Contents

[**1. Introduction**](#_w76yc1qlr7fh) **1**

[1.1. Project Context](#_imgg5oem7seq) 1

[1.2. Problem Description](#_1zkvqeh16zx1) 1

[1.3. Objectives](#_ke9udhwte1a7) 1

[1.4. Project Structure](#_tu82ghe20fta) 1

[**2. Context Survey**](#_ux9dq4grkvvw) **2**

[2.1. Legal Background](#_t46oe0kb6nwx) 2

[2.1.1. Privacy and the GDPR](#_o66tngs5ceip) 2

[2.1.2. Consent](#_90oyty1c3wr5) 2

[2.1.3. Legitimate Interests](#_rqmyzruzhbxk) 2

[2.2. Computer Science Background](#_lao6bwpsrtqv) 2

[2.2.1. Standardisation of Privacy Policies](#_qjeh0mg3ttwa) 2

[2.2.2. Technical Analysis of Cookie Consent Notices](#_ome4t9ifhwlb) 2

[2.2.3. Topic Modeling in Privacy Policies](#_tg6437cnpy9w) 2

[2.3. Research Gaps](#_f2te2q8ebtnj) 2

[**3. Design**](#_f2oxkbrf1qs3) **2**

[3.1. Design overview](#_1cedejbsl0i7) 2

[3.2. Website Gathering](#_mif4z4lzqeg6) 2

[3.3. Policy Extraction](#_xarasonx3lwl) 2

[3.4. Policy Sanitisation](#_b77glnme3dhu) 2

[3.5. Policy Analysis](#_qqkdp9rqngod) 3

[**4. Implementation**](#_3ada6paucnsx) **3**

[4.1. Website Gathering](#_pj0nnfmjnsoa) 3

[4.2. Web Scraper](#_xu0alq6fffe) 3

[4.3. Policy Sanitisation](#_5uf0nq6i2mz) 3

[4.4. Analysis](#_p79pi0fgev0z) 3

[**5. Evaluation and Critical Appraisal**](#_jmrx9a2wji6y) **3**

[5.1. Evaluation of Approaches](#_kiof2372w88n) 3

[5.2. Achievements & Limitations](#_6fgq5fw78rde) 3

[**6. Conclusion**](#_9nb8hiyce7az) **3**

[6.1. Project Summary](#_8lt656kfrfek) 3

[6.2. Suggestions for Future Work](#_fowt0ww6un6z) 3

[**References**](#_4b857h9rxqo) **3**

[**Appendix**](#_ohpywuthmmw8) **4**

[Appendix A: Ethics Self-Assessment Form](#_g203acwhpuhz) 4

# 1. Introduction

## 1.1. Project Context

## 1.2. Problem Description

## 1.3. Objectives

## 1.4. Project Structure

# 2. Context Survey

## 2.1. Legal Background

### 2.1.1. Privacy and the GDPR

### 2.1.2. Consent

### 2.1.3. Legitimate Interests

## 2.2. Computer Science Background

### 2.2.1. Standardisation of Privacy Policies

### 2.2.2. Technical Analysis of Cookie Consent Notices

### 2.2.3. Topic Modeling in Privacy Policies

## 2.3. Research Gaps

# 3. Design

## 3.1. Design overview

## 3.2. Website Gathering

## 3.3. Policy Extraction

## 3.4. Policy Sanitisation

## 3.5. Policy Analysis

# 4. Implementation

## 4.1. Website Gathering

## 4.2. Web Scraper

## 4.3. Policy Sanitisation

## 4.4. Analysis

# 5. Evaluation and Critical Appraisal

## 5.1. Evaluation of Approaches

## 5.2. Achievements & Limitations

# 6. Conclusion

## 6.1. Project Summary

## 6.2. Suggestions for Future Work

# References

# Appendix

## Appendix A: Ethics Self-Assessment Form