



Manika
Lamba

Postdoctoral Research Associate
School of Information Science



Hendrik
Erz

Doctoral Student
Institute for Analytical Sociology



Thanking the World

Exploring Gender-Based
Differences in Acknowledgment
Patterns and Support Systems in
Theses

23rd December 2023





JD Carpentieri @jdcarpentieri · Apr 30, 2019

In my initial draft I thanked my wife and daughter for their support. When they read it they laughed out loud & said “What support?!” They were right – so I rewrote it to say that I would’ve finished years earlier if not for them. (But would have been far less happy!)

1

1

6

...

...



Dr. RebekahD @DRebekahd · Apr 30, 2019

That's fantastic!

1

1

...

...

...



Shane Gleason @shanegleasonPhD · Apr 29, 2019

I thanked my cats. Because they deserved to be thanked. #catsofpolisci

1

1

3

...

...



TaniaM Jenkins@sciences.social @TaniaM Jenkins · Apr 30, 2019

I actually thanked the local Starbucks for fueling my dissertation writing!

1

1

4

...

...



Bryan @bryanmwilcox · Apr 30, 2019

Two dogs, two group text threads, CrossFit all in the acknowledgment section of mine!

1

1

1

...

...



Edward Kammerer @ProfKammerer · Apr 30, 2019

I definitely thanked the staff of the coffee shop I wrote my dissertation in. They were definitely more important to the successful completion than friends and family, and not *that* far behind my committee.

1

1

1

...

...

I would also like to express my deepest gratitude to my family members and friends Sara and Jihene who always stood by my side and encouraged me all along the way. Finally, a special thanks to the biggest love BTS whose existence brought comfort and healing in my life.

Thank you to the people that understand me without even saying anything and that gave me happiness in a swirling day, placing all of the smiles in the world in my hands. To those that still hold my hand with warmth even when I run out of breath on a steep road and even when I lose my path during a cold day.

Introduction

- Lots of research on acknowledgment sections of scientific papers since the 1990s
- Lack of publications analyzing electronic theses and dissertations (ETD) acknowledgment sections (Bangani et al., 2020)
- Cronin et al. (1992) identified positive correlation:
 - More Acknowledgments → More Citations in the future
- Since 2008, WoS has been providing information on funding acknowledgments in WoS-indexed journals
- Currently, there is no database that provides data on ETDs' acknowledgments
- We focus on the field of LIS as a case study
- The significance of this study lies in shedding light on the performance of academic support systems as indicated by the acknowledged types of support
- By uncovering the nuances of support systems and gender dynamics, we aim to contribute insights that can inform mentoring practices and promote gender equality in academia

Research Questions

RQ 1 What are the important support systems for male and female researchers to finish their master's or doctoral research?

RQ2 What are the major sentiments that researchers felt throughout their journey of research?

Methodology

Data

Database: **PQDT Global**

Query: **SU.exact("LIBRARY SCIENCE")**

Period: **1921 to 2020**

Language: **English**

No. of ETDs Downloaded (Manually b/w 2019-2021): **4000**

No. of Acknowledgement Sections: **1252 or 20,834** individual sentences

Data Extraction & Cleaning

- Text Extraction with Tesseract
- Challenges in Acknowledgment Section Extraction
 - Initial Script Using Heuristics Extracted Only 1/3 of Theses
 - Varied Acknowledgment Styles Complicated Extraction
- Manual Section Extraction
 - Due to Variances, Manual Extraction of Sections Required
 - Ensured Comprehensive Corpus Coverage
- Corpus Refinement
 - Narrowed Corpus to Theses from 1960 to 2020
 - Excluded 1921-1959 Period Due to Data Limitations
 - Could not determine the gender of researchers and advisors in this period (missing data)

2.1 Introduction

Interactive computer graphics has advanced to a stage where a choice of hardware and software approaches relative to many potential economic and practical applications can be made with considerable justification. In this chapter, we review commonly used systems in their various forms, and outline a hardware choice which appears the best on economic grounds. For this configuration, software support is discussed.

2.2 Hardware Configurations

The commonest approaches taken in hardware considerations for interactive graphics can be broadly classified under two headings:

- (1) Directly-Coupled Displays- a display console (or console) directly connected to a medium-to-large computer; or,
- (2) Indirectly-Coupled Displays- a display (or displays) coupled to a peripheral computer, which is itself coupled to a larger system.

