her that the first part (the one where citations from articles are concentrated) is where the bulk of the data is included and where the context of the author’s theory is built with information that might or might not be connected with the subject of the monograph in an obvious way (e.g. Articles from newspapers of the time written by the first war correspondents from the Crimean War might be cited to illustrate the development of a journalism style, which the author might later argue that influenced the style of the Victorian novel in the second half of the century). On the other hand, the part where citations from monographs are more common is the discussion portion of the work. Erica decides to have a more in-depth view of this discussion, so she switches to the second view of the tool (the note span) where she zooms in to the portion of the monograph she is interested in by clicking on the representation of that portion in the column representing the entire text of the monograph. located to the far right of the window. She is left then with the list of citations included in that portion of the monograph and with the visual representation of the note spans and the text spans that are not supported by citations on the main scene. She notices that the lengthiest span of text is preceded by a succession of complex footnotes with more than two citations. She skims through both the span of text with no citations (which she identifies as the location where the author describes her theory about the sources of the Victorian novel) and the note spans that precede it and which contain references to competing ideas from the field. To get a clear idea about how these references relate to each other and to the ideas of the citing author, Erica switches to the third view (hierarchical view) where she can see these relations more clearly.

By going back to the second view of the tool and repeating the analysis for the second largest span of text with no citations, she discovers that the author of the monograph used two different approaches: in the first, she described at length in text the views that might undermine her theories, cited them amply and included in the footnotes works and authors that didn’t agree with those ideas, only to dissect and disprove them at length in the text span where she exposes her own opinions and ideas on the matter, while in the second case she discussed at length both the ideas for and against her theory, with an emphasis on the ideas that were closer to her own.

By repeating this analysis for other portions of the monograph as well as for other monographs in the field, Erica might get an idea about how argument is built in the filed of Victorian novel studies.

Dear all,

Please find attached a more detailed scenario of research using our tool. It is not intended to be comprehensive, as the tool will afford different avenues of inquiry, but it shows how the tool might be used in a particular instance.

I am going to expand on this some more, and I am not sure how useful it is at this point, but Jennifer asked me to describe how users might transition between the different views of the tool, so here it is:

* Erica is asked to consult with the Social Sciences and Humanities Research Council for the development of a set of standards for a potential Canadian version of the UK Research Assessment exercise, in which scholars, departments and universities are assessed periodically in terms of academic performance and scholarly excellence. Her task is to assess the weigh citation counts should play in the analysis of academic performance in the humanities and to support her recommendations in front of a board composed of non-specialists. She is employing the visualization to demonstrate the complexity of attribution in the humanities in terms of role, origin, and time span of sources cited. She uses comparison to determine the degree of predictability of citation patterns in humanist scholarship, thus the relevance of citation counts for measuring academic performance.
* For the graduate seminar she is facilitating, Erica uses the visualization tool to map the development of the argument in the discussion chapter of an archaeology monograph for the benefit of her students who have very diverse backgrounds. She has gathered information about the significant, controversial, or often cited works in said chapter and is using the tool to illustrate how the monograph represents a turning point from one school of archaeological thought to another by visualizing aspects of citation like language, age, type of publication, function in context, etc – while going back and forward between the text of the monograph and the graphic visualizations of citations.

# Raheem Premji – PhD student in Ancient History

Raheem is a 1st year PhD student in Ancient Mesopotamian History. He is still in the initial phase of his PhD studies but is quite advanced in gathering sources for his intended research. He usually takes detailed notes when reading and keeps them organized by subject of interest and intended purpose. Raheem is still adapting to the North-American scholarly environment and customs , and is a little uncertain about what it’s appropriate / desirable to cite in his field or not. (i.e. How can he demonstrate his expertise on the subject as a young, emerging scholar)

Raheem is quite interested in visualization and is aware of some projects that apply this technique to his field but would characterize himself as a novice in the area of information visualization. He heard of our citation visualization tool from a talk on campus and he is aware that the database of monographs linked to the tool includes some writings in his area of interest.

## Goals:

* Have a visual, interactive representation of the sources cited in a monograph
* Have continuous access to the text of the monograph while the visualization runs
* Be able to filter, locate in text and compare different features of the materials cited (e.g. date of publication, type of publication, function, etc.)
* Be able to contrast and compare the function of citations from different authors (positive or negative view)

## Scenarios:

* Recently, Raheem has been invited to contribute a book chapter on the subject of “Neo-Babylonian slave selling contracts” and he is a bit anxious as this is his first substantial contribution. He is a little worried about attribution because in his field, aside from a number of passing, often conflicting recommendations from his research supervisors and peers, there doesn’t seem to be a clear, widely accepted attribution policy (if you discount the citation conventions for primary sources).  
  He is using our visualization tool to look at significant monographs in his field (which are included in our collection of marked up monographs) hoping that the visualization will help him identify the do’s and don’t of citation from consecrated authors in his field. Some of the things he is looking at are whether or not it is appropriate to cite older secondary sources, or to which extent are digital sources (like the database of Neo-Sumerian texts) cited in “serious” scholarship. On the finer points of the argument, Raheem is also looking at how much space is customary to spend disproving opposed theories and how much of the discussion is usually relegated to the footnote.

Lidia Silverio – undergraduate science student

Lidia is a 2nd year Pre-med student. She is fulfilling her humanist classes requirements by taking a class in 19th century English novel. She is far more used to reading scientific literature containing parenthetical citations and finds the long readings in her humanist class difficult to follow, especially because of the extensive footnotes and endnotes which she feels are interrupting her reading flow thus hampering her understanding of the arguments.

## Goals:

* Have a visual, interactive representation of the sources cited in a monograph
* Being able to switch smoothly between the full text of the monograph and the visualization of citations
* Being able to visualize at a glance the structure of complex footnotes with multiple citations. (which would increase her understanding of the text – or at least the speed of her understanding)

## Scenarios:

She heard from a colleague that one of the required readings in the English novel class is included in this visualization tool. As she is quite familiar with other forms of information visualization from her Biology classes, she decides to give it a try in the hope that using the tool as a reading aid will help her navigate the annotations more easily and understand and read the text faster.