Privacy Impact Assessments in the Wild:
Scoping Review Protocol

Leonardo Horn Iwaya, Ala Sarah Alaqra, Simone Fischer-Hübner

November 2022

Document versioning control			
Editor	**Version**	**Comment**	**Date**
Anom	v1.0	First version of the protocol finalised.	07/11/2022
Anom	v1.1	Version of the protocol approved by all reviewers.	14/11/2022

1 Introduction

As stated in Article 35 of the European General Data Protection Regulation (GDPR) [1], organisations are expected to carry out a Data Protection Impact Assessment (DPIA) – also better known as Privacy Impact Assessment (PIA) – whenever the processing of personal data is likely to result in a high risk to the rights and freedoms of individuals. Although there are various definitions of PIAs, we can adopt Wright's proposal of *"privacy impact assessment as a methodology for assessing the impacts on the privacy of a project, policy, programme, service, product or other initiative and, in consultation with stakeholders, for taking remedial actions as necessary in order to avoid or minimise negative impacts"* [2]. Since PIA methodologies often adopt a risk-based approach for identifying privacy threats and selecting adequate controls, they are usually considered a basic step for achieving privacy by design. PIAs also generate important documentation along its process supporting organisations to demonstrate compliance with privacy authorities. Organisations can also choose to provide public PIA Reports to their customers and users, enhancing the system's transparency to the general population.

Given the growing importance of PIAs as a strategy for privacy-by-design and for demonstrating compliance, this study aims to understand the state-of-the-practice and identify the challenges organisations face when undertaking PIAs. To do so, we will conduct a Scoping Review (ScR) to start addressing our research objective by consulting the body of knowledge of scientific literature. For this ScR, we follow the Preferred Reporting Items for Systematic reviews and Meta-Analyses extension for Scoping Reviews (PRISMA-ScR) guideline [3], also considering other guidelines such as the PRISMA 2020 Statement [4]. For this protocol, we also rely on the PRISMA-P guidelines [5] to ensure the completeness and transparency of the review process.

2 Justification for the ScR

A preliminary search on the Scopus database using the keywords `''privacy impact assessment''` and `''data protection impact assessment''` returns a total of 309 unique entries. No systematic reviews were identified, further motivating the proposed study. In this initial test, a total of 25 studies have been identified, suggesting a feasible "critical mass" for a proper ScR (see Table 1).

3 Research Questions

As part of a formative study, this ScR on the topic of "PIAs in the Wild" aims to carefully identify and synthesise the existing literature about PIAs in practice. By practice, we mean empirical experience, application and practical use of PIAs. Here, the focus is on understanding the practical approaches,

Table 1: Initial list of studies identified.

Ref	Titles	Year
[6]	Data Protection Impact Assessments in Practice: Experiences from Case Studies	2022
[7]	A perfect match: Converging and automating privacy and security impact assessment on-the-fly	2021
[8]	Can data subject perception of privacy risks be useful in a data protection impact assessment?	2021
[9]	The Never-Ending Story: On the Need for Continuous Privacy Impact Assessment	2020
[10]	An organizational scheme for privacy impact assessments	2018
[11]	Privacy impact assessment in practice: The results of a descriptive field study in the Netherlands	2017
[12]	Integrating privacy impact assessment in risk management	2014
[13]	Introducing a privacy impact assessment policy in the EU member states	2013
[14]	Privacy impact assessment reports: A report card	2012
[15]	How Siemens assesses privacy impacts	2012
[16]	Vodafone's approach to privacy impact assessments	2012
[17]	Privacy impact assessment: Optimising the regulator's role	2012
[18]	Privacy impact assessment	2012
[19]	Privacy impact assessment – Great potential not often realised	2012
[20]	Privacy impact assessment in the UK	2012
[21]	Privacy impact assessment in New Zealand – A practitioner's perspective	2012
[22]	Privacy impact assessments in Canada	2012
[23]	Findings and recommendations	2012
[24]	Auditing privacy impact assessments: The Canadian experience	2012
[25]	Challenges to managing privacy impact assessment of personally identifiable data	2012
[26]	Understanding the drivers and outcomes of healthcare organizational privacy responses	2011
[27]	An evaluation of privacy impact assessment guidance documents	2011
[28]	The emergence of privacy impact assessments	2010
[29]	Incorporating privacy outcomes: Teaching an old dog new tricks	2008
[30]	Privacy Impact Assessments: International experience as a basis for UK Guidance	2008

challenges, and opportunities concerning PIAs that have been reported in the scientific literature. With that in mind, the Research Questions (RQs) for this ScR are:

- **RQ1:** *What is the literature and research that discusses and/or evaluates the use of PIAs in practice?* **Objective:** To identify the existing research and to understand the types of studies and types of contributions in the literature, methodological approaches used, main venues and researchers.
- **RQ2:** *What are the main methodologies, positive and negative aspects of PIAs in practice as reported in studies?* **Objective:** To identify relevant characteristics, such as the main methodologies reported in the studies, compile shared experiences, focusing on understanding the main positive and negative aspects related to PIAs in practice.
- **RQ3:** *What is the state of existing primary research studies on the topic?* **Objective:** To further examine and critically appraise existing primary research and identify the studies (or the lack thereof) that have proposed valid and reliable instruments related to PIAs.

The PCC mnemonic, which stands for Population, Concept, and Context [31], can also be used as a guide to frame the review questions:

- The **population** refers to individuals and organisations (such as a company, an institution, or

an association, comprising one or more people and having a particular purpose) involved in conducting PIAs for working systems.

- The **concept** refers to PIAs with emphasis on their application in practice.
- The **context** refers to any organisation undertaking PIAs, yet looking at research publications (any study design) that discuss the topic through varied viewpoints, e.g., researchers, consultants, privacy officers, privacy authorities, industry professionals, etc.

Based on our preliminary search (see Table 1), we envision that this ScR will be composed of three broad types of published research: (i) studies on the perspectives of various stakeholders (e.g., data subjects, DPOs, DPAs, privacy engineers, etc) about PIAs in practice; (ii) studies that share practical experiences about undertaking PIAs in specific projects, jurisdictions, or as part of a broader process in the organisation; and, (iii) user studies that evaluate PIA methodologies and their artefacts in a given population. Notwithstanding, other types of research may be revealed during the review process, and if deemed appropriate, they shall be included in the ScR.

4 Eligibility Criteria

The publications retrieved through the searches will have their titles and abstracts screened by two reviewers and selected following the inclusion criteria that link back to the research objectives. Table 2 provides a summary of the inclusion and exclusion criteria.

Table 2: Eligibility criteria.

Inclusion	**Rationale**
1. Studies on PIAs related to how they are conducted in practice	Publications need to address both aspects: (1) PIAs or DPIAs; and their (2) practice, application, and empirical experience.
2. Studies that perform an empirical evaluation of PIA methodologies in practice	We also consider studies that discuss and/or more profoundly evaluate the use of PIAs in practice through the lens of many stakeholders (e.g., privacy officers, policy-makers, data subjects, etc.).
Exclusion	**Rationale**
1. Studies on PIAs that are only theoretical, not solidly linked to practice	Many publications address the topic of PIAs by proposing new methodologies, theoretically analysing and comparing approaches, or discussing them in terms of hypothetical use cases. Such publications will not be considered since they are not strongly linked to a practical observation or evaluation.
2. Foreign language studies	Papers published in foreign languages that the authors do not master will be acknowledged, and their existence documented with 'language' recorded as the reason for exclusion. The languages that the authors can review are English, Portuguese, Spanish, Swedish, Dutch, and German.

4.1 Types of Sources

As mentioned, this ScR will consider studies of any research design. Quantitative studies will consider both experimental and quasi-experimental study designs, including randomised controlled trials, non-randomised controlled trials, before and after studies and interrupted time-series studies. In addition, analytical observational studies, including prospective and retrospective cohort studies, case-control studies and analytical cross-sectional studies, will be considered for inclusion. This review will also consider descriptive observational study designs, including case series, individual case reports and descriptive cross-sectional studies for inclusion. Qualitative studies will also be considered that focus on qualitative data including, but not limited to, designs such as phenomenology, grounded theory, ethnography, qualitative description, action research and feminist research. In addition, systematic reviews that meet the inclusion criteria will also be considered, depending on the research question. Text and opinion papers will also be considered for inclusion in this scoping review.

