

# Network Analysis and Topic Modelling to Identify Influential Authors

## Assessing the Impact of Law Research in South African Courts



### INTRODUCTION

The Faculty of Law at the University of Pretoria (UP) is interested in understanding the underlying insight of the Fontes Juris data. In assisting the faculty on their mission, the project aims at using machine learning (ML) techniques to assess the impact of Law Research in SA courts.

### RESEARCH PURPOSE

The purpose of this research is to assess the impact of law research in South African courts using machine learning techniques. Network Analysis and Topic Modelling are employed. The goals of the research are:

- Identify and rank authors using citations
- Identify and analyse topics cited

### METHODOLOGY

The initial task was to transform unstructured data from the Fontes Juris book that contains source references of the law noted in South African court cases from 1828 to 2015 stored in a pdf format to structured data. Feature engineering and EDA was done on structured data. The next steps topic modelling and network analysis..

### EXPERIMENT

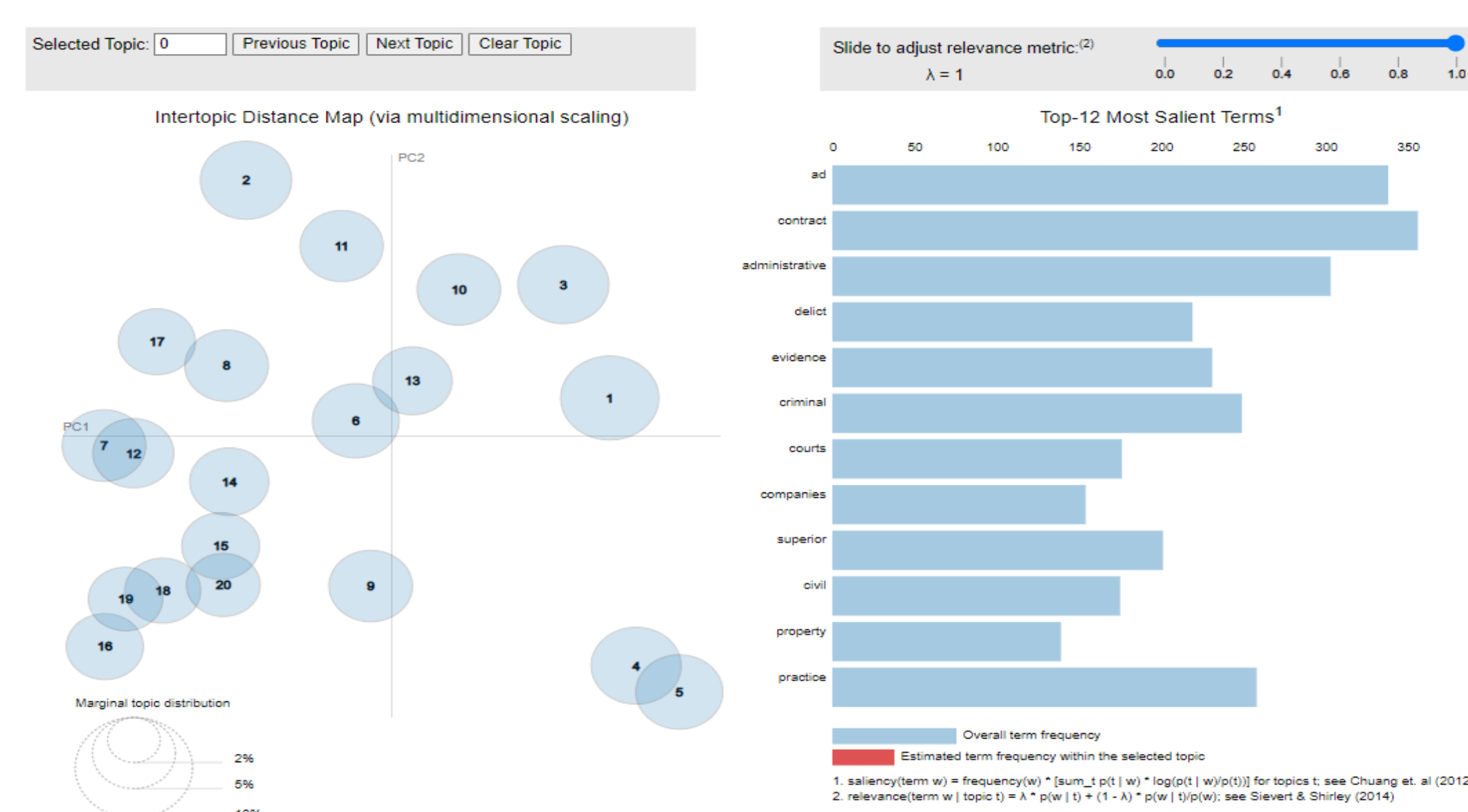
The first experiment was to conduct topic modelling using LDA and identifying an appropriate number of topics that represent the corpus of publications cited.

