Source Code-Driven GDPR Documentation: Supporting RoPA with Assessor View

Mugdha Khedkar*, Michael Schlichtig*, Eric Bodden*†
* *Heinz Nixdorf Institute, Paderborn University, Germany*
† *Fraunhofer IEM, Paderborn, Germany*
{mugdha.khedkar, michael.schlichtig, eric.bodden}@uni-paderborn.de

***Abstract*—The General Data Protection Regulation (GDPR) requires collaborative assessment of software to ensure lawful processing of personal data. Such collaboration, central to the Data Protection Impact Assessment (DPIA), involves diverse stakeholders but is hindered by limited tool support and manual documentation practices.**

In prior work, we introduced Assessor View, a tool that visualizes Android app behavior using a data privacy vocabulary and offers tailored views for legal, privacy, and technical experts. This paper extends Assessor View with new features to directly support stakeholders: a partially automated generation of the Record of Processing Activities (RoPA), and a collaborative dashboard to support communication between privacy experts and developers.

We evaluate these enhancements through semi-structured interviews with 6 privacy experts, demonstrating how the tool can better support GDPR-aligned documentation. Our findings indicate that Assessor View's RoPA feature provides a good overview of the processed data, and enhances understanding and interdisciplinary communication between stakeholders.

Demonstration video: https://youtu.be/rVwJDHLj7Pw.

Artifacts: https://doi.org/10.5281/zenodo.16151588.

***Index Terms*—data protection, privacy-aware reporting, usability, static program analysis.**

I. Introduction

Any software targeting the EU market, including Android apps, must comply with the *General Data Protection Regulation* (*GDPR*) [1]. The GDPR defines personal data as *"any information relating to a natural person"*, and regulates its access, storage, and processing. Non-compliance can lead to severe financial penalties [2].

Assessing GDPR compliance, particularly for software systems, requires collaboration among groups with a diverse expertise. Article §35 [3] of the GDPR describes the *Data Protection Impact Assessment* (*DPIA*) [4] as *"an assessment of the impact of the envisaged processing operations on the protection of personal data."* The data controller (e.g., software provider) is responsible for conducting the DPIA, with oversight from a designated *Data Protection Officer (DPO)*, typically a legal expert.

The DPIA involves a systematic analysis of software to identify and mitigate data protection risks, requiring collaboration between privacy experts, legal experts, and technical experts (such as app developers). DPIA documentation is created for and by multiple stakeholders, yet it must remain consistent with the source code. Despite its central role in GDPR compliance, two key **research gaps** remain: (1) the need to better understand and support collaboration and communication among diverse stakeholders involved in DPIA activities; and (2) the absence of tools that can accurately generate reporting documents directly from software artifacts.

Our prior work [5], [6] introduced Assessor View, a tool designed to address the **first research gap** by visualizing Android app code in terms of a data privacy vocabulary [7] and offering multiple stakeholder-specific views. However, it did *not* directly support the generation of GDPR-related reporting documents.

This paper extends Assessor View to address the **second research gap** by introducing new features to *directly support privacy experts, especially DPOs*. The enhanced Assessor View partially automates the generation of *Record of Processing Activities* (***RoPA***) [8]), a legally mandated inventory detailing how personal data is processed. It further provides a dashboard designed to facilitate communication between privacy experts and technical stakeholders.

We evaluate these enhancements through semi-structured interviews with 6 privacy experts. Our findings indicate that Assessor View's RoPA feature provides a good overview of the processed data, and has the potential to reduce manual workload for DPOs. It is a promising next step toward enhancing interdisciplinary communication between stakeholders.

The source code and all artifacts are available at https://doi.org/10.5281/zenodo.16151588, and the demonstration video is available at https://youtu.be/rVwJDHLj7Pw.