5 Methods

The proposed scoping review will be conducted in accordance with the PRISMA-ScR guideline [3].

5.1 Search Strategy

We will search four bibliographic and full-text databases: Scopus, Web of Science, IEEE Xplore, and ACM Digital Library. The databases will be searched, without any specific period of time, for peer-reviewed publications. A structured search strategy will be used based on keywords and ordered vocabulary relevant to our study objectives. Keywords relevant to our inclusion criteria include: (1) privacy impact assessment; and (2) data protection impact assessment. Studies published in English, Portuguese, Spanish, Swedish, Dutch, and German will be considered since these are the languages that the reviewers master. Studies in other languages will, however, be acknowledged and documented.

We removed the acronyms as search terms, i.e., PIA and DPIA, because in our initial tests, they were returning many publications in completely unrelated areas. We also decided to exclude other limiting key terms (e.g., practice, empirical, application) since the number of returned results was already of a feasible size for the screening process. Therefore, the generic search string is structured as follows: `"privacy impact assessment*" OR "data protection impact assessment*"`. The search strategy, including all identified keywords, will be adapted for each included database and/or information source.

We also plan to conduct backward searches (screening references cited in included studies) and forward searches (exploring studies that cite included studies using Google Scholar). A preliminary search in the grey literature (e.g., unpublished work, reports, website information, newspaper articles) using the OpenGrey [1] literature database has not returned any results for the keywords.

5.2 Study/Source of Evidence selection

To manage the screening process, we will export search results from each database and then import them to the RAYYAN [2] software, allowing both researchers to select papers independently (i.e., double-blinded) and manage conflicts by a third reviewer. Duplicated publications can also be removed using RAYYAN during the selection process.

The full-text of selected citations will be assessed in detail against the inclusion criteria by two or more independent reviewers. Reasons for excluding sources of evidence in full-text that do not meet the inclusion criteria will be recorded and reported in the scoping review. Any disagreements that arise between the reviewers at each stage of the selection process will be resolved through discussion or with an additional reviewer. The results of the search and the study inclusion process will be reported in full in the final scoping review and presented in a Preferred Reporting Items for Systematic Reviews and Meta-analyses extension for scoping review (PRISMA-ScR) flow diagram [3].

5.3 Data Extraction

After selecting the relevant studies, the reading in full of the publications starts. The data extracted will include specific details about the participants, concept, context, study methods and key findings relevant to the research questions. The primary reviewer will extract significant data from each publication and initiate the data charting at this stage. A secondary reviewer will perform quality control, assessing the consistency for a representative sample of selected publications. This pre-trial will allow us to resolve any conflicts, discuss and settle among reviewers, and a new data charting iteration can be performed following an agreed and consistent process. We will use a charting form to include study details, characteristics, and key findings related to the review question. This process of reading, extracting and charting data constitutes an iterative process in which researchers can continuously criticise, agree, and update the charting form as needed. To ensure transparency in the reporting, we will take notes to explain the rationale for the charting form creation and updates.

Preliminary components of the charting form are:

[1] OPENGREY.EU – Grey Literature Database (`https://opengrey.eu/`)

[2] Rayyan – Intelligent Systematic Review (`https://www.rayyan.ai/`)

- Demographics of the publications: author(s), publication date, title, journal, and citation. Most of these data points can be easily obtained from the final list of publications after the selection process using RAYYAN software (i.e., exported to Excel format).
- Potential facets for data charting: publication type, research type, contribution type, organisation type, PIA methodologies and artefacts, relevant regulations, types of organisations and industry sectors, countries and jurisdictions.
- Main contributions and summary of the work.
- Study results and relevant conclusions regarding PIAs in practice.
- (If appropriate) Critical appraisal of quantitative and qualitative studies.
- Potential references cited in the study for backward snowballing.

