

Compliance Audit Form

Organization Name: CyberNations//Protexxa

Audit Date: [23/04/2025]

Auditor Name: Tamara Clarke, Angela Walrond, Latonia Belle & Staicia Yarde

Department/Unit Audited: Data Cleaning

Contact Person & Details: Sophia Pitt-Browne & Gilbert Millar

Purpose of Audit: _____

Regulatory Frameworks Applicable (Check all that apply):

☒ Barbados Data Protection Act (BDPA)

☐ PIPEDA

☐ AIDA

☒ GDPR

☒ Other: NIST, EU AI Act

Section 1 (PART 2): Data Cleaning & Processing (To be answered by Data Cleaning ONLY)

Data Integrity & Quality Control

Comments

1. Is there a structured data cleaning process that is version controlled or documented for reproducibility?

☒ Yes/☐ No

There is a data cleaning checklist

2.Are scripts or automated tools used to check data quality (e.g., for duplicates, missing values, or formatting issues)?

☒ Yes/☐ No

Within scripts have data quality checks, either code blocks or external scripts

3.How is missing data handled? Are nulls, blanks, and duplicates properly identified and managed?

They are handled by removing them and placing them in garbage files if it is a duplicate or missing value

4.After data cleaning, is the output reviewed manually or validated using tools to ensure accuracy before use?

☒ Yes/☐ No

They are reviewed both manually and through analysis tool

5.How is potentially compromised or illegal data handled during cleaning (e.g., flagged, quarantined, discarded)?

All data utilized by the team is compromised so they are just cleaned

6.Are cleaning rules reassessed regularly based on feedback, new patterns, or issues found?

☒ Yes/☐ No

7.Are there performance bottlenecks or known limitations in the current cleaning workflow?

☒ Yes/☐ No

Hardware limitations such as RAM and processing speed

Total Score for Section 1: [40/40] **Add 5 to the final score**

Section 2: Data Storage, Security & Privacy

Security & Risk Awareness

1.If your team handles compromised or leaked data, how is risk managed during acquisition and storage. (3 pts max)

☒ Yes/☐ No

Comments

1) The team has a process in place for handling compromised data

2.Are regular vulnerability assessments conducted? (3pts max)

☐ Yes/☒ No

2) There is currently no regular vulnerability assessments conducted.

3.Is there a centralized repository for finalized scripts and collected information or datasets? How is version control managed? (3 pts max)

☒ Yes/☒ No

3) Finalized scripts are maintained in a GitHub repository. However, there is no centralized repository for collected datasets, which could result in version control and data fragmentation issues.

4. Are backup systems in place for scripts, scraped information or datasets? Are they encrypted and stored offsite or in a secure cloud location? (4 pts max)

☒ Yes/☐ No

4) Backup systems are in place and confirmed to be secure

5. Is there a process for revoking access when a team member leaves or changes roles? (2 pts max)

☐ Yes/☒ No

5) There is no formal process for revoking access when team members leave or change roles

6. Are there clear policies for retention and secure disposal? (3 pts max)

☒ Yes/☒ No

6) There are no clearly defined policies for data retention or secure disposal

7. If datasets are stored across different platforms (e.g., Teams, GitHub, local machines), how is cross platform consistency ensured, so people are not working on outdated information? (3 pts max)

☒ Yes/☐ No

7) The team uses naming and date conventions to maintain consistency across platforms.

8. Are you aware whether members have standard protections in place? Multi-Factor Authentication (MFA), etc.? (4 pts max)

☒ Yes/☐ No

8) Standard security measures such as multi-factor authentication and antivirus protection are in place due to company policies.

Total Score for Section 2: [17/25]

Section 3: Compliance & Accountability Measures

Breach & Incident Reporting

Comments

1.Are formal data breach reporting processes in place?

☐ Yes/☒ No

Not persons in AI cohort

2.Any compliance violations or incidents before this audit?

☐ Yes/☒ No

3.Is there a documented incident response plan in case of a data breach or unauthorized access? If yes, who is responsible for executing it?

☐ Yes/☒ No

4.Has your team encountered any challenges in storing or organizing data securely (e.g., decentralized files, missing logs, lost local data)?

☒ Yes/☐ No

Decentralized files. Due to the lack of centralized areas to store data safely, information is being held by the contractor and not on a hub.

5.Is there a dedicated point person responsible for ensuring your team remains up to date on compliance obligations?

☐ Yes/☒ No

There should be an update on how the compliance sector is interacting with the data cleaning team.

Total Score for Section 3: [20/25]

Section 4: Continuous Improvement & Compliance Culture

Transparency

☐ Yes/☒
No

Comments

1.Has your team finalized a document outlining the roles and responsibilities of each subgroup or team member?

The team is awaiting feedback to finalize the roles and responsibilities document, which remains unapproved, with no firm timeline for completion yet.

2.What specific improvements (if any) has your team made in the past 3–6 months to strengthen data practices or compliance?

The team relies on code reviews for quality assurance but lacks formal post-data collection reviews, though they plan to implement structured review processes soon.