II. The Original Assessor View

The initial prototype of Assessor View comprised four key components: the User Interface (Module Ⓐ in Figure 1), the Privacy Slice Visualizer (Module Ⓑ), the DPV Transformer (Module Ⓒ), and the Views (Module Ⓓ). In this section, we describe only the original

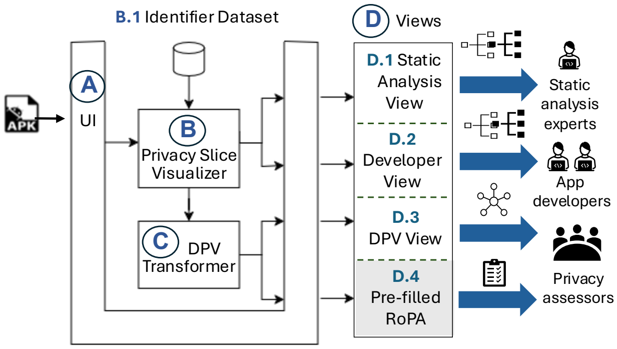


Fig. 1: Assessor View architecture. Colored box indicates the tool extension described in Section III.

prototype components (shown in white in Figure 1). The new extension D.4, highlighted in grey, is introduced later in Section III.

Initially, the *User Interface* Ⓐ accepts an APK, and forwards it to the *Privacy Slice Visualizer* Ⓑ for further analysis. The Privacy Slice Visualizer first identifies privacy-relevant data sources in the app's code using the *Identifier Dataset* (B.1) that we introduced in prior work [9]. It next uses Jicer [10], a static Android slicer, to extract program slices originating from these sources. These slices are visualized graphically to support inspection and interpretation by static analysis experts and developers. They are then passed on to the *DPV Transformer* Ⓒ, which converts them into graphical representations familiar to privacy assessors (legal experts, privacy experts, and DPOs).

After completing its analysis, Assessor View's user interface displays the results on a dashboard that provides a high-level, privacy-focused overview of risky sources detected in the APK, and the corresponding program slices. Users can explore these slices across the following views Ⓓ, each tailored to different stakeholder needs and levels of technical expertise:

- *Static Analysis View* (D.1): Presents the app code as graphs using Java intermediate representation, aimed at static analysis experts who need a detailed, low-level perspective.
- *Developer View* (D.2): Displays the slices in a source code-like format, making the information accessible to Android developers more familiar with Java-like representations.
- *DPV View* (D.3): Maps the code (from the Developer View) to concepts from the Data Privacy Vocabulary (DPV) [7], offering a high-level summary suited for privacy assessors. This view presents the detected processing operations through multiple complementary representations, including a DPV diagram (cf. Figure 2) and corresponding textual explanations. This view also flags potential GDPR violations and provides relevant warnings, suggestions, and legal recommendations.

A. Limitations of the Original Prototype

To assess the viability of the original Assessor View, we conducted an interview-based user study with 16 privacy experts, focusing on its potential use in privacy assessments (including DPIAs [4]) and GDPR compliance. This study, along with accompanying correctness checks of Assessor View's output, is currently under submission [6].

In the study, most participants reported relying heavily on manual effort to prepare GDPR-related documents. Ten of the sixteen participants mentioned using standard templates and spreadsheets to complete GDPR documentation, though one noted the limitations of generic templates: *"We can't just give them (clients) a template to begin with, right? It has to be customized to what they do."* Fourteen participants identified Record of Processing Activities (RoPA) [8] as the most important document required for GDPR compliance, often used as starting point for DPIAs.

Article §30 [11] of the GDPR requires organizations to document how personal data is collected, stored, and processed, including details such as processing purposes, data categories, recipients, and retention periods. This documentation, known as the **Record of Processing Activities (RoPA)**, serves as both a compliance record and a foundation for conducting DPIAs. As a legal researcher explained: *"Completing the RoPA is a data mapping exercise which allows stakeholders to envisage potential threats."* Participants recommended aligning the tool's output with the structure of a RoPA. For example, one suggested: *"(The tool can) help fill RoPA using static analysis. If you get data in a RoPA template, DPOs would be very happy."*

To summarize, while the initial prototype of Assessor View addressed the **first research gap** by providing tailored visualizations for different stakeholders, it did *not* address the **second research gap**—namely, the need to *accurately generate GDPR reporting documents*. This limitation and the direct feedback from participants motivated the design of the new RoPA extension (D.4 in Figure 1), which we introduce in the following section.