The draft data extraction tool will be modified and revised as necessary while extracting data from each included evidence source. Modifications will be detailed in the scoping review. Any disagreements that arise between the reviewers will be resolved through discussion or with an additional reviewer. If appropriate, authors of papers will be contacted to request missing or additional data, where required. The critical appraisal of the studies may be considered appropriate if a significant number of primary research studies (qualitative or quantitative) are selected.

5.4 Data Analysis and Presentation

The data charting process will provide us with a structured dataset that can be used to summarise the research results. Quantitative data (e.g., year of publication, citations) and qualitative data (e.g., publication type, contribution type) can be discussed and reported using tables and figures when necessary. This first step of the data synthesis will allow us to create a map of the body of knowledge.

As a second step, a narrative summary will accompany the tabulated and/or charted results and will describe how the results relate to the review's objective and questions. The narrative review will also describe key elements of the PIAs in the wild, such as methodologies used, related organisational processes, and positive and negative aspects. Besides, the search process will be described using the PRISMA flow chart diagram.

Preliminary components of the data synthesis:

- Demographics of the publications.
- Narrative summary of the literature.
- Main positive and negative aspects for PIAs in practice.
- (If appropriate) summary of the critical appraisal of studies.

6 Dissemination

The results of this study will be made available in scientific venues, such as conferences and journals. We also plan to disseminate findings among key stakeholders, such as international associations of privacy practitioners, research communities and industry partners.

References

[1] European Commission, "Regulation (EU) 2016/679 of the European Parliament and of the Council of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation)," *Official Journal of the European Union*, no. April, 2016.

[2] D. Wright, "The state of the art in privacy impact assessment," *Computer law & security review*, vol. 28, no. 1, pp. 54–61, 2012.

[3] A. C. Tricco, E. Lillie, W. Zarin, K. K. O'Brien, H. Colquhoun, D. Levac, D. Moher, M. D. Peters, T. Horsley, L. Weeks *et al.*, "PRISMA extension for scoping reviews (PRISMA-scr): checklist and explanation," *Annals of internal medicine*, vol. 169, no. 7, pp. 467–473, 2018.

[4] M. J. Page, J. E. McKenzie, P. M. Bossuyt, I. Boutron, T. C. Hoffmann, C. D. Mulrow, L. Shamseer, J. M. Tetzlaff, E. A. Akl, S. E. Brennan *et al.*, "The PRISMA 2020 statement: an updated guideline for reporting systematic reviews," *Systematic reviews*, vol. 10, no. 1, pp. 1–11, 2021.

[5] D. Moher, L. Shamseer, M. Clarke, D. Ghersi, A. Liberati, M. Petticrew, P. Shekelle, and L. A. Stewart, "Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-p) 2015 statement," *Systematic reviews*, vol. 4, no. 1, pp. 1–9, 2015.

[6] M. Friedewald, I. Schiering, N. Martin, and D. Hallinan, "Data protection impact assessments in practice: Experiences from case studies," *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, vol. 13106 LNCS, pp. 424–443, 2022.

[7] D. Papamartzivanos, S. A. Menesidou, P. Gouvas, and T. Giannetsos, "A perfect match: Converging and automating privacy and security impact assessment on-the-fly," *Future Internet*, vol. 13, no. 2, p. 30, 2021.

[8] S. Dashti, A. Santana de Oliveria, C. Kaplan, M. Dalcastagne, and S. Ranise, "Can data subject perception of privacy risks be useful in a data protection impact assessment?" in *Proceedings of the 18th International Conference on Security and Cryptography - SECRYPT,*, INSTICC. SciTePress, 2021, pp. 827–832.

[9] L. Sion, D. Van Landuyt, and W. Joosen, "The never-ending story: On the need for continuous privacy impact assessment," in *2020 IEEE European Symposium on Security and Privacy Workshops (EuroS&PW)*. IEEE, 2020, pp. 314–317.

[10] K. Vemou and M. Karyda, "An organizational scheme for privacy impact assessments," in *European, Mediterranean, and Middle Eastern Conference on Information Systems*. Springer, 2018, pp. 258–271.