2.2.1 Directly-Coupled Display

For several years, the cathode ray tube (CRT) display console

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Red: Insertion of Unwanted Characters due to Speckles
Yellow: Insertion of Random Words/Characters not present
Green: Misspelled Words

IV CHARTER II
hardware/software CONSIDERATIONS FOR A DISPLAY console SYSTEM
2.1 Introduction
 Interactive computer graphics has advanced to a stage where a choice of hardware and software approaches relative to many potential economic and practical applications can be made with considerable justification; In this chapter, we review commonly used systems in their various forms, and outline a hardware choice which appears the best on economic grounds. For this configuration, software support is discussed.

The commonest approaches taken in hardware considerations for interactive graphics can be broadly classified under two headings :

- (1) Directly -Coupled Displa spfys- a display console (pp consoles) directly connected to a medium-to-large computer; or,
- (2)

Indirectly -Coupled Displays - a display (or displays) coupled to a peripheral computer, which is itself coupled to a larger system.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Lamba, M., & Madhusudhan, M. (2023). Exploring OCR Errors in Full-Text Large Documents: A Study of LIS Theses and Dissertations. *Library Philosophy and Practice (e-Journal)*. <https://digitalcommons.unl.edu/libphilprac/7824>

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER I

THE PROBLEM AND DEFINITIONS OF TERMS USED

Guilford, Connecticut, with its long history, also possesses a long library history. Although Guilford is particularly rich in this respect, little has been done to preserve and publicize the information concerning the various libraries.

Certain characteristics and trends in library philosophy have continued through the years despite the changes in titles and personnel, and can be traced from the earliest library to the present organization, the Guilford Free Library.

I. THE PROBLEM

Statement of the problem. It was the purpose of this paper to gather together from many sources as possible, all information concerning the development of libraries in Guilford. Using the history of the libraries as a point of departure, six characteristics and trends were traced from their beginnings to their evidences in the present-day library organization, to outline of some study. Prior to this study, the information concerning the libraries of Guilford was scattered in many places. This paper attempts to bring together all the pertinent available data which has appeared previously in print, in addition to information from primary sources which were evidently either overlooked or not known to earlier writers.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

Reproduced with permission of the copyright owner. Further reproduction prohibited without permission.

CHAPTER I

NATURE OF THE PROBLEM

Significance of the Problem

The problem of discovering a good plan for state aid to local school libraries is becoming increasingly important not only because of the need but also because of the recognition of the need for good library facilities. Due to the modern program of education, the library has become the center of the school. From this center radiate the materials for better study, provided the facilities are adequate. The present day departure from basal texts demands a larger supply of supplementary books and modern courses of study are demanding a wider range of material than single textbooks can supply. Hence the library becomes more important in the modern school than it was in the traditional school.

Modern theory believes that the state is not obligated to finance local school libraries alone, but that it is the responsibility of the state to aid the locality in financing them.

The problem of this thesis is to study the different types of state and local aid given to local school libraries within the states of Connecticut, Louisiana, Missouri, Virginia, Wisconsin, Iowa, Maryland, Nebraska, and South Carolina. The plans in the nine states mentioned will be analyzed and desirable features in each of the plans will

Blue: Concatenation of Words

Yellow: Misspelled Words

Green: Space between Words

Page 1

CHAPTER I

NATURE OF THE PROBLEM

Significance of the Problem

The problem of discovering a good plan for state aid to local school libraries is becoming increasingly important not only because of the need but also because of the recognition of the need for good library facilities. Due to the modern program of education, the library has become the center of the school. From this center radiate the materials for better study, provided the facilities are adequate. The present day departure from basal texts demands a larger supply of supplementary books and modern courses of study are demanding a wider range of material than single textbooks can supply. Hence the library becomes more important in the modern school than it was in the traditional school.

Modern theory believes that the state is not obligated to finance local school libraries alone, but that it is the responsibility of the state to aid the locality in financing them.

