Master's Project

Bias and Fairness in Digital Archival System

Rafael Biehler
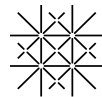
Supervised by:
Prof. Dr. Heiko Schuldt
Fynn Farouz

Abstract

Todo

Contents

University
of Basel
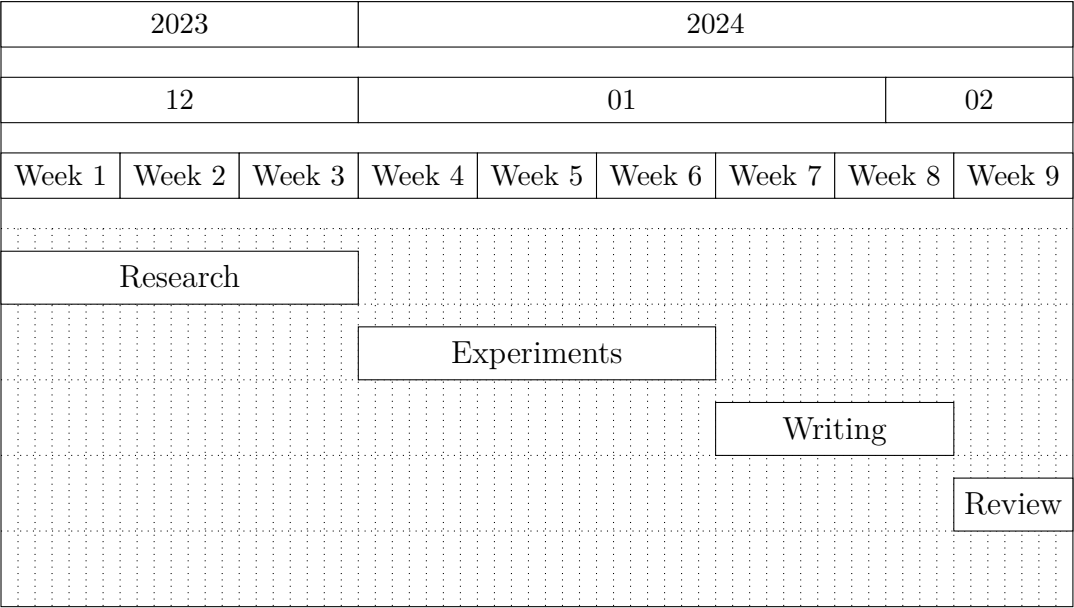
Center for
Innovative Finance

Plagiatserklärung

Ich bezeuge mit meiner Unterschrift, dass meine Angaben über die bei der Abfassung meiner Arbeit benutzten Hilfsmittel sowie über die mir zuteil gewordene Hilfe in jeder Hinsicht der Wahrheit entsprechen und vollständig sind. Ich habe das Merkblatt zu Plagiat und Betrug vom 22. Februar 2011 gelesen und bin mir der Konsequenzen eines solchen Handelns bewusst.

Rafael Biehler

1   Time Management

The following chart depicts my schedule from today (15.12.2023) to the submission date (12.02.2024). Research and Writing are spread around to each three weeks, because in this time I will be studying and working on other lectures as well. Writing and the final review process are relatively short, as I will be able to work full-time.

2023			2024					
12			01					02
Week 1	Week 2	Week 3	Week 4	Week 5	Week 6	Week 7	Week 8	Week 9
Research								
			Experiments					
						Writing		
								Review

2   Introduction

In the context of Galleries, Libraries, Archives, and Museums (GLAM), a prevalent challenge is the imbalance between the volume of work and the available staffing. This often leads to a substantial amount of image data being left unlabeled or uncategorized. To mitigate this issue, there is increasing interest in the use of multimedia retrieval systems that enable querying image data by using natural language descriptions.

This paper will examine the fairness and bias by using a dataset from the GALT Museum & Archives in Lethbridge. This will be done by utilizing the Contrastive Language-Image Pretraining (CLIP) model for

classification tasks to explore whether First Nations people are unfairly disadvantaged in the search results.

We plan to establish a methodology for assessing the fairness and bias using our dataset. This allows us to measure and compare different models and model versions in the future.

TODO: Write exact content of paper

3 Background

Todo

4 Methods

Todo

5 Discussion

Todo

References