III. Extended Assessor View

Building on feedback from the previous study (cf. Section II-A), we identified and prioritized improvement options focused on supporting RoPA documentation. This resulted in an enhanced version of Assessor View, featuring a *Pre-filled RoPA* module (D.4 in Figure 1) that extends the existing DPV View to automatically generate a partially completed Record of Processing Activities

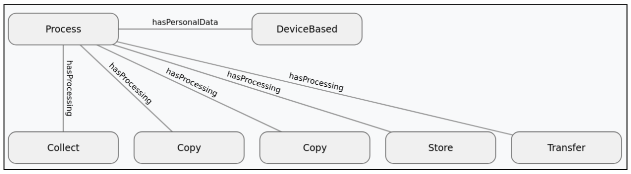
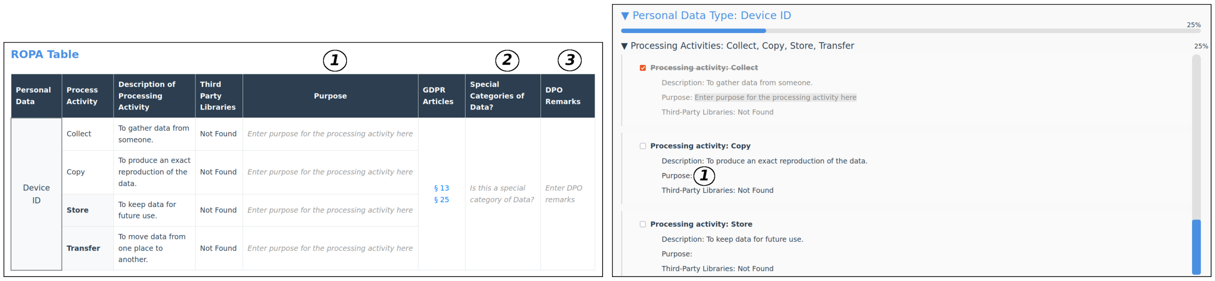


Fig. 2: DPV diagram for the demo app [12], summarizing the collected data and processing operations using DPV constructs such as *hasPersonalData* and *hasProcessing*.



ROPA Table

Personal Data	Process Activity	Description of Processing Activity	Third Party Libraries	Purpose ①	GDPR Articles	Special Categories of Data? ②	DPO Remarks ③
Device ID	Collect	To gather data from someone.	Not Found	*Enter purpose for the processing activity here*	§ 13 § 25	*Is this a special category of Data?*	*Enter DPO remarks*
	Copy	To produce an exact reproduction of the data.	Not Found	*Enter purpose for the processing activity here*			
	Store	To keep data for future use.	Not Found	*Enter purpose for the processing activity here*			
	Transfer	To move data from one place to another.	Not Found	*Enter purpose for the processing activity here*			

▼ Personal Data Type: Device ID

25%

▼ Processing Activities: Collect, Copy, Store, Transfer

25%

☑ ~~**Processing activity: Collect**~~

Description: To gather data from someone.

Purpose: Enter purpose for the processing activity here

Third-Party Libraries: Not Found

□ **Processing activity: Copy**

Description: To produce an exact reproduction of the data.

Purpose: ①

Third-Party Libraries: Not Found

□ **Processing activity: Store**

Description: To keep data for future use.

Purpose:

Third-Party Libraries: Not Found

(a) A part of the RoPA table.

(b) A part of the RoPA form.

Fig. 3: RoPA formats for the demo app [12], showing editable fields for ① the purpose of processing, ② special categories of data, and ③ DPO remarks (present in both formats but partially cropped in Figure 3b).

(RoPA) [8]. The module retains the *distinct* processing operations identified in the DPV diagram (cf. Figure 2) and represents them in a structured, editable RoPA format tailored to the needs of DPOs. Unlike generic templates, this customizable RoPA is automatically tailored to the structure of the analyzed APK, reflecting practices described by several participants who currently rely on manual templates.