The problem of this thesis is to study the different types of state and local aid given to local school libraries within the states of Connecticut, Louisiana, Missouri, Virginia, Wisconsin, Iowa, Maryland, Nebraska, and South Carolina. The plans in the nine states mentioned will be analyzed and desirable features in each of the plans will

Yellow: Misspelled Words

Green: Insertion of Random Words/Characters not present in PDF

Blue: Insertion of Hyphen

Red: Concentration of Words

Orange: Insertion of Space between Words

CHAPTER I INTRODUCTION AND THE PROBLEM

"J's resulted in a lack of coherent management of one of the most important items in the music educators' repertoire. It is true that personnel, as well as physical and financial limitations and variances, tend to create individual administrative problems; however, the "... successive steps which correspond to the journey a composition takes from the time it is ordered until it has been used in the band [music] room and replaced in the files," (22:165) should be more fully investigated and discussed in the hope that a unified system of cataloging and filing instrumental music

I . THE PROBLEM
 Statement of the problem. It was the purpose of this study (1) to compare several existing methods of cataloging and filing instrumental music; (2) to determine the best procedure inherent in each of the systems examined; and (3) to recommend a catalog and file system emanating

Yellow: Insertion of Random Words/Characters not present in PDF
 Green: Space between Words
 Blue: Misspelled Words

Magenta Box: Sentences are skipped

Qualitative Coding

Qualitative Coding

A. Support Coding

Category	Definition
Academic Support	It involves guidance from Professors or Doctors who, as supervisors or committee members, provide crucial assistance to students or researchers. Acknowledgment is typically done using their full names and academic titles.
Moral Support	It encompasses encouragement and comfort provided by colleagues, friends, and family in private settings. This form of support is typically expressed using first names and serves to uplift and strengthen individuals emotionally during challenging times.
Financial Support	It refers to assistance provided by funding agencies or universities to enable individuals to complete their academic theses or projects. This support involves monetary contributions that cover various expenses associated with research, education, or specific project needs.
Technical Support	It involves assistance with specific skills or tasks related to writing, proofreading, and adopting new techniques. It includes guidance and help provided to individuals to enhance their technical abilities, refine written work, and acquire proficiency in novel methods or approaches.

Access to Data	It refers to the provision of assistance or support in obtaining the necessary information or datasets required for a particular purpose. This support may involve facilitating the acquisition, retrieval, or permission to use specific data sets, enabling individuals or organizations to gather relevant information for analysis, research, or other purposes.
Religious Support	It involves drawing on quotes from religious texts and making references to the names of deities or higher powers for inspiration, guidance, or comfort. This form of support often includes expressions rooted in faith and spirituality to provide strength and solace in various situations.
Library Support	It refers to the assistance provided by libraries and librarians to help individuals complete searches, tasks, or research endeavors. This support may involve guidance on navigating library resources, locating relevant materials, and utilizing research tools effectively to enhance the quality and efficiency of the information retrieval process.
Other	It refers to miscellaneous or additional information not covered by specific categories. In this context, it encompasses diverse sentences or details that do not fit into predefined categories.

B. Sentiment Coding

"very positive", "somewhat positive", or "neutral"

- **Creating a "Gold-Standard" Dataset**
 - Sampled 903 Sentences from the Corpus
 - 4 Coders Manually Labeled Support Type and Sentiment
 - Represented Approximately 4% of the Total Corpus

Automated Coding

Automated Coding

- Employed Large Language Model (LLM) for Classification
- Applying Labeling Strategy to the Entire Corpus
 - Utilized a Sequence-Classification Task
 - Leveraged RoBERTa-base models (Liu et al., 2019) to Learn and Apply Human Annotations (Vaswani et al., 2017)

Table 2: Metrics for the support category classifier (rounded). Categories are: academic, moral, technical, data access, library, financial, religious, and unknown/other support. Best model is number 14.

#	F1 (avg)	F1 (acad)	F1 (moral)	F1 (tech)	F1 (data)	F1 (lib)	F1 (fin)	F1 (rel)	F1 (other)	Accuracy
1	0.22	0.76	0.71	0.0	0.0	0.0	0.0	0.0	0.27	0.65
2	0.27	0.78	0.77	0.0	0.0	0.0	0.0	0.012	0.58	0.70
3	0.36	0.79	0.78	0.0	0.0	0.0	0.0	0.67	0.64	0.72
4	0.38	0.78	0.73	0.0	0.0	0.25	0.0	0.67	0.59	0.69
5	0.5	0.82	0.68	0.0	0.0	0.6	0.55	0.67	0.68	0.72
6	0.5	0.85	0.71	0.25	0.0	0.33	0.55	0.67	0.62	0.70
7	0.57	0.82	0.70	0.33	0.33	0.33	0.67	0.67	0.68	0.71
8	0.57	0.85	0.72	0.4	0.4	0.31	0.6	0.67	0.6	0.72
9	0.59	0.83	0.70	0.29	0.4	0.44	0.75	0.67	0.62	0.72
10	0.58	0.83	0.71	0.29	0.4	0.55	0.6	0.67	0.60	0.71
11	0.55	0.80	0.67	0.25	0.33	0.36	0.67	0.67	0.64	0.69
12	0.51	0.81	0.68	0.0	0.33	0.33	0.67	0.67	0.60	0.69
13	0.53	0.81	0.69	0.0	0.33	0.46	0.67	0.67	0.63	0.69
14	0.60	0.81	0.68	0.44	0.4	0.46	0.67	0.67	0.68	0.71
15	0.60	0.80	0.68	0.44	0.4	0.46	0.67	0.67	0.64	0.70

