Research Document

GDPR Compliance Audit Tool

Cormac Casey ~ C00283808

BCs (Hons) Cybercrime & IT Security

SETU Carlow

1. Table of Contents

Contents

1. Table of Contents 2

2. Introduction 4

2.1 Motivation 4

2.2 Problem Statement 5

2.3 Objectives 6

2.4 Project Scope 6

3 Project deliverables & System Overview 7

3.1 Core Deliverables 7

3.2 Secondary Deliverables 7

3.3 Project Description 9

3.4 Target Users 9

4 Key Features 11

4.1 Overview of Tools Capabilities 11

4.2 Report Generation and Visualisation 12

5. System Success Criteria 13

5.1 Functionality 13

5.2 Usability 13

5.3 Reliability 13

5.4 Performance 14

5.5 Supportability 14

6 Research Methodology 15

6.1 Literature Review 15

6.1.1 Overview of GDPR and Organisational Responsibilities 15

6.1.2 GDPR Challenges for SMEs 15

6.1.3 Metric based Scoring Models 16

6.1.4 Reporting and Visualisation in Compliance Tools 17

6.1.5 Summary of Literature Findings 17

6.2 Overview of GDPR Principles 18

6.3 Existing GDPR Compliance Tools 20

7 Ethics and Legality 21

7.1 Ethical Considerations 21

7.2 Terms and Conditions 23

7.3 Data Protection 23

8 Technologies used. 25

8.1 Programming Languages and Frameworks 25

8.2 Secure Data Handover 25

9 Project Infrastructure 26

9.1 Data Fields Overview 26

9.2 Module design 28

9.2.1 Data Input and Processing 28

9.2.2 Compliance Checking Engine 29

9.2.3 Reporting and Visualization 29

9.3 Implementation Workflow 29

9.4 System Evaluation and Development 30

9.4.1 Version 1 30

10 Constraints & Dependencies 31

11. Project Plan 32

12. References 32

2. Introduction

2.1 Motivation

My motivation for this project is largely influenced by the lack of support for SME's (Small & Medium sized enterprises) with regards to ensuring GDPR Compliance. In todays tech dominated world GDPR plays a huge role for the majority of SME's as they tend to deal with vast amounts of user data. SME's need to be compliant with GDPR or they could be subject to large fines, security risks & reputational damage. However Enterprises on the smaller side don't usually have the necessary resources to conduct efficient GDPR Compliance audits and in-house audits. This leads to many SME's struggling to meet GDPR Compliance regulations. My research tells me the existing

tools on the market tend to share some common traits. These traits include expensive solutions, high complexity, & high time consumption. The solutions for GDPR Compliance audits currently on the market are not on par with the needs for the Smaller Enterprises. This motivated me to create an affordable, easy to use & efficient GDPR Compliance Audit Tool that isn't too complex for SME's.

"GDPR compliance involves a great number of policies and procedures, often at a big expense, leaving many small and medium-sized enterprises (SMEs) finding it difficult to comply. Data security methods such as encryption can be confusing, especially for small businesses without technical expertise." *(1)*

2.2 Problem Statement

In todays digital landscape, organizations collect and process very large amounts of personal data through websites and applications. With the introduction of General Data Protection Regulation (GDPR), organizations are legally required to ensure transparency, consent and secure handling of user data.

Achieving + Maintaining compliance with these regulations presents challenges for organizations.

- Complexity of Regulation: GDPR consists of numerous articles and requirements that can be difficult to interpret for smaller organizations lacking legal expertise,
- Manual Auditing effort + Cost: Manual GDPR auditing is a time consuming process which is prone to human error and also not cost efficient.

- Ethical + Legal Boundaries: Automated scanners can unintentionally cross ethical boundaries if used on unauthorized websites, this highlights the need for an ethical and legal GDPR compliance tool for organizations.

Therefore, there is a need for an ethical and automated system capable of auditing authorized data against key GDPR compliance indicators and generate a measurable compliance scoring metric.

2.3 Objectives

Create an easy to use, cheap, effective & secure GDPR Compliance Audit Tool that will be effective and useful for SME's. A main objective for me with regards to the finished project is to output an easy to interpret metric containing a compliance score and best practice recommendations to improve current GDPR Compliance for the said SME. The output and metric should be easy to interpret as many SME's don't have the skilled/educated personal with regards to GDPR and cyber legislation.

