Audit Planning

Audited Company	Nguyen Auditing Services
Reason for Audit	Deployment at client sites
Date	17/07/2025
Auditor	Khuong Nguyen
Approval: (signature of lead auditor)	K. Nguyen
Date of approval	20/07/2025

Table of Contents

1. Audit Plans

1.1 Scope of Audit

- Audit covers:
  - The on-premise lakehouse (data storage & processing environment)
  - The public-facing site
  - The GitHub repo and CI/CD pipeline (if applicable)
  - Security, GDPR, Risk Management

1.2 What's NOT Covered

- No full penetration test
- No network stress/performance testing

1.3 Reference Materials

- ISO 27001 Annex A
- NIST RMF & 800-53 control families (low/medium baseline)
- GDPR (privacy principles & data subject rights)

1.4 Audit Deliverables

- Checklist for Records
- Risk Register
- Audit Report

---

2. Understanding The Environment

2.1 Assets and information classification

- Employee List & Political Landscape

    - Number of employee: 1 (due to personal innovation project)
- Physical Assets
    - Host PC: 1 (Lenono P340 Tiny)
    - GPU: 1 (RTX A2000 12GB)
    - Power adapter: 1 (Lenovo)
    - Client PC: 1 (Macbook Pro M1)
    - Physical Storage: 2 (SSDs)
- Digital Assets
    - Documents related to Cyber Risks: 3 files
    - Source code of the project: 1 repo

2.2 Technical Details

- The lakehouse architecture
    - Servers: Hosted on-premise
    - Databases:
        - SSDs as local storage
        - PostgreSQL for querying
    - Computing clusters: 1 GPU
- Data sources (raw vs processed).
- Network segmentation
    - public: API calls to dawum.de
    - private: whole lakehouse is hosted on-prem
- Access controls
    - git acccount: nbkhuong (Developer)

2.3 Organizational Settings, Policies & Compliance

- Access management
  - Physical Security
    - No central data center; hosting via GitHub Pages (inherently secured by provider).
    - No badge access system needed (fully remote project).
  - Network & System Access
    - PostgreSQL access limited to core team only.
    - MFA not yet enforced (High-Risk finding — see Risk Register).
    - Audit logs captured for all queries.
  - User Access Rights
    - Role-based access for engineers and analysts.
    - Quarterly reviews of permissions; immediate revocation for departed contributors.
- Incident response procedures
  - Incidents currently managed ad hoc; logs maintained
  - Incident response plan is checked annually
- GDPR handling
  - Data Subject Rights:
    - No formal documented process for access/erasure requests (Medium-Risk)
    - Must implement a tracking and response mechanism (30-day SLA).
  - Breach Notification:
    - 72-hour rule to notify the responsibles
  - Privacy Policy & Cookies:
    - Google Analytics present, but no consent banner or public privacy policy
- Change & Code Management

- GitHub-based deployments
- Branch protection
- CI/CD pipeline uses mechanism for dependency scanning and secrets scanning

3. Data Policies

3.1 Data Processing

- Basic web analytics data (IP address, browser metadata, page views)
- No names, emails, or direct identifiers are collected

3.2 Data Storage

- Analytics logs are stored on secure servers managed by the hosting provider
- No raw personal data is permanently stored by the project team

3.3 Access Controls

- Only core team members with GitHub repository access can view aggregated analytics
- Access is controlled via GitHub authentication and role-based permissions

3.4 External Exposure

- Aggregated, anonymized statistics may be displayed publicly
- No personally identifiable information (PII) is shared or sold