Through the dashboard, users can now select the data sources they wish to document. For each selected data source, the generated RoPA identifies associated processing operations, with descriptions drawn from the DPV [7]. Assessor View supports two interchangeable RoPA formats: a table (cf. Figure 3a), and a form (cf. Figure 3b). Both include editable fields where DPOs can specify the business purpose of processing (① in Figure 3a), indicate whether special category data is involved (② in Figure 3a), and add custom remarks (③ in Figure 3a). DPOs can switch between views, edit content, and export the output as a printable PDF—a feature one participant highlighted as useful for communication with supervisory authorities. The form view also allows users to check off completed sections and monitor progress via a progress bar (cf. Figure 3b).

We also extended the dashboard to better support collaboration. We introduced two warning scales: one for developers, focusing on technical issues (e.g., third-party data sharing or missing encryption), and one for privacy experts and DPOs, highlighting GDPR-specific risks (e.g., lack of data minimization or missing safeguards), derived from the Developer View and DPV View, respectively. We also added assessment flags for each program slice to help stakeholders track review status (e.g., “Addressed”, “Needs review”, or “Flagged”). To support informed decision-making, we included documentation explaining the warning categories, GDPR articles, DPV concepts, and potential technical measures.

To evaluate these new features, we conducted a follow-up study with 6 privacy experts. The findings from this study are presented in the following section.

IV. Interview-Based Study

We now discuss the study we conducted with the enhanced version of Assessor View to answer the following research question: *To what extent does Assessor View support privacy experts in completing the Record of Processing Activities?*

A. Experimental Setup

To answer the research question, we conducted semi-structured interviews with 6 participants. The study design was approved by the Ethics committee of our institution. Our consent form and privacy policy were approved by the Data Protection Office of our institution.

Interview Recruitment. We contacted 6 privacy experts from the initial expert pool of 16 participants (cf. Section II-A), particularly DPOs based in the EU and UK. All of them accepted our invitation to participate.

Interview Design. At the start of each interview, we gave participants all the information about the study, and they signed the *informed consent* form. We began by asking about their current practices for completing

TABLE I: Participant experience summary. Filled circles indicate expertise from 1 (Beginner) to 5 (Expert).

ID	Job title	Country	Exp in privacy/ law (yrs)	Familiarity				
				Data Visualization	GDPR	GDPR Compliance	DPIA	DPV
P01	Corporate DPO	Germany	9	●●●●○	●●●●●	●●●●●	●●●●○	●●○○○
P02	DPO	Germany	12	●○○○○	●●●●●	●●●●●	●●●●●	●●●●○
P03	DPO	Cyprus	7	●●●○○	●●●●●	●●●●●	●●●●●	●●●●●
P04	Data Protection Advisor	Germany	3.5	●●●○○	●●●●●	●●●●○	●●●○○	●●●●●
P05	External DPO	UK	15	●●●○○	●●●●●	●●●●●	●●●●●	●●●●○
P06	Privacy Counsel	Germany	8	●●●○○	●●●●○	●●●●○	●●●●○	●○○○○

the RoPA. Next, participants used Assessor View to automatically generate a RoPA for a given demo app [12] with documented data flows. Finally, we asked questions about their experience with the tool—specifically, the usability, clarity, and reliability of the generated RoPA—and their impressions of the two RoPA formats (table and form) provided by Assessor View.

Interview Process. We used Zoom for all interviews. All interviews were conducted by the first author in July 2025, and audio-recorded via Zoom's functionality.

Data Analysis. We recorded and transcribed all interviews using NoScribe [13], then reviewed transcripts for errors. The first author used iterative open coding [14] to develop a codebook, and the second author double-checked the codes.

B. Results

We refer to our participants as P01–P06.

Participant Experience. Participants self-reported familiarity on a 5-point scale (1 = Beginner, 5 = Expert), reporting median scores of 5 for GDPR, GDPR compliance; 4.5 for DPIA; 4 for DPV; and 3 for data visualization (cf. Table I). They have a median of 8.5 years of experience in privacy and/or law, with all holding additional privacy certifications. P02 holds a doctoral degree in Law.

Existing support for RoPA. All participants reported using standard templates for RoPA, including *checklists* (P02, P04) and *questionnaires* (P03). While many (P01, P02, P04, and P06) used *stakeholder interviews* to gather necessary information, P01 and P04 also relied on a *questionnaire-style tool* to guide the RoPA process.

Participants identified several challenges in completing a RoPA, such as *balancing abstraction layers* (P01), *time consuming manual process* (P02), and *difficulties in collecting necessary information* (P04).

Assessor View-generated RoPA. All participants considered the generated RoPA a *useful starting point* for documenting activities. Several (P01, P03, and P05) found it *easy to understand*, and many (P03, P04, P05, and P06) felt it could *reduce manual workload* during DPIAs or pre-DPIA screening. P01 said, *"In theory, it should start at this level. In practice, this is not possible because this granularity is missing. This would be the first representation I've seen of a concept that actually does the proper theory on designing processing activities."* P03 noted, *"Maybe it could provide a good understanding of how and where data has been used, as opposed to just having to trust whatever the developers tell you."* P04 and P06 emphasized its value in preparing for stakeholder interviews: *"Because you already have a lot of pre-information, we can make better questions for the colleagues"* (P04). P05 emphasized its usefulness for external DPOs without source code access, adding, *"The bigger plus to me would be that it reduces the chance of something being missed"*.

RoPA Formats. All but one participant (P04) preferred the table format (cf. Figure 3a) for documenting activities. P03 explained, *"It (table) provides a good view of what data you have in your codebase, and you can see everything at the same time."* Most also favored the table for developer discussions, though P01 found the form view (cf. Figure 3a) more suitable for interviews. P05 observed, *"I think it's good to allow both options, because different people would work in different ways"*.

Other feedback. Participants suggested minor refinements to terminology and structure (P01, P04), including clarifying legal bases for data collection such as contracts or user consent (P02). They also proposed a future enhancement: including risk analysis to automatically determine whether a full DPIA is required (P04, P06).

Finding. The RoPA feature in Assessor View shows promise for improving interdisciplinary communication among stakeholders. It provides a good overview of the processed data, and has the potential to reduce manual workload of DPOs.

V. Related Work

Existing open-source static analysis tools [15]–[17] assist developers in verifying whether application source code and privacy policies comply with data protection requirements. However, these tools are developer-centric and lack support for the broader set of stakeholders involved in compliance processes.

Proprietary solutions such as Privado [18] automate privacy compliance checks through local source code and data-flow analysis. In contrast, Assessor View operates on compiled APKs, enabling external auditing when source code is unavailable. While such tools demonstrate the feasibility of automated privacy auditing, they remain closed-source and focus primarily on developer-facing reports, offering little insight into how privacy assessors actually interpret or use such outputs in practice. Similarly, other proprietary platforms such as SonarQube [19] focus exclusively on software quality and security, lacking the support for GDPR-specific analyses or RoPA documentation generation. Assessor View addresses this gap by integrating source code analysis with assessor-oriented RoPA generation, evaluated with DPOs.

In contrast, commercial platforms such as OneTrust [20] focus on compliance documentation management but do not analyze source code.

As a result, *there remains no open-source tool that bridges source-code analysis with automated RoPA documentation*, leaving assessors without adequate technical support. Assessor View fills this gap by enabling code-aware RoPA generation.

VI. Threats to Validity

While our interview-based study is qualitative, we mitigated generalizability concerns by recruiting a diverse participant pool through direct outreach to over 175 experts across 43 countries. A broader quantitative evaluation remains a valuable direction for future work. Prior research suggests that as few as six interviews can be sufficient when the participant group is relatively homogeneous and the study objectives are focused [21].

VII. Conclusion

In this paper, we have extended Assessor View to partially automate RoPA generation, and improve collaboration between different stakeholders. We have conducted an interview-based study with 6 privacy experts to evaluate these enhancements.

Our findings suggest that Assessor View's RoPA feature provides a good overview of the processed data, and has the potential to reduce manual workload for DPOs. It is a promising next step toward enhancing interdisciplinary communication between stakeholders involved in DPIA activities.

References

[1] "The european parliament and the council of the european union. general data protection regulation (gdpr)," 2018. [Online]. Available: https://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679

[2] "Gdpr penalties." 2018. [Online]. Available: https://gdpr-info.eu/issues/fines-penalties/

[3] "Gdpr article 35," 2018. [Online]. Available: https://gdpr-info.eu/art-35-gdpr/

[4] "Information commissioner's office. 2018. data protection impact assessments (dpias)." 2018. [Online]. Available: https://ico.org.uk/for-organisations/guide-to-data-protection/guide-to-the-general-data-protection-regulation-gdpr/data-protection-impact-assessments-dpias/

[5] M. Khedkar, M. Schlichtig, and E. Bodden, "Advancing android privacy assessments with automation," in *Proceedings of the 39th IEEE/ACM International Conference on Automated Software Engineering Workshops*, ser. ASEW '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 218–222. [Online]. Available: https://doi.org/10.1145/3691621.3694953

[6] (2025) Assessor view: Introducing tool support for android privacy assessments. [Online]. Available: https://doi.org/10.21203/rs.3.rs-6323701/v1

[7] "Data privacy vocabulary." 2025. [Online]. Available: https://w3c.github.io/dpv/2.1/dpv/

[8] (2018) Records of processing activities. [Online]. Available: https://gdpr-info.eu/issues/records-of-processing-activities/

[9] M. Khedkar, A. K. Mondal, and E. Bodden, "Do android app developers accurately report collection of privacy-related data?" in *Proceedings of the 39th IEEE/ACM International Conference on Automated Software Engineering Workshops*, ser. ASEW '24. New York, NY, USA: Association for Computing Machinery, 2024, p. 176–186. [Online]. Available: https://doi.org/10.1145/3691621.3694949

[10] F. Pauck and H. Wehrheim, "Jicer: Simplifying cooperative android app analysis tasks," in *2021 IEEE 21st International Working Conference on Source Code Analysis and Manipulation (SCAM)*, 2021, pp. 187–197.

[11] "Gdpr article 30," 2018. [Online]. Available: https://gdpr-info.eu/art-30-gdpr/

[12] "Demo application from taintbench," 2025. [Online]. Available: https://github.com/TaintBench/godwon_samp

[13] "Noscribe," 2025. [Online]. Available: https://github.com/kaixxx/noScribe

[14] J. Corbin and A. Strauss, "Grounded theory research: Procedures, canons and evaluative criteria," *Zeitschrift für Soziologie*, vol. 19, no. 6, pp. 418–427, 1990. [Online]. Available: https://doi.org/10.1515/zfsoz-1990-0602

[15] S. Arzt, S. Rasthofer, C. Fritz, E. Bodden, A. Bartel, J. Klein, Y. Le Traon, D. Octeau, and P. McDaniel, "Flowdroid: Precise context, flow, field, object-sensitive and lifecycle-aware taint analysis for android apps," in *Proceedings of the 35th ACM SIGPLAN Conference on Programming Language Design and Implementation*, ser. PLDI '14. New York, NY, USA: Association for Computing Machinery, 2014, p. 259–269. [Online]. Available: https://doi.org/10.1145/2594291.2594299

[16] Z. Tan and W. Song, "Ptpdroid: Detecting violated user privacy disclosures to third-parties of android apps," in *2023 IEEE/ACM 45th International Conference on Software Engineering (ICSE)*, 2023, pp. 473–485.

[17] K. Zhao, X. Zhan, L. Yu, S. Zhou, H. Zhou, X. Luo, H. Wang, and Y. Liu, "Demystifying privacy policy of third-party libraries in mobile apps," 2023. [Online]. Available: https://arxiv.org/abs/2301.12348

[18] "Privado.ai," 2022. [Online]. Available: https://www.privado.ai/data-safety-report

[19] (2025) Sonarqube. [Online]. Available: https://www.sonarsource.com/sem/products/sonarqube/

[20] (2025) Onetrust. [Online]. Available: https://www.onetrust.com/

[21] G. Guest, A. Bunce, and L. Johnson, "How many interviews are enough?: An experiment with data saturation and variability," *Field Methods*, vol. 18, no. 1, pp. 59–82, 2006. [Online]. Available: https://doi.org/10.1177/1525822X05279903