Supplementary Materials for Mining and Classifying Privacy and Data Protection Requirements in Issue Reports Manuscript

1 Additional examples for privacy requirements identification (Section 4.1)

This section provides more examples of the privacy requirements derived in the privacy requirements identification process.

It is important to note that we have followed the GBRAM to extract and refine requirements from narrative statements. For some statements, they are straightforward since privacy requirements can be directly derived from them. However, we have restructured and refined some privacy requirements to emphasise the functionalities that should be provided by software systems. For example, Art. 7(1) in the GDPR states "Where processing is based on consent, the controller shall be able to demonstrate that the data subject has consented to processing his or her personal data". From this statement, we derive requirement *R17* ***SHOW*** *the relevant stakeholders the consent given by the data subjects to process their personal data.*

The lawfulness of processing is one of the key principles for protecting privacy of data subjects. Based on the GDPR Art. 6, the processing of personal data will be lawful if the data subjects give consent *or* the processing is required for other legal conditions, such as the processing is necessary for the performance of contract, compliance with a legal obligation and protecting vital interests. R35 in the taxonomy addresses the requirement of obtaining consent from the data subjects for the processing. For other legal conditions, the data controllers must inform the data subjects if they need to process personal data under those conditions. Requirements R38 and R39 in the taxonomy cover these scenarios.

After lawfully obtaining personal data, the data controllers must provide the data subjects with mechanisms to execute their individual rights in the system. All the privacy requirements related to the rights of data subjects are covered in our taxonomy (i.e., the right to be informed (R12, R24, R26, R30, R31 and R50), the right of access (R1), the right to rectification (R45), the right to erasure (R44), the right to restriction of processing (R4), the right to data portability (R33), the right to object (R3), and the rights in relation to automated decision making and profiling (R21)).

A number of requirements (e.g. R35, R39 and R60) can be triggered when the software is dealing with special category data/more sensitive data. In addition to consent, there are other conditions for a lawful processing of special categories of personal data (e.g. necessary for protecting vital interests and public interests). Requirement R35 covers the consent condition. R39 and R60 in our taxonomy address the remaining conditions specified in GDPR Art. 9. R39 requires the data controllers, if not obtaining the consent, must provide the data subjects the purpose(s) of the processing of special categories of personal data to the data subjects. The controllers also require to protect those personal data with appropriate measures (R60). The key privacy requirements related to international data transfers are also covered as follows: i) the data subjects must be informed about the transfer of their personal data to a third country or an international organisation (R9); ii) the personal data must be appropriately protected (R60) and iii) the transfer of personal data must comply with local requirements (R65).

We note that an article, a section or a statement can lead to the identification of more than one requirements. For example, we derived 17 privacy requirements (i.e. R1-R4, R6, R9, R20-22, R27, R29, R34, R37, R39, R44-R45 and R55) from the Art. 13 in GDPR[1]. The following example demonstrates two privacy requirements that were derived from a statement. A statement in Section 23-6 in Thailand PDPA states "In collecting the Personal Data, the Data Controller shall inform the data subject, ... (5) information, address and the contact channel details of the Data Controller, where applicable, of the Data Controller's representative or data protection officer;", we derived two requirements from this statement which are *R20 PROVIDE the data subjects with the contact details of a data protection officer (DPO)* and *R22 PROVIDE the data subjects with the identity and contact details of a controller/controller's representative.*

We also note that we excluded the articles related to the DPOs in our study due to the following reasons. Firstly, the articles related to DPOs in GDPR mainly focus on their duties, tasks and responsibilities (i.e. Art. 37 - 39). The DPOs requirements are related to governance aspect rather than software requirements aspect. Secondly, the main role that directly determines the activities related to the processing of personal data is the data controllers. This makes the data controllers the key stakeholder in governing how personal data is processed and how the processing activities should be done in software development level. Finally, we have not found any DPOs requirements reported in issue reports in our study. This finding implies that the DPOs requirements were not reflected as software requirements in this context.

[1]See the file *Privacy-requirements-references* in the replication package for more details [1].

2 Comparing the level of abstraction between the privacy requirements derived from GDPR, ISO/IEC 29100, Thailand PDPA and APEC privacy framework (Section 4.2)

The privacy requirements derived from GDPR, ISO/IEC 29100, Thailand PDPA and APEC privacy framework are in fact at the same level of abstraction. Table 1 below presents the examples demonstrating the privacy requirements that were derived from different regulations and frameworks, but refer to similar things. For example, we derived ***PROVIDE*** *the data subject the recipients or categories of recipients of the personal data* from Art. 13(1)(e) in GDPR, ***PROVIDE*** *the types of persons whom the PII can be transferred* from openness, transparency and notice principle in ISO/IEC 29100, ***INFORM*** *the data subject the categories of Persons or entities to whom the collected Personal Data may be disclosed* from Section 23 in Thailand PDPA and ***PROVIDE*** *the types of persons or organisations to whom personal information might be disclosed* from Point 21 in APEC framework. These four requirements demonstrate that they are at the same level of abstraction and aim to achieve the same goal. They can be merged in the requirements refinement process. Meis et al. has also confirmed that GDPR and ISO/IEC 29100 are at the same level of abstraction [2].

