



REPORT FOR QUALIFYING EXAM

Understanding the flow of Temporal Context Shift in Data Voids

Manisha Keim

manisha-keim@uiowa.edu

Advisor

Dr. Rishab Nithyanand

rishab-nithyanand@uiowa.edu

Department of Computer Science,
The University of Iowa

Contents

1	Abstract	3
2	Introduction	3
3	Research Papers	3
3.1	Statistically Significant Detection of Linguistic Change	3
3.2	Specious Sites: Tracking the Spread and Sway of Spurious News Stories at Scale	4
3.3	Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change	4
4	Conclusion	4

1 Abstract

While browsing information online, users come across search results which are full of misinformation, narratives and conspiracy theories which are exposed to users which could lead them to believe or share this misinformation. However, not all queries return this kind of behavior, there are certain queries which causes these kind of issues, these queries are called as data voids. Data voids are identified while searching terms for which relevant information is either limited, non-existent, or deeply problematic. For example, unlike a typical web search for "basketball" that yields a vast amount of data, data voids present a problematic scenario such as searches on these terms may return no results or irrelevant, often inaccurate information. Broadly, data voids can be of two types - firstly, there can be new terms which were never seen before and other can be terms which have changed their meaning over time. For this report, we are focusing on the latter one, to understand the behavior of data voids where context drift is happening, following paper have been chosen:

- [?]

2 Introduction

- What is a Data Void?
- Example of a data void?(Migrant Caravan)
- Lifecycle of data void (from long tail to spike)
- Types of Data Void
- Technical Challenge - analyzing data void involves context drift over historical time. (How context drift happens with data void.)

3 Research Papers

3.1 Statistically Significant Detection of Linguistic Change

- About
- Research Questions
- Methodology
- Results

3.2 Specious Sites: Tracking the Spread and Sway of Spurious News Stories at Scale

3.3 Diachronic Word Embeddings Reveal Statistical Laws of Semantic Change

4 Conclusion

Mention how these paper helped me in this project.

References