https://doi.org/10.2352/EI.2025.37.3.MOBMU-328


# PrivacyBuddy: An Android Privacy Dashboard for Detecting Excessive Data Collection with a Focus on Location Data

***Toon Dehaene*, Maxime Bellis, Tristan Pelgrims, Vincent Naessens*, Bert Lagaisse*; KU Leuven (DistriNet*); Leuven and Ghent, Belgium; firstname.lastname@kuleuven.be***

## Abstract

*This paper presents* ***PrivacyBuddy****, an innovative Android privacy dashboard designed to detect excessive data collection, focusing on location data. It addresses the growing concern over apps that collect more personal information than necessary, often leading to privacy violations. Current Android privacy dashboards are inadequate, lacking effective visualizations and insights into app behavior. PrivacyBuddy enhances user awareness by providing* intuitive visualizations *of data collection frequency, volume, and type, along with a* Privacy Score *for each app. The dashboard employs user-centered design principles to cater to diverse user needs, from casual users to privacy-conscious individuals. Key features include detailed tracking of location requests, distinguishing between foreground and background access, and offering actionable insights for users to manage their privacy settings effectively. A* secure implementation *extends the Android OS to monitor data requests made by other apps. A user study validates the design, demonstrating improved usability and user engagement compared to existing solutions. This work contributes to the field of visual data analytics by merging design principles with practical applications, ultimately empowering users to take control of their data privacy in an increasingly data-driven world.*

## Introduction

Excessive tracking of users has real-life impact, whether they consider personal privacy to be important or not. Data collection can be advantageous for tailoring individualized experiences and services. A health app tracks your activity to suggest diet and workout plans. A photo gallery app analyzes your pictures to automatically categorize family pictures, notes, screenshots, and other types. However, data collectors frequently amass information that exceeds what is necessary, with an extent and frequency that surpasses the requirements for improving service. The business plan of many such data collectors includes the selling of this data which enables misuse by other parties [9]. In one case, the FTC charged the developers of a flashlight app for collecting users' location data [6], which they naturally didn't need. Even when the necessity is apparent, apps collect data with higher accuracy or frequency than necessary. Although almost all apps benefit from having access to the clipboard, many apps have been caught procedurally scraping the user's clipboard for tracking purposes [2]. It is therefore important not only to analyze whether apps have the ability to collect data, but also to what extent they do. Many more examples exist, and more importantly, little effective systems are in place to prevent this kind of misuse. Awareness about which privacy-sensitive data is frequently collected by apps on a smart phone is the first line of defense against excessive data collection and tracking of users.

When apps track users, they use either data on the device or sensors to collect personal data. In most systems, the access to sensors and the file system must pass through the operating system. System designers provide tools to deny data requests to certain collectors and apps. In the Android operating system, users have to grant permissions to apps. A study [4] shows that users are indeed more likely to deny permissions that don't make sense for a given app, e.g. a flashlight app asking for location permission. However, even when the permission makes sense at installation time, misuse is possible when apps request a lot of data frequently.

*A key problem is that users lack a good insight on which app collects what data at which frequency while executing, hence after installation and permission approval of the app*. We demonstrate in this paper that the current privacy dashboards in Android and its open-source variants do not meet the users' needs.

As a solution, we propose PrivacyBuddy, which provides **trustworthy visualization and data analysis** of runtime data collection by apps that can be of assistance to many types of users. We identify the following contributions of our overall technical solution, related to visual design, data analysis and trustworthy deployment of the solution on Android phones.

- We determine which key data collection properties are required to visualize tracking behavior, including the frequency, volume and type of data collection as a key property to express excessive tracking.
- We calculate and provide a privacy score per app based on data analysis of the key data collection properties and their priority and importance related to privacy-awareness for the user.
- We design and evaluate intuitive and effective interfaces based on key design frameworks such as User-Centered Design, the C-HIP model and Gestalt principles. We focus on high priority visualizations first and use detail-on-demand to cater to a wide-ranging customer base, from laypeople to individuals with advanced technical expertise.
- We implement and deploy a trustworthy data flow architecture for the different components of our solution leveraging the internal Android security and permission system.

This work is unique in the visualization community because it blends privacy engineering insights and design principles to create visual tools that aid privacy awareness and enable users to take action to protect their privacy. Upon examining the visualizations, individuals can more readily take action against intrusive apps by leveraging existing tools such as restricting data access, blocking access, or completely uninstalling them. Furthermore, this work contributes a trustworthy and secure implementation that is

built into the Android system. We also provide the source code as open-source for easy extensibility.

This work provides visualizations for data types that can be requested by an Android app. The modular dashboard allows for extensions, where other system developers and app developers can contribute widgets. For this proof-of-concept, we validate our approach with a case study on location data for detailed analysis. The reasons why this work limits the scope to location data are as follows:

- We do not have enough resources to produce custom visualizations for all data types shared and used on mobile devices.
- Mobile device users consider location data to be one of the most sensitive types of information in the category of types that do not directly identify them. [16]
- Location data is one of the most tracked data types on mobile applications. [12]
- Focusing on location lends itself to unique visualizations that greatly improve efficacy. This provides strong argument that custom visualizations of tracking behavior can be useful.