Table 1: Some sample requirements derived from GDPR, ISO/IEC 29100, Thailand PDPA and APEC privacy framework demonstrate that they are at the same level of abstraction.

Requirements derived from GDPR statements	**GDPR reference**	**Requirements derived from ISO/IEC 29100**	**ISO/IEC 29100 reference**	**Requirements derived from Thailand PDPA**	**Thailand PDPA reference**	**Requirements derived from APEC framework**	**APEC framework reference**
PROVIDE the existence of the right to withdraw consent	13(2)(c)	ALLOW a PII principal to withdraw consent	5.2	ALLOW the data subject to withdraw his or her consent	19-5	None	None
PROVIDE the data subject the recipients or categories of recipients of the personal data	13(1)(e)	PROVIDE the types of persons whom the PII can be transferred	5.8	INFORM the data subject the categories of Persons or entities to whom the collected Personal Data may be disclosed	23-5	PROVIDE the types of persons or organisations to whom personal information might be disclosed	21-4
PROVIDE the data subject the purposes of the processing	13(1)(c)	PROVIDE the PII principal the purpose of the processing of PII	5.8	INFORM the data subject the purpose of the collection, use, or disclosure of the Personal Data	19-3	PROVIDE the individuals with the purpose their information is to be used	(21-23)-1
PROVIDE the data subject the categories of personal data concerned	15(1)(b)	PROVIDE the PII principal the specified PII required for the specified purpose	5.8	INFORM the data subject the Personal Data to be collected	23-4	PROVIDE the individuals with what personal information is collected	(21-23)-1
IMPLEMENT appropriate technical and organisational measures to protect personal data	32(1)(a)	PROTECT PII with appropriate controls	5.11	PROVIDE appropriate security measures for preventing the Personal Data	37-2	IMPLEMENT organizational controls to prevent from the wrongful collection or misuse of personal information	20-1, 28-1, 28-2

3 Top 10 privacy requirements and their descriptive statistics (Section 7.1)

The following tables show the top 10 privacy requirements mined in Chrome and Moodle projects together with their frequency based on issue type.

Table 2: A table summarising top 10 privacy requirements with their frequency based on issue type in Chrome project.

Project	Requirement	Category	Subcategory	Issue type	Frequency	Total frequency
Chrome	R30	2) Notice	2.1) Data subjects	Bug	151	209
				Bug-regression	22	
				Bug-security	1	
				Feature	32	
				Task	1	
				Unspecified	2	
Chrome	R44	1) User participation	-	Bug	151	204
				Bug-regression	15	
				Bug-security	4	
				Feature	21	
				Task	1	
				Unspecified	12	
Chrome	R60	7) Security	-	Bug	80	135
				Bug-regression	4	
				Bug-security	25	
				Feature	17	
				Task	1	
				Unspecified	8	
Chrome	R8	3) User desirability	3.1) Consent 3.3) Preference	Bug	89	144
				Bug-regression	4	
				Bug-security	25	
				Feature	17	
				Task	1	
				Unspecified	8	
Chrome	R36	3) User desirability	3.2) Choice	Bug	74	119
				Bug-regression	7	
				Bug-security	4	
				Feature	32	
				Task	0	
				Unspecified	2	
Chrome	R45	1) User participation	-	Bug	58	73
				Bug-regression	1	
				Bug-security	0	
				Feature	11	
				Task	0	
				Unspecified	3	
Chrome	R53	4) Data processing	4.4) Erasure	Bug	46	70
				Bug-regression	4	
				Bug-security	2	
				Feature	9	
				Task	0	
				Unspecified	9	
Chrome	R1	1) User participation	-	Bug	31	49
				Bug-regression	6	
				Bug-security	0	
				Feature	11	
				Task	0	
				Unspecified	1	
Chrome	R26	2) Notice	2.1) Data subjects	Bug	20	29
				Bug-regression	1	
				Bug-security	0	
				Feature	6	
				Task	1	
				Unspecified	1	
Chrome	R41	4) Data processing	4.1) Collection	Bug	12	17
				Bug-regression	1	
				Bug-security	0	
				Feature	3	
				Task	0	
				Unspecified	1	