2.4 Project Scope

- Generate synthetic GDPR Data for testing
- Create a data ingestion module to sort data into correct format
- Generate a compliance checker engine
- Create a custom scoring metric
- Create a report generator with scores and recommendations

3 Project deliverables & System Overview

3.1 Core Deliverables

- A complete GDPR Compliance Audit Tool that offers SMEs an option to conduct in-house GDPR Compliance Audit Checks.
- A complete Metric Scoring Design taking many GDPR Variables into consideration
- A Report Generator that will display reasons for the scoring and recommendations for future improvements in GDPR Compliance in said SME.

3.2 Secondary Deliverables

User Interface

The main secondary deliverable I would like to create a User Interface for the GDPR Compliance Audit Tool, this will improve usability of the tool and make the tool more user-friendly. The UI would include buttons, forms, visual feedback.

Report Graphics

Extended reporting features are another Secondary Deliverable I would like to implement if time shall allow. This would be the addition of charts, graphs, or heatmaps for individual fields or overall GDPR compliance. Inclusion of more export options like (PDF Excel, HTML) would be a worthwhile addition also.

AI-Assisted Narrative Generation

An additional enhancement considered for this project is the integration of an AI-driven narrative module. This component would generate human-readable explanations of compliance results, translating technical scoring outputs into accessible summaries for non-technical users.

The purpose of this feature is to improve usability and accessibility, particularly for SMEs that may lack cybersecurity expertise. By converting structured audit results into contextual explanations, the system would support better decision-making and reduce the cognitive effort required to interpret compliance findings.

While not essential to the core system, this feature represents a valuable extension that aligns with emerging trends in AI-assisted cybersecurity tooling.

Host Telemetry Integration

Another potential extension involves the integration of host-level telemetry data to enhance the accuracy of compliance assessments. This would allow the system to incorporate real system indicators such as open ports, active services, or patching status into the compliance evaluation process.

Currently, the tool relies on structured synthetic inputs. However, integrating telemetry data would shift the system towards a more realistic and dynamic auditing approach, improving the validity of results and reducing reliance on manually supplied data.

Due to ethical considerations and project scope limitations, this functionality is not implemented in the current version but remains a strong candidate for future development.

3.3 Project Description

This project focuses on the design and development of a GDPR Compliance Audit Tool aimed at supporting small and medium-sized enterprises (SMEs) in conducting internal data protection assessments. The tool enables organizations to evaluate key GDPR-related criteria using a structured set of metrics, producing a quantified compliance score alongside clear explanations and improvement recommendations.The system comprises a fully implemented scoring engine that analyses multiple GDPR variables across areas such as security measures, transparency, user rights, and internal processes. Results are automatically compiled into a professionally structured PDF report that outlines scoring rationales, highlights areas of strength and weakness, and offers actionable guidance for enhancing data protection practices. Collectively, these components aim to provide SMEs with an accessible, practical, and repeatable method for understanding and improving their GDPR compliance posture.

3.4 Target Users

- Small & Medium Enterprises (SMEs)
    - Organizations lacking the internal expertise or staffing to conduct full GDPR compliance audits.

  - Businesses without dedicated Data Protection Officers (DPOs) or compliance teams.
  - Companies seeking an affordable alternative to outsourcing full external audits.

- Start-Up Businesses Handling Client Data
  - Early-stage companies collecting personal data but without established compliance processes.
  - Start-ups that need a straightforward tool to identify risks before scaling operations.
  - New businesses aiming to demonstrate data protection maturity to clients and investors.

- Non-Technical Teams Needing Accessible Compliance Tools
  - Staff without cybersecurity knowledge who require more user-friendly audit support.
  - HR, finance, and operations teams who routinely handle sensitive data.

- Businesses Preparing for Third-Party or Regulatory Audits
  - Organizations wanting to assess their compliance posture before an official audit.
  - Companies seeking documentation and automated reporting to demonstrate accountability.