The second experiment was then to model node level network analysis conducted on author and judgments pair and on author and topic pair. The following centrality measures were calculated to identify influential authors:

- Degree Centrality
- Eigen Vector Centrality
- Closeness Centrality

### RESULTS

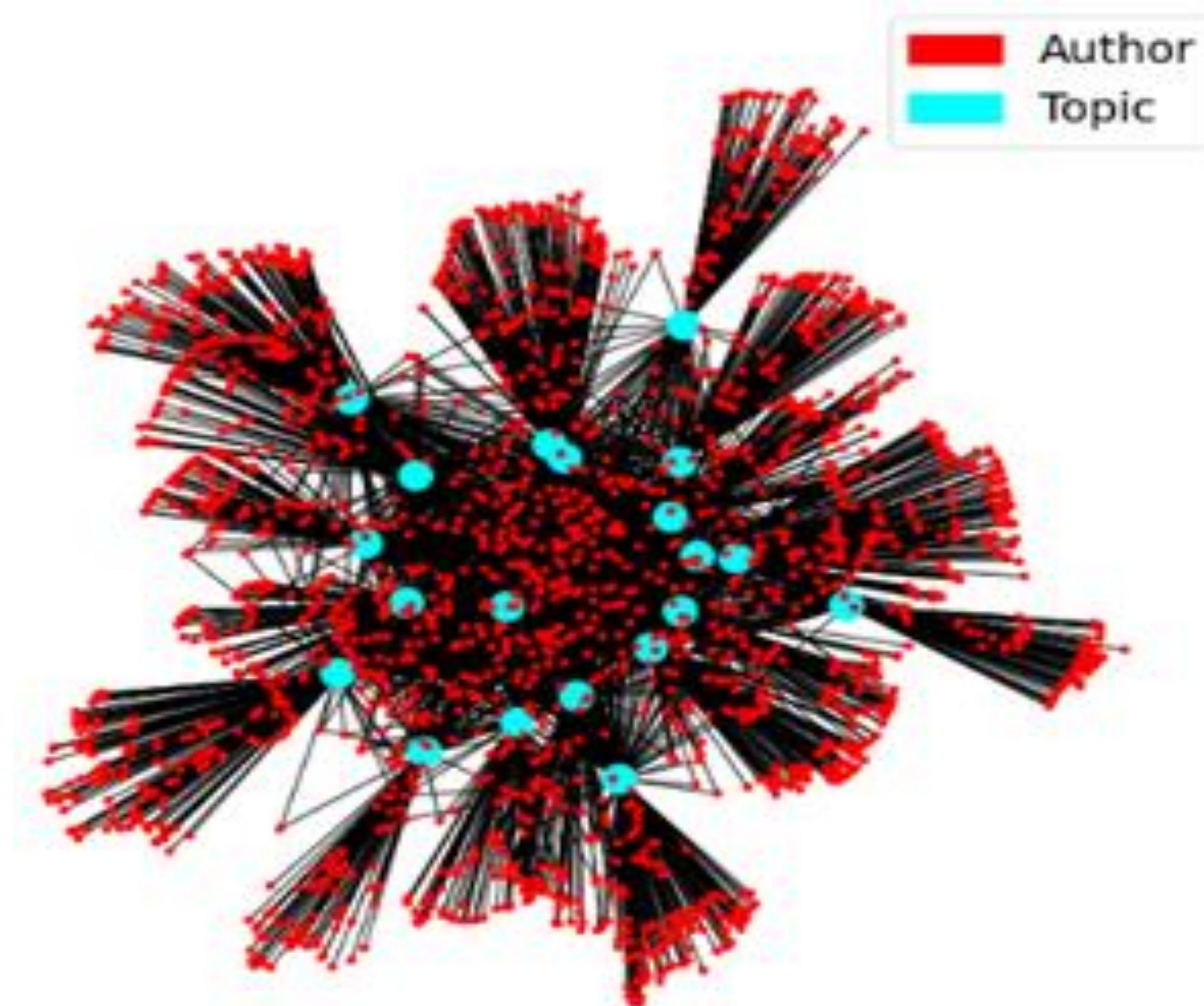
- Top 5 influential authors based on degree of centrality **VOET J, HALSBURY, DIGEST OF JUSTINIAN, VAN LEEUWEN J** and **GROTIUS**.
- **Publications** cited in court judgments reduced in recent years and is not topic specific.
- Topic modelling shows large influence of "Old Laws", "Roman-Dutch" and "English Law" cited in court judgments.
- Over the period 2000-2015 South African law research were cited more containing the following top topics: "Criminal law", "Administration law" and "Company act".



Topic distribution (2000-2015)

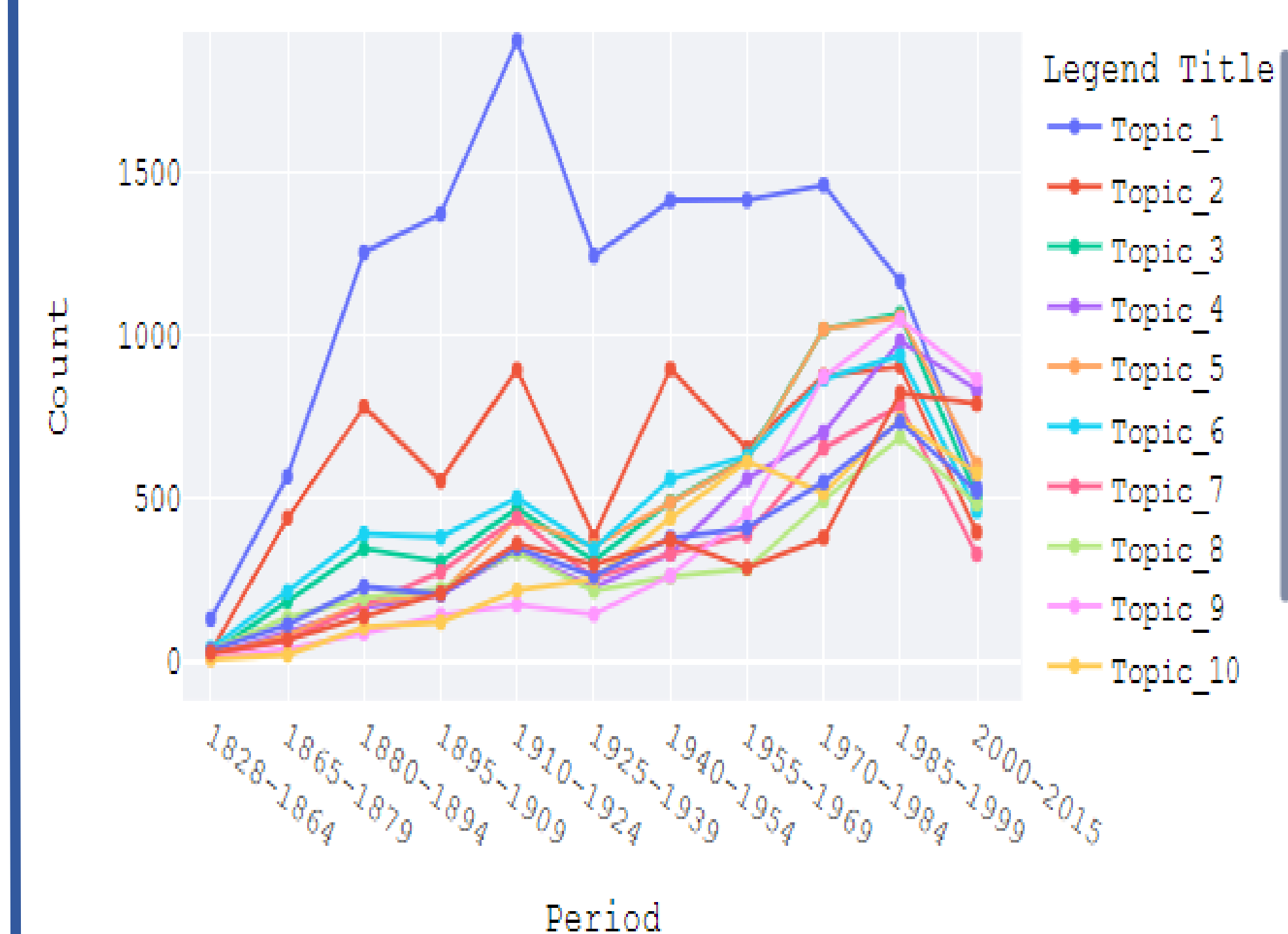


PUBLICATION WORDCLOUD (2000-2015)



AUTHOR & TOPIC NETWORK

Citation count per 15 year period



TOPICS OVER THE YEARS

Arno du Preez, Moyahabo Rabothata

Department of Computer Science

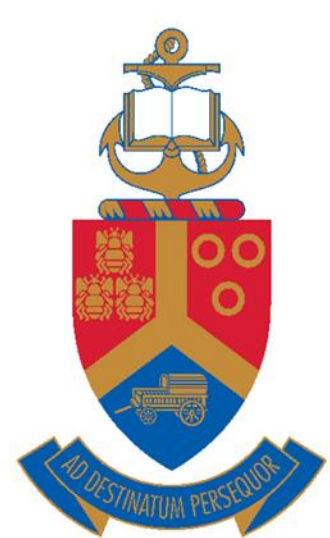
Faculty of Engineering, Built Environment and Information Technology

Fakulteit Ingenieurswese, Bou-omgewing en Inligtingstegnologie / Lefapha I a Boetšenera, Tikologo ya Kago I e Theknolotši ya Tshedimošo

Capstone Project – MIT 808

Course Coordinators:  
Dr. Vukosi Marivate (vukosi.marivate@cs.up.ac.za)  
Abiodun Modupe (abiodun.modupe@cs.up.ac.za)

Scan me



UNIVERSITEIT VAN PRETORIA  
UNIVERSITY OF PRETORIA  
YUNIBESITHI YA PRETORIA