Table 3: Metrics for the sentiment classifier (rounded). Best model is number 9.

#	F1 (average)	F1 (very positive)	F1 (somewhat positive)	F1 (neutral)	Accuracy
1	0.36	0.48	0.45	0.15	0.44
2	0.43	0.36	0.78	0.16	0.66
3	0.43	0.04	0.77	0.47	0.64
4	0.58	0.51	0.76	0.45	0.67
5	0.48	0.21	0.76	0.46	0.64
6	0.55	0.45	0.77	0.44	0.67
7	0.58	0.49	0.76	0.49	0.67
8	0.57	0.51	0.74	0.46	0.65
9	0.59	0.52	0.75	0.49	0.66
10	0.57	0.52	0.76	0.41	0.67
11	0.56	0.5	0.71	0.47	0.63
12	0.57	0.51	0.74	0.46	0.65
13	0.56	0.45	0.75	0.49	0.66
14	0.55	0.45	0.74	0.47	0.65
15	0.56	0.46	0.74	0.47	0.65

Liu, Y., Ott, M., Goyal, N., Du, J., Joshi, M., Chen, D., Levy, O., Lewis, M., Zettlemoyer, L., & Stoyanov, V. (2019). RoBERTa: A Robustly Optimized BERT Pretraining Approach. arXiv:1907.11692 [Cs]. <http://arxiv.org/abs/1907.11692>

Vaswani, N., Shazeer, N., Parmar, J., Uszkoreit, L., Jones, L., Gomez, A. N., ... Polo-sukhin, I. (2017). Attention Is All You Need. arXiv:1706.03762 [cs]. arXiv:1706.03762

Gender Analysis

Gender Analysis

- Gender determination for supervisors and researchers which did match almost all names
- Utilized a list of commonly female & male given names from the corresponding data files that can be found at
<https://www.geeksforgeeks.org/python-gender-identification-by-name-using-nltk/>
- The gendered nouns ("he", "she", etc.) have been collected from the data files "male_pairs.txt" and "female_pairs.txt" from this github repository:
<https://github.com/nikhgarg/embeddingdynamicstereotypes/tree/master/data> (Garg, N., Schiebinger, L., Jurafsky, D., & Zou, J. (2018). Word embeddings quantify 100 years of gender and ethnic stereotypes. *Proceedings of the National Academy of Sciences*, 115(16), E3635–E3644. <https://doi.org/10.1073/pnas.1720347115>)
- Enhancing our understanding of gender-based discrepancies

Findings

Sentiment Analysis

- ***Sentiment Distribution***
 - Most Sentences: "Neutral" Sentiment
 - Followed by "Somewhat Positive"
 - Lastly, "Very Positive"
- ***Formalized Acknowledgment Style***
 - Limited Sentiment Variation
 - Suggests a Consistent, Formalized Style
 - Allows for Academic Expression with Emotional Tone

Support Systems

- ***Most Frequent Support Categories***
 - Academic Support: 26.72%
 - Moral Support: 11.22%
 - Library Support: 3.40%
- ***Emerging Trend: Religious Support***
 - Historically Insignificant
 - Increasing Frequency in the 21st Century
 - Notable shift in acknowledgment patterns
- ***Diversity in Support Systems***
 - "Other" Category Dominant: 53.93%
 - Support Systems are more diverse than expected

