Ethics, Privacy, and Regulation in Informatics

Tópicos de Informática para Automação

Mário Antunes

December 22, 2025

Universidade de Aveiro

Table of Contents i

Ethics in Informatics

RGPD (GDPR) and Privacy Assurance

European AI Act

How to Protect Ourselves and Our Users

Ethics in Informatics

The Impact of Collecting and Sharing Data i

- **Data is Power**: In the digital age, “Money is one thing, but data is power”.
- **Real-world Consequences**:
    - **Political Manipulation**: Data analytics firms (e.g., Cambridge Analytica) have used personal data to micro-target voters, potentially influencing democratic outcomes like Brexit.
    - **Financial Fraud**: Leaked or mishandled data can lead to identity theft and financial loss, where institutions might blame users for “phishing” despite systemic vulnerabilities.

The Impact of Collecting and Sharing Data ii

- **The Ethical Imperative**: As developers, we must recognize that behind every data point is a human being with fundamental rights.

Indirect Identification i

- **The Myth of "Anonymous" Data**: Removing names is often insufficient to protect identity.
- **Definition**: A person is "identifiable" if they can be distinguished not just directly (name, ID), but **indirectly** by combining factors like location, age, gender, or physical characteristics.
- **Re-identification Risks**:
  - Studies show that 87% of the US population could be uniquely identified using only three data points: **Zip Code, Birth Date, and Sex**.

Indirect Identification ii

  - **Example**: The “Netflix Prize” dataset was anonymized, but researchers re-identified users by cross-referencing movie ratings with public IMDb profiles.
- **Lesson**: Always assume data can be re-linked.

RGPD (GDPR) and Privacy Assurance

What is RGPD? i

- **Regulation (EU) 2016/679**: The General Data Protection Regulation (RGPD/GDPR).
- **Core Philosophy**: Privacy is a **fundamental human right**, not a luxury.
- **Scope**: Applies to *any* entity processing data of EU residents, regardless of where the processing takes place.

The Level of Privacy We Must Provide i

Developers must ensure systems adhere to these key principles:

1. **Lawfulness, Fairness, & Transparency**: No hidden processing; users must know what is happening.
2. **Purpose Limitation**: Data collected for “Project A” cannot be used for “Project B” without new consent.
3. **Data Minimization**: Collect only what is strictly necessary.
4. **Accuracy**: Data must be correct and up-to-date.
5. **Storage Limitation**: Delete data when it is no longer needed.

The Level of Privacy We Must Provide ii

6. **Integrity & Confidentiality**: Ensure security against unauthorized access or loss.

Privacy by Design & Default i

- **By Design**: Privacy measures must be embedded into the architecture of the software from the very start, not added as a patch later.
- **By Default**: The strictest privacy settings should apply automatically without user intervention (e.g., a social media profile should be private by default).
- **Accountability**: The controller must be able to *demonstrate* compliance through documentation and logs.

European AI Act

Overview of the AI Act i

- **Risk-Based Approach**: The AI Act categorizes AI systems based on the potential risk they pose to users' safety and fundamental rights.
    - **Unacceptable Risk**: Banned (e.g., social scoring, real-time remote biometric identification in public spaces by law enforcement, with exceptions).
    - **High Risk**: Permitted but strictly regulated (e.g., AI in education, employment, critical infrastructure).
    - **Limited Risk**: Transparency obligations (e.g., chatbots must reveal they are AI).
    - **Minimal Risk**: Unregulated (e.g., spam filters).

Privacy and Auditability in AI i

- **Relation to Privacy**: High-risk AI systems must run on high-quality data to avoid discrimination and must adhere to GDPR principles like data governance.
- **Auditability Requirements**:
  - **Logging**: Systems must automatically record events (logs) to trace functioning and identify risks.
  - **Technical Documentation**: Developers must maintain detailed documentation to prove compliance to authorities.
  - **Human Oversight**: Systems must be designed so that natural persons can oversee their operation and override decisions.

How to Protect Ourselves and Our Users

Guaranteeing Protection: Technical Measures i

- **Pseudonymization**: Processing data such that it can no longer be attributed to a specific subject without additional information (key), which must be kept identifying separate.
- **Anonymization**: Irreversible removal of identifiers. *Note: Truly anonymized data falls outside GDPR scope, but it is hard to achieve*.
- **Encryption**: Mandatory for sensitive data transmission and storage to prevent unauthorized access.

Guaranteeing Protection: Process Measures i

- **Data Protection Impact Assessment (DPIA/AIPD)**:
  - Before starting a project with high risks, you must assess the impact on data privacy.
  - **Steps**:
    1. Describe the processing operations.
    2. Assess necessity and proportionality.
    3. Identify risks to rights and freedoms.
    4. Define measures to mitigate those risks.
- **Consent Management**: Consent must be **free, specific, informed, and explicit**. Pre-ticked boxes are invalid.

Protecting Ourselves (As Professionals) i

- **Clear Liability**: Understand who is the “Controller” (determines purpose) vs. “Processor” (technical handler). As a developer, you often act on behalf of a controller, but you must ensure your tools are compliant.
- **Continuous Vigilance**:
    - Monitor re-identification risks in big data.
    - Stay updated on adequacy decisions for international data transfers (e.g., data stored on US servers).

Further Resources

1. **Official Legal Texts**:
   - GDPR (EU 2016/679) Full Text
   - European AI Act Text
2. **Handbooks**:
   - *Handbook on European Data Protection Law* (FRA/Council of Europe).
3. **Institutional Guidance**:
   - Data Protection Officer (DPO) contacts at your institution here
   - CNPD (Comissão Nacional de Proteção de Dados) guidelines.