3.Do you conduct internal reviews or debriefs after completing major collection or cleaning cycles?

☐ Yes/☒
No

Sprint retrospectives are set to begin in May, allowing the team to reflect on completed sprints and identify improvements, with Gilbert emphasizing their value as learning tools.

4.When was the last time your team reviewed or discussed compliance expectations in a meeting or formal setting?

The team hasn't had recent formal discussions on compliance but expects designated individuals to provide guidance, acknowledging that integrating compliance is challenging, especially with data cleaning responsibilities. They are open to improving compliance awareness as roles become clearer.

5.What tools or support do you feel would help your team become more compliant or secure?

The team discussed the need for a centralized data repository, monitoring software, and antivirus tools to improve compliance and security, with a preference for automation and recommendations from leadership, while recognizing the challenges of managing individual device security.

Total Score for Section 4: [/10]

For Official Use Only:

Section 6: Continuous Improvement & Recommendations

Best Practices & Future Improvements

Feedback Area	Suggested Action	Compliance Risk Level	Responsible Party	Deadline	Status
What audit finding is being addressed?	What corrective action is proposed?	High, medium, limited, low, minimal risk etc.	Who owns the action?	By when?	Open, Closed, In progress etc.
***No dataset repository	Improve use of existing SharePoint and Microsoft Teams ecosystem	High	Team Leads/Data Librarian		Open
***No vulnerability assessments	Schedule bi-annual vulnerability scans of cleaning tools and environments	High	Team Leads/Compliance Liaison		Open
No retention/disposal policy	Draft clear data retention timelines and secure disposal SOPs	Medium	Team Lead/Scrum Master		Open
No incident response plan	Develop and disseminate a simple IRP covering compromised accounts or tool misuse	High	Compliance Team		Open
No designated compliance liaison	Appoint a member responsible for tracking obligations and reporting issues	Medium	Team Lead		Open
***Incomplete roles document	Finalize and circulate the roles/responsibilities guide	Medium	Team Leads/Support		Open
No post-sprint debriefs yet	Begin the planned retrospectives in May	Medium	Scrum Master		Open

Feedback Area	Suggested Action	Compliance Risk Level	Responsible Party	Deadline	Status
	and include compliance checks				
***Need for tools and leadership support	Provide automation tools, antivirus, and storage management guidance	High/Medium	Leadership		Open

***Notes:

1. Due to the large size of datasets and existing device limitations, maintain use of SharePoint but adopt a structured hybrid approach: store only dataset links or compressed data partitions centrally. Additionally, the team can explore splitting datasets into more manageable sections or storing lightweight versions of datasets accompanied by documentation/cleaning rules. The team can further explore designating a Data Librarian or a rotating team member to maintain the integrity of the SharePoint,

folder structure, naming conventions and any version control. This balances accessibility with consistency and may help mitigate strain on local devices.

2. Vulnerability assessments can be carried out in a light weight but structured manner to examine the following:

- Devices used, tools and scripts used for cleaning, location of critical scripts/information and backups, data pipelines and cleaning practices, folder structure and file naming. This would help in determining what systems need updating for protection and where vulnerabilities might arise.
- Risks should be documented and classified in a log of issues with each vulnerability, potential impact and urgency (e.g., “Low – Improper file naming” vs “High- Important scripts stored locally with no back up”)

3. This has since been completed April 25th, 2025. Any major changes or updates should be provided to the Compliance team.




4. The team currently operates with limited external support, relying heavily on their own initiative and internal coordination. To strengthen their capacity and reduce risk, leadership should prioritize providing basic tools such as antivirus software, automated utilities to assist with routine cleaning tasks, and better guidance around managing storage for large datasets. These improvements would not only reduce friction in day-to-day workflows but also help address ongoing concerns around performance bottlenecks and data fragmentation. By enabling the team with proper resources and oversight, long-term sustainability and compliance will be easier to maintain.


The Data Cleaning team demonstrates technical strength and an improving compliance culture. However, the absence of structured documentation and reactive governance processes creates medium-level compliance risks. With leadership support and a few focused improvements, the team is well-positioned to reach full compliance in the next cycle.

Final Compliance Score Calculation

Each 'Yes' = 5 pts, 'Partial/In Progress' = 3 pts, 'No' = 0 pts; N/A = not included in score.

Section	Max Score	Score Achieved	Compliance Status
Data Cleaning & Processing	40	[40]	[Compliant]
Data Storage, Security & Privacy	25	[17]	[Partially Compliant]
Compliance & Accountability Measures	25	[20]	[Partially Compliant]
Continuous Improvement & Compliance Culture	10	[6]	[Partially Compliant]
Total	100	[83]	[Partially Compliant]

Compliance Score (%)	Final Rating
90 - 100%	 Compliant – No major risks, minor improvements suggested.
70 - 89%	 Partially Compliant – Some risks identified; action needed.
Below 70%	 Non-Compliant – Does not meet requirements, immediate action required.

Final Rating: [ **Partially Compliant** – Some risks identified; action needed.]

Auditor's Signature:

Team Lead's Signature:

Date: