

## LaTeX Example Doc

Changrui Li

CS 800

*Current draft: 2/9/26 at 8:16pm EDT*

Look at the source main.tex to see how this is done.

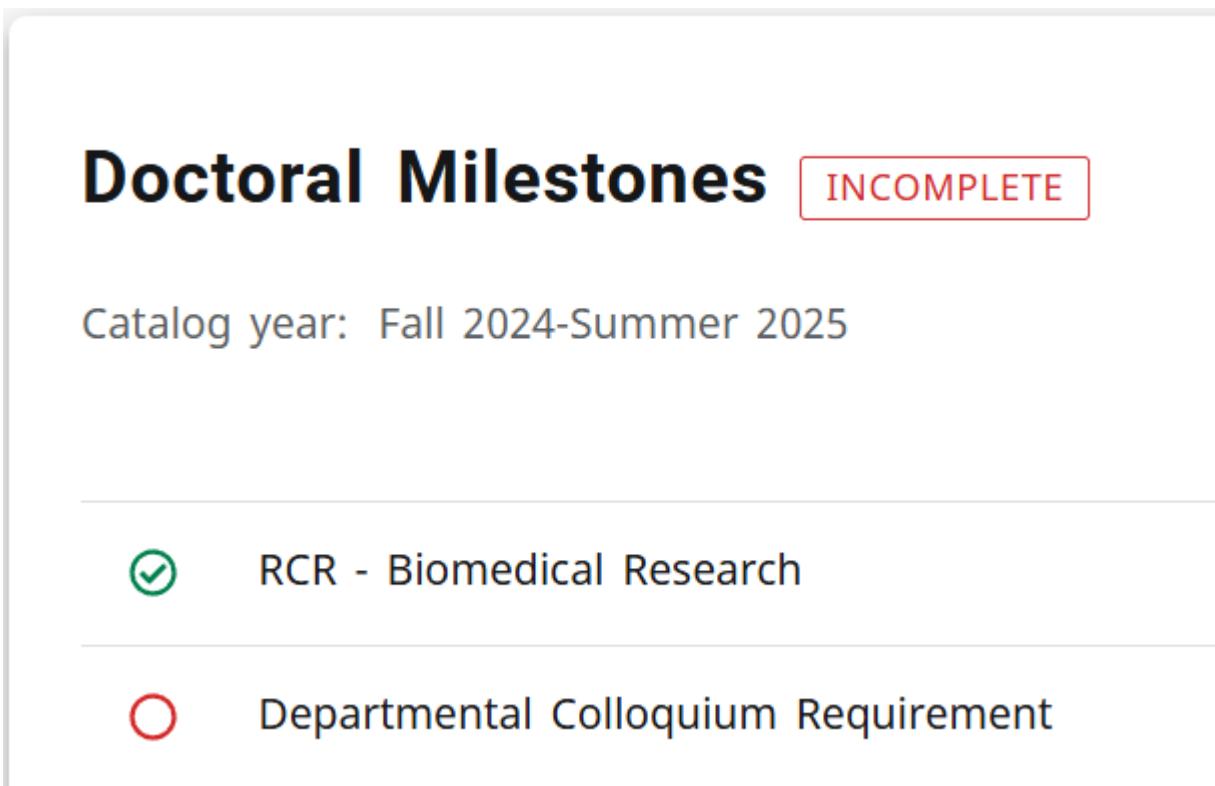
## 1 URIs

This is a formatted, clickable link to my webpage: [https://www.cs.odu.edu/~cs\\_cli027/](https://www.cs.odu.edu/~cs_cli027/)

## 2 Images

All figures must have a caption and must be referenced in the text. See the example below.

Figure 1 shows an original PNG with no scaling or cropping. The original dimensions are 609 x 390. Figure 2 shows an example of cropping the image using the `trim`, `clip` options to `includegraphics`.



**Figure 1:** Original PNG

Figure 3 shows the same cropping as Figure 2 but scaled up. It's blurry because the original image (Figure 1 was a low resolution.)

We can insert PDFs into the document in the same way as images. Figure 4 is the first page of

# Doctoral Milestones

INCOMPLETE

Catalog year: Fall 2024-Summer 2025



RCR - Biomedical Research

**Figure 2:** Cropped PNG - 0.25in from left, 0.5in from bottom, 1in from right, 0.3in from top

# Doctoral Milestones

INCOMPLETE

Catalog year: Fall 2024-Summer 2025



RCR - Biomedical Research

**Figure 3:** Cropped and scaled PNG

an academic paper. I've added the `\frame` command to show where the boundaries are. Figure 5 shows the margins trimmed off so that the text can be larger (scaled up).