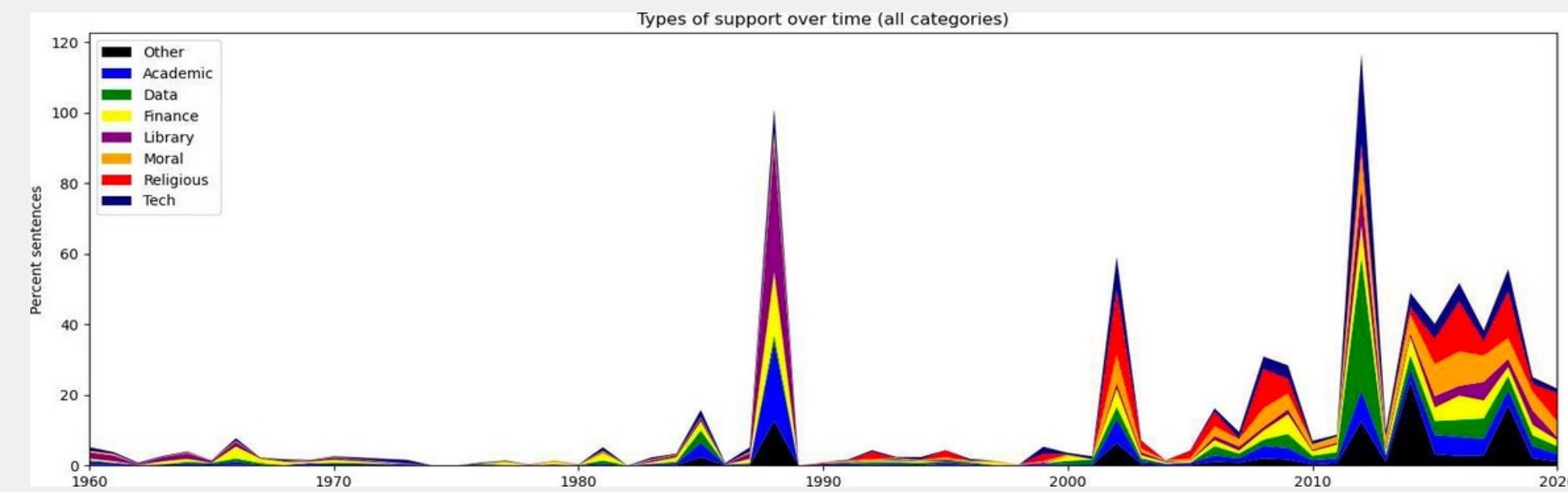


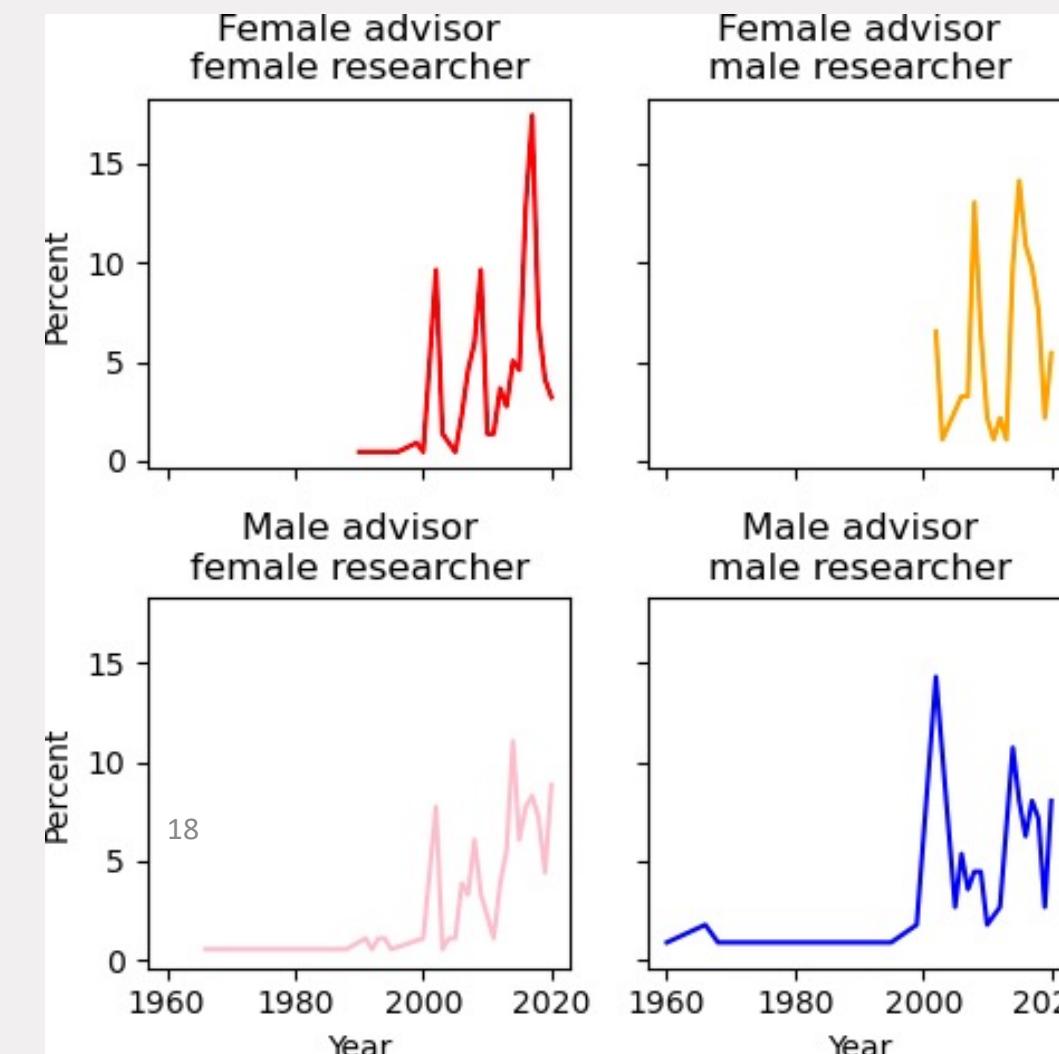
Fig: Evolution of the acknowledgment of various types of support systems over time (1960-2020)

Gender Analysis

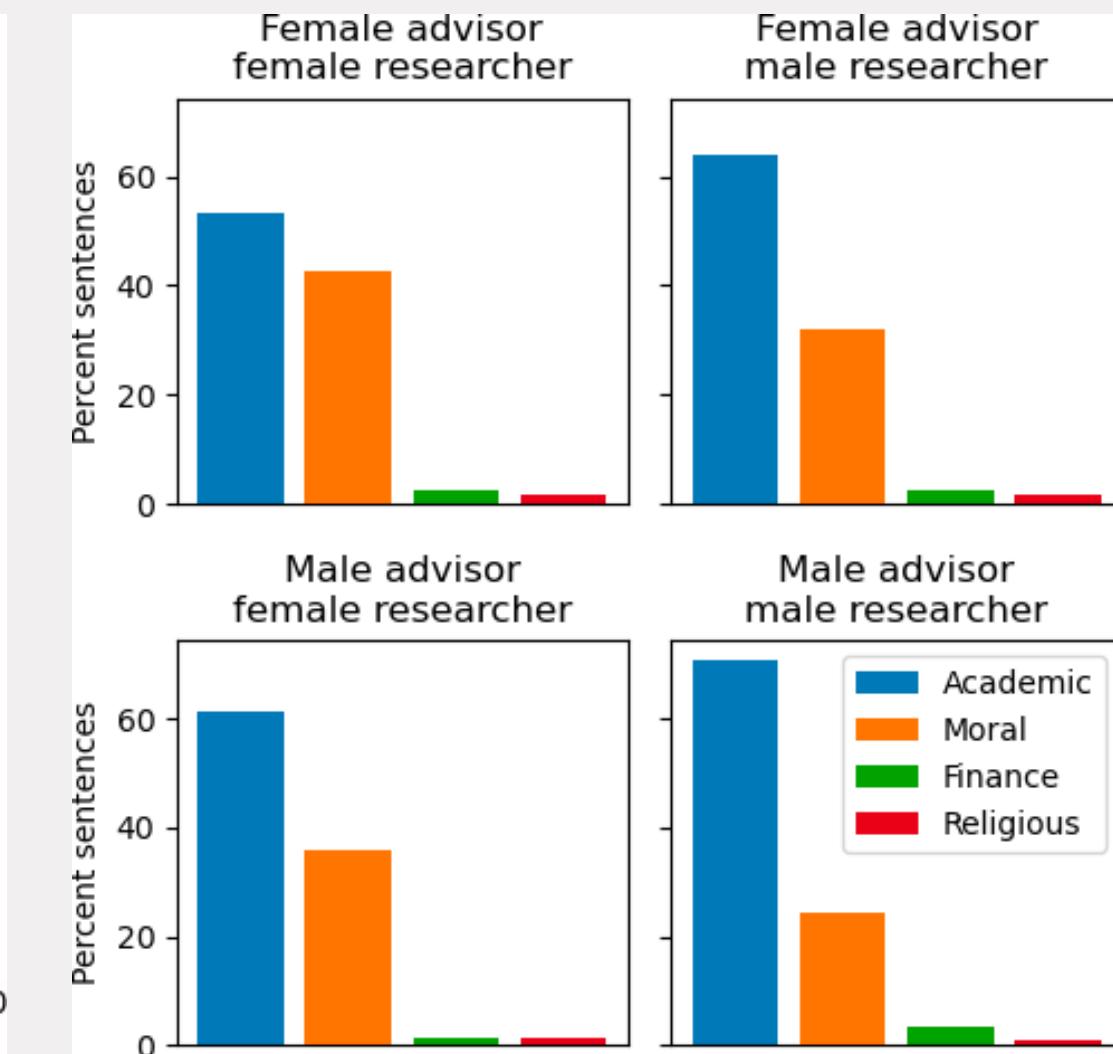
- After omitting researcher-advisor pairs where any participants' gender could not be determined, we were left with 603 pairs

- Gender Distribution***

- Advisors' Gender Balance:**
 - Female: 51%
 - Male: 49%
- Researcher Predominance:**
 - Female Researchers: 66%
 - Male Researchers: 34%



(a) Breakdown of advisor-researcher pair proportions over time (1960-2020)



(b) Breakdown of advisor-researcher pair proportions by type of support (1927-2020)

Figure 2: Analysis of acknowledgment sections by gender, broken into the four combinations of female-female, female-male, male-female, and male-male

Gender Analysis

- Regardless of Gender, **Advisors Prefer Same-Gender Researchers** (*Fig. 2a*).
- **Female researchers** tended to acknowledge both their family/friends and supervisors/committee members more frequently than male researchers in their theses
- **Female researchers** acknowledge their academic support system more than their moral support system (60% versus 40%) when they are supervised by a male advisor as opposed to when they are supervised by a female advisor (50% versus 40%) (*Fig. 2b*)
- **Male researchers** acknowledge their moral support system more when they are supervised by a female supervisor (60% versus 30%) as opposed to when they are supervised by a male supervisor (70% versus 20%) (*Fig. 2b*)
- **Religious Support:**
 - Male vs. Female Supervisor: No strong gender-based differences
- **Financial Support:**
 - **Male Researchers Supervised by Male Advisors:**
 - More frequent acknowledgment of **financial support**

Gender Analysis

- ✓ Understanding these gender-based differences in acknowledgment patterns provides insights into the complex dynamics of support systems and the unique challenges faced by researchers of different genders
- ✓ This finding raises important questions about gender disparities in research supervision within the field of LIS
- ✓ It highlights the need for further investigation and intervention to ensure equitable support and opportunities for all researchers
- ✓ This suggests that female researchers may rely on a broader support network encompassing both personal and academic spheres

Limitations

Focus on the field of LIS may limit the generalizability of the findings to other academic disciplines

1

Analysis of acknowledgment sections relies on the accuracy of the data extraction process, which may be subject to errors, particularly in cases of OCR inaccuracies

2

There might be errors during the assignment of gender to the advisors' and researchers' names

3

Even though the metrics for our classifier show strong support for our interpretations, better results might be obtained by annotating more examples, e.g., via utilizing Active Learning

4

Future Directions

Exploring Linguistic Discourse and Potential Biases

- Investigate **adjectives** used when acknowledging female vs. male supervisors
- Analyze whether **certain adjectives** exhibit gender-associated patterns
- Examine potential biases in expressions of gratitude

1

Tracking Language Evolution Over Time

- Study the evolution of language in acknowledgment sections over the years
- Analyze **shifts** in expressions of **gratitude** and recognition
- Uncover **linguistic trends** and changes in academic writing

2

Correlation with Academic Achievements

- Investigate the connection between acknowledgment section vocabulary and academic success
- Identify specific **vocabularies linked** to subsequent citation recognition

3

Discerning Variations Across Disciplines and Languages

- Explore variations in acknowledgments across diverse academic fields
- Analyze differences in language use in acknowledgment sections in various languages

4



Thank You