[11] J. van Puijenbroek and J.-H. Hoepman, "Privacy impact assessments in practice: Outcome of a descriptive field research in the Netherlands," in *In Alamo, JM del (ed.), IWPE 2017: International Workshop on Privacy Engineering: Proceedings of the 3rd International Workshop on Privacy Engineering, co-located with 38th IEEE Symposium on Security and Privacy (S&P 2017)*, San Jose (CA), USA, 2017, pp. 1–8.

[12] D. Wright, K. Wadhwa, M. Lagazio, C. Raab, and E. Charikane, "Integrating privacy impact assessment in risk management," *International Data Privacy Law*, vol. 4, no. 2, pp. 155–170, 2014.

[13] D. Wright and K. Wadhwa, "Introducing a privacy impact assessment policy in the EU member states," *International Data Privacy Law*, vol. 3, no. 1, pp. 13–28, 2013.

[14] K. Wadhwa, "Privacy impact assessment reports: a report card," *info*, 2012.

[15] F. Thoma, "How Siemens assesses privacy impacts," in *Privacy Impact Assessment*. Springer, 2012, pp. 275–284.

[16] S. Deadman and A. Chandler, "Vodafone's approach to privacy impact assessments," in *Privacy Impact Assessment*. Springer, 2012, pp. 285–304.

[17] B. Stewart, "Privacy impact assessment: Optimising the regulator's role," in *Privacy Impact Assessment*. Springer, 2012, pp. 437–444.

[18] D. Wright and P. De Hert, *Privacy impact assessment*. Springer, 2012, vol. 6.

[19] N. Waters, "Privacy impact assessment – great potential not often realised," in *Privacy Impact Assessment*. Springer, 2012, pp. 149–160.

[20] A. Warren and A. Charlesworth, "Privacy impact assessment in the UK," in *Privacy Impact Assessment*. Springer, 2012, pp. 205–224.

[21] J. Edwards, "Privacy impact assessment in New Zealand – a practitioner's perspective," in *Privacy Impact Assessment*. Springer, 2012, pp. 187–204.

[22] R. M. Bayley and C. J. Bennett, "Privacy impact assessments in Canada," in *Privacy Impact Assessment*. Springer, 2012, pp. 161–185.

[23] D. Wright and P. D. Hert, "Findings and recommendations," in *Privacy Impact Assessment*. Springer, 2012, pp. 445–481.

[24] J. Stoddart, "Auditing privacy impact assessments: The Canadian experience," in *Privacy impact assessment*. Springer, 2012, pp. 419–436.

[25] C. Onwubiko, "Challenges to managing privacy impact assessment of personally identifiable data," in *Information Assurance and Security Technologies for Risk Assessment and Threat Management: Advances*. IGI Global, 2012, pp. 254–272.

[26] R. Parks, C. H. Chu, H. Xu, and L. Adams, "Understanding the drivers and outcomes of healthcare organizational privacy responses," in *32nd International Conference on Information System 2011, ICIS 2011*, 2011, pp. 245–264.

[27] R. Clarke *et al.*, "An evaluation of privacy impact assessment guidance documents," *International Data Privacy Law*, vol. 1, no. 2, pp. 111–120, 2011.

[28] D. Tancock, S. Pearson, and A. Charlesworth, "The emergence of privacy impact assessments," 2010, last accessed 29 September 2022. [Online]. Available: https://www.hpl.hp.com/techreports/2010/HPL-2010-63.pdf

[29] E. Brown and T. A. Kosa, "Incorporating privacy outcomes: Teaching an old dog new tricks," in *2008 Sixth Annual Conference on Privacy, Security and Trust*. IEEE, 2008, pp. 232–239.

[30] A. Warren, R. Bayley, C. Bennett, A. Charlesworth, R. Clarke, and C. Oppenheim, "Privacy impact assessments: International experience as a basis for UK guidance," *Computer Law & Security Review*, vol. 24, no. 3, pp. 233–242, 2008.

[31] M. D. Peters, C. Godfrey, P. McInerney, Z. Munn, A. C. Tricco, H. Khalil *et al.*, "Chapter 11: scoping reviews (2020 version)," *JBI manual for evidence synthesis, JBI*, vol. 2020, 2020.