4 Key Features

4.1 Overview of Tools Capabilities

The GDPR Compliance Audit Tool provides a structured and efficient way for organizations (particularly SMEs) to evaluate their current level of GDPR adherence using synthetic but realistic data inputs. The tool is capable of ingesting structured datasets, validating each field against expected formats, and applying a rule-based compliance scoring model that reflects key GDPR requirements across security, transparency, user rights, and internal governance. The system analyses a range of GDPR-relevant variables such as HTTPS usage, password storage practices, cookie consent mechanisms, privacy policy clarity, DSAR procedures, retention policies, and breach-response readiness. Each field influences the overall compliance score through a weighted metric design, enabling the tool to highlight strong areas and identify weaknesses that require attention. Once analysis is complete, the tool automatically generates a detailed PDF compliance report. This report includes the organization's overall GDPR score, individual control scores, explanations for each scoring outcome, and tailored recommendations to improve compliance posture.

Overall, the tool provides SMEs with an accessible, affordable, and non-technical method to conduct internal GDPR assessments, helping them understand their risk exposure and take informed steps toward better data protection practices.

4.2 Report Generation and Visualisation

A key component of this project is the automated generation of a structured, user-friendly GDPR compliance report. The tool outputs a PDF document that contains all assessment results, including the compliance metrics score, individual control scores, and a set of tailored recommendations. This report is designed to provide users with a clear, actionable understanding of their current GDPR posture.

The report generation module compiles the results produced by the scoring engine and transforms them into an easily readable format.

In addition to raw scores, the report also provides context-aware recommendations based on the organisation's specific deficiencies. For example, if the system detects weak password storage practices or the absence of a DSAR process, the generated report includes targeted guidance on how to meet these GDPR expectations. This ensures the output is not only informative but also practically valuable to an SME.

Overall, this combination of automated report generation and data visualisation enhances usability, supports decision-making, and strengthens the real-world applicability of the GDPR compliance auditing tool.

5. System Success Criteria

5.1 Functionality

The system must successfully perform all core tasks defined for the GDPR Compliance Audit Tool. This includes accepting synthetic GDPR-related datasets, validating each field against expected data types, applying rule-based scoring logic, and generating a final compliance score. The tool must also highlight weak areas, provide actionable recommendations, and produce a structured PDF report. Full functionality is achieved when each stage—input, processing, scoring, and reporting executes accurately and consistently without manual intervention.

5.2 Usability

Although the system does not include a GUI, usability remains a key criterion. The tool should be straightforward to operate via a command-line or script-driven interface, with clear instructions on how to supply input data. Output reports must be easy to interpret, using clear headings, structured sections, and consistent formatting. Users should not require advanced technical knowledge to understand the compliance results. Usability is considered successful if a user can complete an audit from input to PDF report with minimal guidance.

5.3 Reliability

Reliability refers to the system's ability to run stably over repeated tests and varied synthetic datasets. The tool must handle missing fields, invalid formats, or incomplete data without crashing, instead providing meaningful error messages or default scoring

behavior. The scoring engine should produce consistent results for identical inputs, ensuring that the assessment logic is deterministic and trustworthy. System reliability is demonstrated through repeated test cycles and validation using diverse synthetic data.

5.4 Performance

The system should process data efficiently, even when presented with larger synthetic datasets. Performance is evaluated based on processing time, resource usage, and responsiveness during report generation. While high-volume scalability is not a core requirement for this academic project, the tool should still complete an audit quickly under normal conditions. Performance success is achieved when the system produces results within a reasonable timeframe without causing excessive CPU or memory consumption.

5.5 Supportability

Supportability focuses on the ease of maintaining, modifying, and extending the system. The codebase must be modular, well-commented, and logically organised to support future enhancements such as adding new GDPR fields, adjusting scoring weights, or incorporating more visualisations. The system should also maintain compatibility with key external libraries, with clear documentation of version dependencies. Supportability is met when future developers or reviewers can understand the architecture and implement changes with minimal effort.

6 Research Methodology

6.1 Literature Review

6.1.1 Overview of GDPR and Organisational Responsibilities

GDPR introduces procedural obligations on organisations who are involved in the processing of personal data. The responsibilities of said organisations are to keep individuals personal data accurate and up to date whilst deleting any data that is no longer deemed necessary. GDPR obligates organisations to respect individual rights by informing them on how and why their data is processed. Organisations have the responsibility to make sure that personal data is handled in a secure way and maintain a record of processing operations. *(Z)*

6.1.2 GDPR Challenges for SMEs

With the introduction of GDPR a change of culture emerged as an initial challenge for many SMEs. SMEs that previously existed outside of any meaningful compliance regime leaving out finance and tax now faced a new level of accountability to national authorities regarding GDPR. For many SMEs taking on this challenge which is GDPR a main struggle was to find the access to practical expertise. With the implementation of GDPR, skilled individuals who have the practical expertise have now entered a competitive market causing their wages and prices to soar and become too high for SMEs to acquire their services. This is the main financial challenge alongside subscription costs to audit/compliance services.

"Even where SMEs decided to try to hire in these skills, there was, and remains, a significant shortage of trained and experienced individuals. For individuals who fit that profile, the demand for their skills has created a competitive market-place where salaries have soared". (8)

In addition to cost and staffing pressures, SMEs often struggle because GDPR is principle-based rather than fully prescriptive. This means organisations must interpret broad legal requirements and translate them into practical technical and organisational controls, which can be difficult without specialist support. The European Data Protection Board has recognised this issue by publishing dedicated SME guidance intended to make GDPR obligations more accessible and easier to apply in practice. *(9)*

6.1.3 Metric based Scoring Models

Metric-based scoring models provide a structured and quantifiable approach to evaluating compliance by assigning weighted values to specific controls and organisational practices. Rather than relying solely on qualitative judgement, these models translate regulatory requirements into measurable indicators, enabling consistent and repeatable assessments.

In cybersecurity and compliance domains, established frameworks such as NIST and ISO/IEC 27001 utilise control-based scoring systems to assess organisational security posture. These approaches allow organisations to identify weaknesses, prioritise remediation efforts, and track improvements over time.

Applying a similar methodology to GDPR compliance enables complex legal requirements to be simplified into actionable metrics. By assigning scores to areas such as security measures, transparency, user rights, and internal controls, organisations can gain a clear understanding of their compliance position.

For SMEs in particular, this approach reduces the need for deep legal or technical expertise by presenting results in an accessible, data-driven format. It also supports continuous monitoring, allowing organisations to benchmark progress and improve their compliance posture in a structured manner. *(6)*

6.1.4 Reporting and Visualisation in Compliance Tools

Modern compliance tools rely on clear reporting and visualisation features to help organisations understand their compliance status. Dashboards, summaries, and simple charts allow non-experts to quickly interpret audit results, risks, and outstanding tasks. Visual elements such as heatmaps and progress indicators highlight areas needing attention, making compliance management easier and more transparent for SMEs with limited resources.

6.1.5 Summary of Literature Findings

Research Area	Key Insight	Relevance to Project

GDPR Responsibilities	Organisations must ensure secure and transparent data processing	Supports need for structured compliance assessment
SME Challenges	SMEs lack expertise, budget, and resources	Justifies focus on accessibility and simplicity
Metric-Based Models	Weighted scoring enables measurable and repeatable evaluation	Supports scoring engine design
Reporting & Visualisation	Clear outputs improve understanding for non-technical users	Supports report generation approach

6.2 Overview of GDPR Principles

The General Data Protection Regulation (GDPR) is built around seven core principles that guide how personal data must be collected, processed, stored, and protected. These principles form the foundation of compliant data handling across all sectors and ensure that organisations manage personal information responsibly and transparently. The principles are:

- Lawfulness, Fairness, and Transparency

  Organisations must process personal data legally, treat individuals' information fairly, and be open about how and why data is used. Individuals should always understand what is happening with their data.

- Purpose Limitation

Data must be collected for specific, explicit, and legitimate purposes. It cannot be reused for unrelated activities without further consent or a valid legal basis.

- Data Minimisation

Only the minimum amount of personal data necessary for the intended purpose should be collected and processed. Excessive or irrelevant data processing is prohibited.

- Accuracy

Personal data must be kept accurate and up to date. Organisations are responsible for correcting or removing inaccurate information without delay.

- Storage Limitation

Data should only be kept for as long as necessary. Once the data is no longer required for its original purpose, it must be securely deleted or anonymised.