Project	Requirement	Category	Subcategory	Issue type	Frequency	Total frequency
Chrome	R30	2) Notice	2.1) Data subjects	Bug	151	209
				Bug-regression	22	
				Bug-security	1	
				Feature	32	
				Task	1	
				Unspecified	2	
Chrome	R44	1) User participation	-	Bug	151	204
				Bug-regression	15	
				Bug-security	4	
				Feature	21	
				Task	1	
				Unspecified	12	
Chrome	R60	7) Security	-	Bug	80	135
				Bug-regression	4	
				Bug-security	25	
				Feature	17	
				Task	1	
				Unspecified	8	
Chrome	R8	3) User desirability	3.1) Consent 3.3) Preference	Bug	89	129
				Bug-regression	4	
				Bug-security	25	
				Feature	17	
				Task	1	
				Unspecified	8	
Chrome	R36	3) User desirability	3.2) Choice	Bug	74	119
				Bug-regression	7	
				Bug-security	4	
				Feature	32	
				Task	0	
				Unspecified	2	
Chrome	R45	1) User participation	-	Bug	58	73
				Bug-regression	1	
				Bug-security	0	
				Feature	11	
				Task	0	
				Unspecified	3	
Chrome	R53	4) Data processing	4.4) Erasure	Bug	46	70
				Bug-regression	4	
				Bug-security	2	
				Feature	9	
				Task	0	
				Unspecified	9	
Chrome	R1	1) User participation	-	Bug	31	49
				Bug-regression	6	
				Bug-security	0	
				Feature	11	
				Task	0	
				Unspecified	1	
Chrome	R26	2) Notice	2.1) Data subjects	Bug	20	29
				Bug-regression	1	
				Bug-security	0	
				Feature	6	
				Task	1	
				Unspecified	1	
Chrome	R41	4) Data processing	4.1) Collection	Bug	12	17
				Bug-regression	1	
				Bug-security	0	
				Feature	3	
				Task	0	
				Unspecified	1	

Table 3: A table summarising top 10 privacy requirements with their frequency based on issue type in Moodle project.

Project	Requirement	Category	Subcategory	Issue type	Frequency	Total frequency
Moodle	R44	1) User participation	-	Bug	48	194
				Epic	0	
				Improvement	44	
				New Feature	57	
				Task	7	
				Sub-task	34	
				Functional Test	4	
Moodle	R1	1) User participation	-	Bug	71	186
				Epic	1	
				Improvement	44	
				New Feature	59	
				Task	6	
				Sub-task	1	
				Functional Test	4	
Moodle	R35	3) User desirability	3.1) Consent	Bug	29	161
				Epic	0	
				Improvement	35	
				New Feature	57	
				Task	4	
				Sub-task	33	
				Functional Test	3	
Moodle	R56	7) Security	-	Bug	23	150
				Epic	0	
				Improvement	33	
				New Feature	57	
				Task	3	
				Sub-task	34	
				Functional Test	0	
Moodle	R38	2) Notice	2.1) Data subjects	Bug	21	112
				Epic	0	
				Improvement	32	
				New Feature	56	
				Task	3	
				Sub-task	0	
				Functional Test	0	
Moodle	R42	2) Notice	2.1) Data subjects	Bug	19	109
				Epic	0	
				Improvement	32	
				New Feature	56	
				Task	2	
				Sub-task	0	
				Functional Test	0	
Moodle	R34	1) User participation	-	Bug	42	57
				Epic	0	
				Improvement	7	
				New Feature	4	
				Task	2	
				Sub-task	0	
				Functional Test	2	
Moodle	R26	2) Notice	2.1) Data subjects	Bug	17	43
				Epic	0	
				Improvement	14	
				New Feature	8	
				Task	4	
				Sub-task	0	
				Functional Test	0	

Project	Requirement	Category	Subcategory	Issue type	Frequency	Total frequency
Moodle	R30	2) Notice	2.1) Data subjects	Bug	26	42
				Epic	0	
				Improvement	11	
				New Feature	1	
				Task	3	
				Sub-task	1	
				Functional Test	0	
Moodle	R60	7) Security	-	Bug	27	40
				Epic	0	
				Improvement	7	
				New Feature	2	
				Task	4	
				Sub-task	0	
				Functional Test	0	

References

[1] P. Sangaroonsilp, H. K. Dam, M. Choetkiertikul, C. Ragkhitwetsagul, A. Ghose, Replication Package, 2022. URL: `https://bit.ly/Mining-privacy-reqs-rev22`.

[2] R. Meis, R. Wirtz, M. Heisel, A taxonomy of requirements for the privacy goal transparency, in: Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), volume 9264, 2015, pp. 195–209.