

Configuration Manual

MSc Research Project
MSc in Cyber Security

Meril Antony
x23103418

School of Computing
National College of Ireland

Supervisor: Kamil Mahajan

National College of Ireland
MSc Project Submission Sheet
School of Computing



Student Name: Meril Antony

Student ID: X23103418

Programme: MSc in Cyber Security

Year: 2024

Module: MSc Research Project

Supervisor: Kamil Mahajan

Submission

Due Date: 29/01/2025

Project Title: Data Privacy in Buy Now Pay Later (BNPL) - A Comprehensive PIA Framework

Word Count: 305

Page Count 2

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature: Meril Antony

Date: 29/01/2025

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST

Attach a completed copy of this sheet to each project (including multiple copies)	<input type="checkbox"/>
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	<input type="checkbox"/>
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	<input type="checkbox"/>

Assignments that are submitted to the Programme Coordinator Office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Configuration Manual

Meril Antony
x23103418

1 Data Mining Tool for privacy policy analysis

1.1 Dependencies:

1. Python - Version 3.8 or higher
2. SpaCy - en_core_web_sm language model.
3. Hugging Face Transformers - facebook/bart-large-cnn
4. Scikit-learn – libraries TfidfVectorizer and Cosine Similarity
5. Matplotlib
6. Google collab is recommended.

1.2 Settings and Parameters:

Input: text only file

Summarization Configuration:

Model: facebook/bart-large-cnn

Input Text Limit: 1024 tokens per chunk.

Text Classification Configuration:

- Predefined Categories:
 - Data Collection
 - Data Sharing
 - Data Retention
 - User Rights
 - Security Measures
 - General Information
- TF-IDF Vectorization:
 - Stop Words Removal: Enabled.
 - Normalization: Tokenized and lowercased text.
 - Output: Sentence vectors and category vectors.
- Similarity Metric:
 - Cosine similarity to map sentences to the most relevant category.

2 Guidelines for using the framework

This section provides instructions on how to effectively use the PIA template for assessing and managing privacy risks within BNPL systems.

1. Preparation

Gather all relevant system and data processing details including the types of data collected, how it is processed, where it is stored, and who has access to it. Identify all stakeholders involved in data collection, processing, and management. Stakeholders may include internal teams, third-party partners, and regulatory bodies.

2. Completing the Template

Carefully fill out each section of the template with accurate data. Ensure every field is comprehensively addressed to identify the landscape of privacy. Contact your organization's Data Protection Officer or compliance team for assistance in completing the template.

3. Risk mitigation

Reference specific ISO/IEC 27701 controls for aligned mitigation strategies.

4. Evaluation and Review:

Establish a review cycle to ensure the PIA is conducted for any system changes or new regulations.

For any missing data processes, contact the technical team for assistance in completing the template. Collaborate with legal, technical and operational team to fill out this template.

.