Security and Privacy

Lab Assignment Worksheet

João Vilela - Manuel E. Correia

2023/2024

— Assignment #3: Privacy Impact Assessment

Due date: April 29, 23:59

Grading: Assignment #3 is worth **3 points**

TO BE DONE IN **GROUPS OF TWO (MANDATORY)**

The goal of this project is for you to investigate the steps and methodology to carry out a Privacy Impact Assessment (PIA). You should then perform a (simplified) PIA by applying relevant methodologies identified in the previous step. The PIA shall be performed for the setup of a real-world project (COP-MODE), that required data collection campaigns from a set of users.

For this, you should:

1. Study related literature/guidelines [1, Chapter 11] [2] to understand the concepts and steps of a methodology to perform a PIA
2. Select a tool for performing PIA [3][4] and/or study examples of PIA reports [5][6] that can serve as a reference
3. Perform the PIA and document your analysis
4. Submit the report in moodle

Essential to a PIA is a description of how PII flows into, through, and out of the system. The workflow diagram of Figure 1 is a typical example workflow of PII processing that summarizes the stages and entities involved. The entities are defined as follows [1]:

- PII principal/data subject: a person that can be identified directly (by reference to an identifier such as name, identification number) or indirectly (through information that can be traceback to the person);
- PII controller: this is the responsible for determining the purposes and means of processing personal data, regardless of whether such data are collected, stored, processed or disseminated by that party or by an agent on its behalf;
- PII processor: the responsible for processing data on behalf and in accordance with the instructions of a controller (may be the same entity as the controller);
- Third party: an entity/person that is authorized to process personal data, other than the data subject, controller and processor.

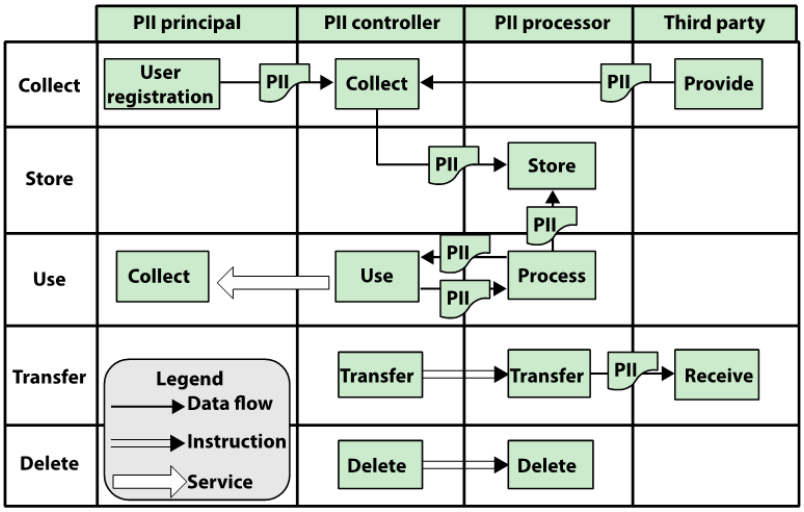

Figure 1: Workflow Diagram of PII Processing (from [1]).

COP-MODE Setup

Consider the setup of the COP-MODE project in which your institution needs to perform data gathering campaigns from users. The PII principals/data subjects are the users that participate in the campaigns, and there are third parties interested in processing personal data. **You are the PII controller/processor, responsible for determining the purposes and means of processing personal data**.

After some initial setup in which the smartphones are pre-installed with users applications, the PII principals/participants carry smartphones for data collection for 1 week, as depicted in Figure 2. During this period several types of data are collected and sent to the COP-MODE server under control of the PII collector/processor.

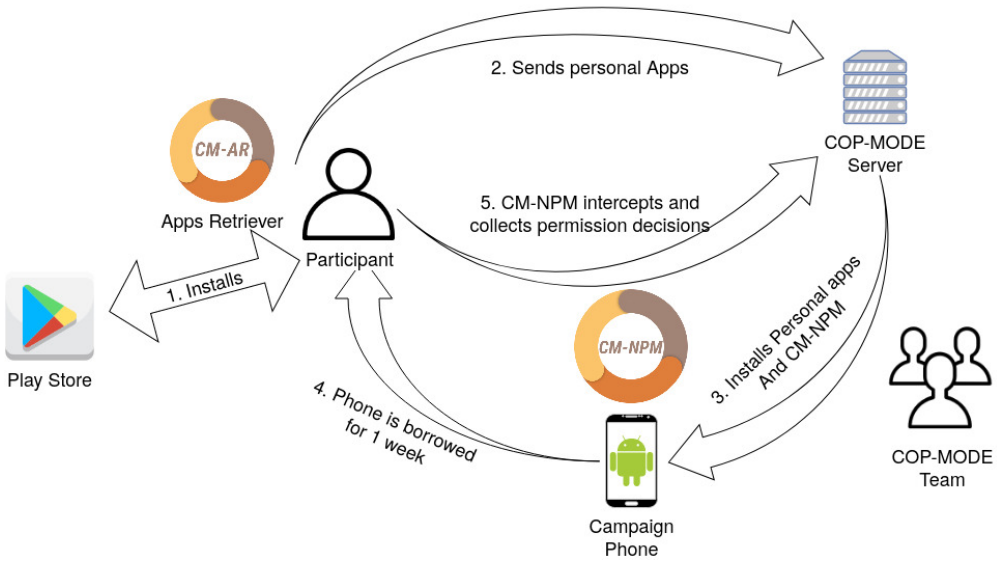

Figure 2: COP-MODE Campaigns Methodology. More details here.