- Integrity and Confidentiality (Security)

Organisations must protect personal data through appropriate security measures to prevent unauthorised access, loss, or damage. This includes both technical and organisational controls.

- Accountability

Organisations are required to actively demonstrate compliance with GDPR. This includes maintaining documentation, implementing policies, and being able to show evidence of responsible data-handling practices. *(4),(5)*

6.3 Existing GDPR Compliance Tools

A widely used example of a GDPR compliance platform is GDPR-Software.com *(3)*, which offers organisations a structured way to manage their GDPR obligations. The tool provides features such as guided compliance questionnaires, policy and document templates, data-processing and breach registers, and audit-ready reporting. These capabilities help businesses organise and document their GDPR responsibilities in a clear and consistent manner.

Platforms like GDPR-Software.com demonstrate the strong demand for accessible GDPR solutions that simplify compliance tasks, especially for organisations that may not have in-house expertise. They offer a comprehensive, centralised approach suitable for a broad range of industries and business sizes.

However, GDPR-Software.com operates on a paid subscription model, which may not be ideal for smaller organisations, start-ups, or teams with limited budgets who still need to assess their compliance position. This creates space in the market for tools that provide a more affordable, focused, or streamlined approach to GDPR auditing.

This context supports the development of alternative GDPR audit tools such as the one proposed in my project, designed to meet the needs of users who require a simpler, more cost-effective option while still gaining meaningful insight into their compliance posture. The inclusion of a custom GDPR compliance score and practical recommendations in the final report supports organisations in understanding their current compliance posture and prioritising areas for improvement.

7 Ethics and Legality

7.1 Ethical Considerations

- Crawling (Proposed but not implemented due to ethical concerns)

  Web crawling was initially considered as a potential feature for automatically detecting compliance indicators such as HTTPS usage, cookie banners, or the existence of privacy policy pages. However, this functionality was deliberately excluded from the final system due to ethical implications. Even lightweight crawling may unintentionally burden websites, violate terms of service, or raise concerns regarding unauthorised interaction with live systems. Additionally, crawling risks the accidental collection of personal or sensitive data, which would contradict the ethical principles and legal obligations of GDPR itself.
  By not implementing this feature, the project maintains a strong ethical posture and avoids the complexities of automated data collection from external systems.

- Data Handling

  Throughout development and testing, only synthetic, anonymised, or entirely fictional data is used. The tool is designed to avoid ingesting or processing any real personal data.

  Data minimisation principles were applied so that only fields strictly required for generating compliance scores are collected and processed. No unnecessary attributes, identifiers, or behavioral data are stored.

  All sample organisations, privacy policies, DSAR descriptions, and retention statements are fabricated specifically for this project. This approach ensures full compliance with ethical research standards and prevents any risk of exposing real individuals or organisations to harm.

- Contracts and Legal Documents

  The system analyses simplified, synthetic representations of legal and regulatory documents such as privacy policies, data handling statements, and DSAR procedures.

  It does not interpret legally binding contracts, does not assess regulatory compliance in a professional capacity, and does not store or review real corporate paperwork.

All outputs and recommendations generated by the tool are strictly for research, demonstration, and educational purposes. They are not legal advice and must not be relied upon for real compliance audits.

7.2 Terms and Conditions

As part of demonstrating GDPR-aligned transparency and user awareness, a simplified Terms and Conditions (T&Cs) document is included within the project. This serves as an illustrative example of how organisations communicate key information regarding data usage, user responsibilities, and service conditions. While the system does not perform legal analysis of such documents, the inclusion of T&Cs reflects the importance of clearly informing users about how their data may be processed within a given service context. This supports the GDPR principle of lawfulness, fairness, and transparency, ensuring that individuals are provided with accessible and understandable information. The use of synthetic and non-binding T&Cs ensures that the project remains ethically compliant while still demonstrating the role such documentation plays in an organisation's overall data protection framework.

7.3 Data Protection

Data protection is a core principle underpinning the design and operation of this GDPR compliance auditing tool. Although the system uses entirely synthetic data for testing and simulation, all processes are structured to demonstrate the application of GDPR-aligned data protection principles in practice.

Synthetic datasets represent fictional companies, users, and GDPR-relevant fields such as HTTPS usage, cookie consent, password storage methods, DSAR processes, retention policies, etc. These datasets allow the system to generate realistic compliance scores and reports without processing any real personal information, eliminating ethical or legal risk.

Internally, the tool handles this synthetic data responsibly. It demonstrates key security practices, including temporary storage, controlled access, and structured organisation of data for processing. The data is stored in local files (e.g., JSON or CSV) and can be encrypted if required, ensuring that even in a simulated environment, sensitive information is protected from unauthorized access. After processing, datasets are deleted, mimicking responsible retention practices.

The system also demonstrates principles of data minimisation and purpose limitation, as only fields necessary for compliance scoring are included and used. No extraneous or irrelevant information is collected or processed. In this way, the project models good data governance practices while remaining fully ethical and compliant with research standards.

In summary, even with synthetic data, this project effectively demonstrates the secure handling, processing, and storage of sensitive information, showcasing the principles of data protection, accountability, and privacy that underpin GDPR compliance.

8 Technologies used.

8.1 Programming Languages and Frameworks

Python 3.14.0 – Primary programming language

Visual Studio Code – Main development environment

Jupyter Notebook – Early prototyping, testing logic, and experimenting with scoring methods or data structures

Github – Version control and repository hosting for tracking progress and safe storing of the project

PDF Libraries (Reportlab) – For generating the structured PDF Output

Matplotlib – Python Library for simple plots and visualisations I may add into the final report

Json / csv modules – for supporting structured synthetic data sets

Pandas - for handling input data and performing basic validation

Datetime - to timestamp audit logs and reports

The chosen tools and technologies remain subject to change as project requirements evolve or dependencies update.

8.2 Secure Data Handover

To ensure that any exchange of client data associated with my GDPR Compliance Audit Tool meets GDPR requirements, a secure, encrypted, and fully auditable handover process is essential. Because the transfer stage carries the highest risk of data exposure, a managed, standards-compliant solution is required.

"The external transfer of sensitive data is a core operational business process of IT organizations. Data in transit is data at risk of interception, unauthorized access or mishandling" *(2)*

I propose using MOVEit Transfer as the recommended option for secure data handover. MOVEit provides strong end-to-end protection through FIPS 140-2 validated AES-256 encryption for data both in transit and at rest. It also maintains a tamper-evident audit trail of all transfer activity, supporting GDPR obligations around accountability, traceability, and access control. Additional features such as MFA support, policy enforcement, and secure protocol handling further strengthen its suitability for regulated environments.

While I will not be demonstrating MOVEit within this project, it stands out as the most robust and compliant solution for organisations requiring secure, GDPR-aligned data transfer processes. Its capabilities ensure that confidentiality, integrity, and regulatory compliance are preserved throughout the handover of sensitive client information.

9 Project Infrastructure

9.1 Data Fields Overview

The GDPR compliance tool utilises a structured set of data fields to simulate an organisation’s security, transparency, and internal control measures. Each field is

carefully designed to reflect key aspects of GDPR compliance, allowing the scoring engine to assess and generate meaningful metrics.

1. Basic Security Measures

- https_enabled (Boolean): Indicates whether the organisation uses HTTPS for secure communications.
- password_storage_method (Enum): Specifies how passwords are stored, e.g., plaintext, SHA256, bcrypt.
- regular_security_testing (Boolean): Reflects whether routine security checks are performed.

2. Transparency & User Rights

- cookie_consent_banner (Boolean): Shows the presence of a user consent mechanism.
- privacy_policy_present (Boolean): Indicates whether a privacy policy exists.
- privacy_policy_clarity (Enum): Assesses how clearly the policy communicates data practices (clear, unclear, missing).
- dsar_response_time_days (Integer): Simulated number of days to respond to Data Subject Access Requests (1–99).
- dsar_process (Enum): Describes the DSAR procedure (documented, partial, missing).

3. Internal Controls

- data_breach_process_present (Boolean): Indicates whether a breach response procedure exists.
- data_retention_policy (String): Describes the retention policy.
- retention_period_days (Integer): Length of data storage in days (1–299).
- has_dpo (Boolean): Whether the organisation has a Data Protection Officer.

Each of these fields serves as an input for the scoring engine, allowing the tool to generate compliance scores, visualisations, and recommendations. The combination of boolean, enumerated, integer, and string fields provides a realistic simulation of a GDPR audit while using entirely synthetic data.

Subject to change during the development of this tool.

9.2 Module design

9.2.1 Data Input and Processing

This module is responsible for accepting, validating, and preparing the GDPR-related synthetic data used by the system. It reads predefined fields such as security measures, transparency indicators, and internal control values, and ensures that each field conforms to the expected type (Boolean, Enum, integer, or string). The module also handles basic error checking, missing values, and schema mismatches to maintain consistent input quality. Once validated, the data is normalised into a structured format that the compliance engine can interpret.

9.2.2 Compliance Checking Engine

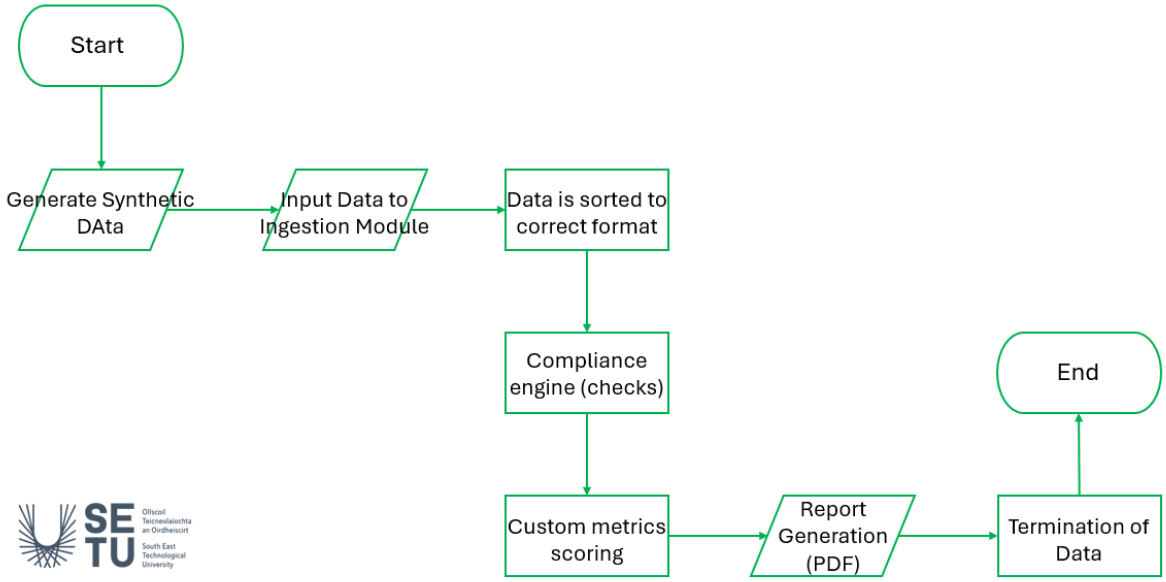
The Compliance Checking Engine is the core logic module of the system. It applies a rule-based approach to evaluate each data field against GDPR-aligned criteria. Every field contributes a weighted score, with Booleans converted to binary values and Enum fields mapped to predefined scoring categories. These individual scores are combined to generate an overall compliance rating. The engine also flags weak areas such as unclear privacy policies or insecure password storage and generates tailored recommendations where deficiencies are detected.

9.2.3 Reporting and Visualization

This module generates the final audit output, including a structured PDF report containing the compliance score, key findings, and improvement recommendations. Visualisation elements may include tables, metric summaries, and simple charts to depict scoring distribution. The reporting module formats the results for readability and ensures that the output remains consistent, well-structured, and suitable for both academic and practical evaluation contexts. Audit logs may also be stored locally to support traceability of assessments.

9.3 Implementation Workflow

Version 1 Workflow



9.4 System Evaluation and Development

9.4.1 Version 1

The initial version (V1) of the GDPR Compliance Audit Tool focused on establishing a functional baseline architecture. The system implemented:

- Basic synthetic data ingestion
- A simple rule-based compliance engine
- Binary or limited scoring logic
- Basic PDF report generation

While V1 successfully demonstrated end-to-end functionality (data → scoring → report), several limitations were identified:

- Over-reliance on binary scoring (true/false)
- Limited maturity modelling
- Minimal dependency handling between fields
- Simplistic report visualisation

- Limited traceability of scoring decisions

10 Constraints & Dependencies

The development of the GDPR Compliance Audit Tool is influenced by several constraints and dependencies that shaped the final design and scope of the project. Acknowledging these constraints and dependencies allows me to plan efficiently with regards to structuring the tool and deciding the tools capabilities.