## 3 Quotation Marks

Quotation marks are weird in LaTeX. Here's using "double quotes". *Not quite right.* Here's the "proper way". It's two backticks and two single quotes: ``proper way''

## 4 Tables

Table 1 shows a simple example table. Table 2 shows an example confusion matrix from [https://en.wikipedia.org/wiki/Confusion\\_matrix](https://en.wikipedia.org/wiki/Confusion_matrix). This employs rows that span multiple columns

(multicol) and columns that span multiple rows (multirow).

**Table 1:** Simple Table

Week	Date	Topic
1	Sep 1, 3	Introduction, What's Vis and Why Do It?
2	Sep 8, 10	Data and Data Cleaning
3	Sep 15, 17	Marks and Channels

**Table 2:** Example Confusion Matrix from Wikipedia

		Actual	
		Cat	Dog
Predicted	Cat	5 (TP)	3 (FP)
	Dog	2 (FN)	3 (TN)

**Figure 4:** Inserted PDF

## Understanding Computational Web Archives Research Methods Using Research Objects

Emily Maemura, Christoph Becker  
*Digital Curation Institute  
 University of Toronto  
 Toronto, Canada  
 e.maemura@mail.utoronto.ca,  
 christoph.becker@utoronto.ca*

Ian Milligan  
*Department of History  
 University of Waterloo  
 Waterloo, Canada  
 i2millig@uwaterloo.ca*

**Abstract**—Use of computational methods for exploration and analysis of web archives sources is emerging in new disciplines such as digital humanities. This raises urgent questions about how such research projects process web archival material using computational methods to construct their findings.

This paper aims to enable web archives scholars to document their practices systematically to improve the transparency of their methods. We adopt the Research Object framework to characterize three case studies that use computational methods to analyze web archives within digital history research. We then discuss how the framework can support the characterization of research methods and serve as a basis for discussions of methods and issues such as reuse and provenance.

The results suggest that the framework provides an effective conceptual perspective to describe and analyze the computational methods used in web archive research on a high level and make transparent the choices made in the process. The documentation of the research process contributes to a better understanding of the findings and their provenance, and the possible reuse of data, methods, and workflows.

**Keywords**-computational methods; web archives; research objects; computational archival science; digital curation

### I. INTRODUCTION

The web has become a key source for scholars studying social and cultural phenomena. Web archives facilitate this work by documenting and preserving snapshots of the web for future research. In areas such as history, scholars have unprecedented opportunities to leverage access to massive primary sources of born-digital artifacts to make sense of individual experiences [1].

These new forms of archives trouble traditional conceptions of historical archival research, which assumes that any material found in archives is historically significant, having been assessed through archival appraisal. The scale of web archives confounds the traditional appraisal process. Web archives are replete with information that may not be significant to the research questions being asked.

The scale of these collections frequently mandates the use of big data analysis techniques to address their research questions. The adoption of these computational methods is a major shift for fields in the humanities. Approaches that have

traditionally focused on close readings of source materials are moving to the use of distant reading methods [2].

We highlight three critical factors to be considered for humanities researchers using web archives:

- 1) *interrogating sources*. – Understanding how these web archives collections were created is key to judging the adequacy, appropriateness and limitations of the source material.
- 2) *understanding new methods*. – The use of emerging computational methods is a key prerequisite to working with large scale data sets.
- 3) *transparency of the research process*. – Findings are dependent on the validity of the computational processes and the adequacy of data.

These factors contribute to the need for a stronger methodological framework for research with web archives to understand the research process in more detail. This can serve as a common vocabulary for discussions of trust in the findings, as well as reuse of data, tools, or analytical techniques.

Our work combines perspectives of a digital humanist engaged in the computational exploitation of web archives and scholars working in the intersection of systems design and digital curation. Our joint interest in research with web archives led to an effort to structure research processes in order to develop detailed descriptions, with an eye to developing a research model. Our motivation for this study is to ask: *how do research projects in digital humanities process web archival material using computational methods to construct their findings?* We follow the work of other researchers in the digital humanities who have called for stronger frameworks for computational methods and the needs of researchers using web archives [3], [4].

To develop the framework, we draw on the concept of the Research Object (RO) developed to address the aggregations of resources and processes used in conducting computational research [5]. We use the RO profile as a structure to characterize three case studies that use computational methods to analyze web archives within digital history research, and

**Figure 5:** Trimmed PDF