In addition to the data necessary for the COP-MODE project, it is also needed to collect a form of contact, in this case e-mail address. This is the main means of communication and therefore needed to be stored until the participant returns the lent smartphone. Figure 3 summarizes the data being collected.

Type of Collection	Data
At beginning	Email address
Snapshot	* Installed Apps and respective permission settings
	Connection Type (wi-fi, cellular, other)
	Device context (idle, in use, in call, etc)
	* Calendar entries (calendar identifier, location and start and end dates)
Continuous	* Geographic Location (latitude, longitude, altitude)
	* Devices nearby (Bluetooth MAC address + WiFi BSSIDs)
At Permission Prompt	Application info (*name, version, category, visibility)
	User decision
	Semantic Location (user input)

Figure 3: COP-MODE Summary of Main Data Collected. Data labeled with a '*' is deemed sensitive.

PIA Goal

For the context set out above, you should follow methodologies to perform a PIA [1, 2] and prepare a report of the PIA process to submit in moodle. You should use the tool in [3] to assist you in this process, but **the report should not be the output of the tool directly**. Your PIA report should include at least the following points:

1) Describe/characterize the system setup that is the subject of this PIA

   This should include an overview of the system, summarizing the system design and personnel involved. It should also focus on what PII is collected/processed, which PII principals are affected, and what systems and processes are involved in the handling of PII.

2) Initial risk assessment for a **plain setup without any corrective measures** (e.g. data stored in clear text, insecure communications) using tool [3]

   This should include the definition of Likelihood and Severity of the identified risks, leading to a privacy risk matrix (risk mapping in [3])

3) A plan of action (corrective measures) to address the identified risks

   This should specify corrective measures that can be put in place to address the situations of relevant risk, as identified in step 2). It should consider techniques that you have learned throughout this course (e.g. cryptographic mechanisms, secure communications).

   Apart from the risks present in the tool [3], you should also consider the following risks and describe mitigations to address them on the form of a two-column table (risk, mitigation) as follows.

Risk	**Mitigation**
Data leak through eavesdropping of communications between smartphones and the project server.	*TBD - your mitigations*
Data leak from unauthorized server/data access. Access to the stored data should be limited to authorized users and the server should be accessible remotely, but not exposed/accessible directly to from the outside.	*TBD - your mitigations*
Data at rest linkage. The data in the server should not be linkable to an individual by third parties. I.e. the email address must be protected (not clear text), but the PII collector/processor must still be able to link it to the data in the server (e.g. for deletion in case the participant requests it).	*TBD - your mitigations*
Sensitive information leakage, in particular the names of the applications that are collected. It is required to know that a given application was used and is the same for the several entries that are collected, but the names of the applications should not be present in the server in raw format.	*TBD - your mitigations*

4) Risk assessment for an **improved setup with application of corrective measures**. Assessment/evaluation of the evolution of the risks with the corrective measures using tool [3].

   This should re-evaluate the levels of risk identified in step 3) to assess and justify the effect of the corrective measures on the attained risk level. The tool [3] allows you to observe in the Risk Mapping matrix the evolution of the risk levels before and after corrective measures.

Submit your report in moodle.

Relevant references:

[1] William Stallings, "Information Privacy Engineering and Privacy by Design", Pearson Addison-Wesley, 2020 (main content available on moodle: "PIA – A Primer")

[2] Guidelines and software for PIA

[3] CNIL PIA tool (**recommended tool**). Tutorial video: https://www.youtube.com/watch?v=5J3h9zIFVDo

[4] https://www.enisa.europa.eu/risk-level-tool/risk

[5] Statistics New Zealand, "Privacy Impact Assessment for the Integrated Data Infrastructure", 2012. https://www.stats.govt.nz/assets/Uploads/Retirement-of-archive-website-project-files/Privacy-Impact-Assessment/Privacy-impact-assessment-for-the-Integrated-Data-Infrastructure/idi-privacy-impact-assessment.pdf

[6] E.U. Smart Grid Task Force, "Data Protection Impact Assessment Template for Smart Grid and Smart Metering Systems", September 2018, https://energy.ec.europa.eu/topics/markets-and-consumers/smart-grids-and-meters/data-protection-impact-assessment-smart-grid-and-smart-metering-environment_en

Evaluation Criteria

- PIA report, including: [75%]
  - Description of the setup/context [10%]
  - Privacy risk assessment [25%]
  - A plan of action (corrective measures) to address the identified risks [25%]
  - Assessment/evaluation of the evolution of the risks with the corrective measures [15%]
- Structure and organization of the report [25%]