Lack of GUI :

Although a graphical interface would improve usability, time limitations meant that a GUI may not be designed or implemented. The project therefore will operate through a script-based workflow.

Access to Sample and Test Data :

The tool depends on synthetic data generated specifically for development and testing. This reliance means the accuracy and usefulness of results are tied directly to the quality and structure of the fabricated datasets.

Development and Testing Time :

A tight schedule significantly constrains feature complexity, refinement, and extensive testing. Only essential components can be implemented reliably within the available time.

Library Versioning & Updates :

The tool will rely on several external Python libraries. Updates or version mismatches may introduce compatibility issues and require additional troubleshooting, placing constraints on development stability.

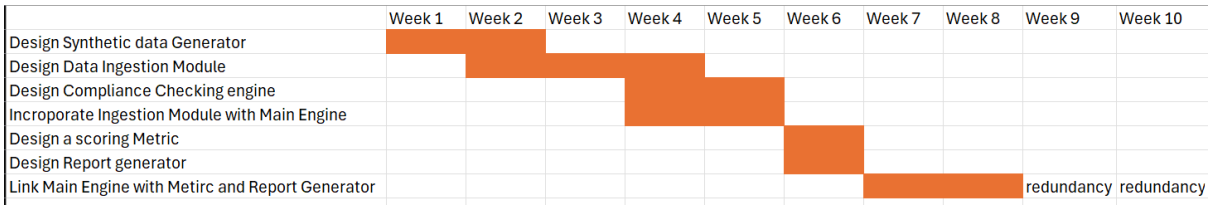
Storage Space for Audit Logs :

Generated reports and audit logs require local storage. Given limited available space, logging is kept lightweight to avoid unnecessary storage consumption.

Designing Effective Metrics :

The scoring and evaluation system depends on well-designed metrics. Developing meaningful, balanced, and consistent scoring criteria is essential for producing valid compliance ratings, making it a core dependency for tool accuracy.

11. Project Plan



12. References

(1) Staff, C. (2024). *What Is GDPR?* [online] Coursera. Available at: https://www.coursera.org/articles/what-is-gdpr

(2) Progress.com. (2023). *GDPR-Compliant Secure File Transfer - Progress MOVEit*. [online] Available at: https://www.progress.com/moveit/gdpr-compliance

(3) Gdpr-software.com. (2025). *GDPR Software*. [online] Available at: https://www.gdpr-software.com/

(4) Özkan, I. (2022). *Data Protection Principles: The 7 Principles Of GDPR Explained*. [online] www.cyberpilot.io. Available at: https://www.cyberpilot.io/cyberpilot-blog/data-protection-principles-the-7-principles-of-gdpr-explained/.

(5) Www.ucd.ie. (2024). *Data Protection Principles & Applications - UCD GDPR*. [online] Available at: https://www.ucd.ie/gdpr/dataprotectionoverview/dataprotectionprinciplesapplications/.

(6) Brodin, M. (2019). A Framework for GDPR Compliance for Small- and Medium-Sized Enterprises. *European Journal for Security Research*, 4. doi: https://doi.org/10.1007/s41125-019-00042-z.

(7) www.edpb.europa.eu. (n.d.). *What are my responsibilities under the GDPR? | European Data Protection Board*. [online] Available at: https://www.edpb.europa.eu/sme-data-protection-guide/faq-frequently-asked-questions/answer/what-are-my-responsibilities-under_en.

(8) Corinna-Tri (2020). *What are the challenges that SMEs are facing in complying with the GDPR? A view from the field*. [online] Trilateral Research. Available at: https://trilateralresearch.com/data-protection/challenges-facing-smes-in-complying-with-the-gdpr-a-view-from-the-field.

(9) Guidelines for SMEs on the Security of Personal Data Processing. (2016).