With the goal of visualizing excessive tracking in mind, users can compare the frequency of location requests between apps. The dashboard also visualizes whether a location request was done by an app in the foreground or background, and whether it was done with high or low location accuracy. We focus on location data as it is often over-requested, and it is typically sensitive data.

In the background section, we cite related work showing that existing tools inadequately support visualizing excessive data collection. The main shortcomings are the poor choice of visualizations, such as using pie charts to show the share of data collection per permission type, and the lack of visualizations that allow users to compare app tracking behavior. To avoid similar pitfalls, we use established design frameworks and lessons from prior publications. Additionally, we propose a modular design with customizable widgets, enabling users to prioritize the visualization of specific data (e.g., location) and select the desired granularity, ensuring compatibility for users with varying technical skills.

In our evaluation we assess the usability and compliance of the dashboards with the requirements by carrying out a study using simulated user data. In this study, we interview individuals with varying levels of technological proficiency to determine if there has been improvement compared to current data collection visualizations. The study is conducted in two phases, with feedback from the initial phase used to refine the design.

The remainder of this paper is structured as follows. We first discuss background and motivation, including information on privacy, related work, as well as motivating examples and design frameworks. We then expand on shortcomings of existing solutions. Our requirements analysis elaborates on different personas and their different needs for privacy awareness. We then detail the design of our Privacy Dashboard and focus on visual design and design principles, the privacy score and the internal system design for Android. We conclude with an evaluation based on the results of our user study and discuss our improvements and limitations.

## Background and Motivation

This section first defines what is considered as excessive tracking of personal data. Then, it discusses some relevant existing solutions, as well as scientific works that have contributed to visualizing privacy dashboards. Lastly, it introduces the employed design frameworks and how they are used for shaping the visualizations in the dashboard this work presents.

### What is excessive tracking?

Personal data encompasses a wide range of information that can be used to identify an individual, either directly or indirectly. This includes, but is not limited to, location data, browsing history, contact information, and health records. The collection and use of such data has become ubiquitous in the digital age, driven by the potential benefits of personalized services and targeted advertising. However, the collection and use of personal data can become excessive when it exceeds what is necessary to provide the intended service or functionality. This is why one of the main principles of GDPR is the principle of data minimization [5]. It mandates that data collectors should only gather the minimum amount of information required to fulfill their objectives. Excessive data collection can lead to privacy violations, discrimination, and other harms, even if the individual does not consider their privacy to be a priority [11].

### Existing privacy dashboards

In this work, we focus on the Android operating system, as it is one of the most widely used mobile platforms globally. Android's open-source nature allows for customization and the development of alternative versions, providing opportunities to address privacy concerns that may not be adequately addressed in the standard Android distribution. The Android ecosystem has implemented several privacy-focused features, including permission systems and data access controls. These features aim to give users more control over their data. Additionally, Android offers visualizations to help users understand and manage their privacy settings better. These visualizations provide a clearer and more concise understanding of the data that apps are accessing, enabling users to make informed decisions about their privacy preferences. However, the effectiveness of these features in helping users manage their data privacy is still a topic of research and debate [4, 7]. To address some of the shortcomings, certain open-source variants of Android have included additional features. CalyxOS, /e/OS, and GrapheneOS [3, 8, 1], for instance, limit or exclude Google Play Services from their OS, while /e/OS and LineageOS [13] provide additional data control features. One such feature allows /e/OS users to share inaccurate or fake locations when apps request them. Despite these improvements, there is still room for growth in the area of visualizations for privacy management. We believe that significant advancements can be made in this area, and elaborate in the section *Shortcomings of existing solutions*.

### Research on privacy dashboards

Prior research has explored different approaches to visualizing and managing personal data collection on mobile devices. Wilkinson et al. [18] explored how smartphone users perceive the structure and granularity of privacy visualizations. They compared two types of privacy visualizations: data-centric and app-centric. Data-centric designs emphasize the data itself, prioritizing the user's awareness of how their personal information is collected, used, and shared across apps. This approach highlights the specific data being shared, regardless of which application is re-

questing it, thus enabling users to make privacy decisions based on the sensitivity of the data itself. In contrast, app-centric designs provide a more application-specific view, allowing users to focus on the behaviors of individual applications and understand which apps are responsible for sharing personal data. The research found that both app-centric and data-centric approaches can be valuable, as users were evenly divided, with a 50/50 preference for each. The concept of Glanceability [14] aims to make important information easily accessible and understandable at a glance. Finding a balance between glanceability and detailed information is key to raising user awareness. The study highlights the importance of using unobtrusive but pervasive visualizations to increase users' understanding of real-time data-sharing practices by mobile apps. By varying the granularity of the visualizations, they found that users' preferences for data- or app-centric designs were shaped by their perceptions of privacy boundaries. Those more concerned about specific data types favored data-centric designs, while users who were focused on the behavior of individual apps preferred app-centric visualizations.

### *Relevant design frameworks*

In this work, we draw upon established design frameworks to guide the development of our privacy dashboard.

**User-Centered Design** (UCD) [15] principles ensure that the system is tailored to the needs and preferences of the target users. UCD emphasizes the importance of understanding users' mental models and the context in which they interact with systems. This approach involves iterative design processes that incorporate user feedback at every stage, ensuring that the final product is both usable and effective. Key aspects of UCD include the need for systems to support users' cognitive processes and the importance of designing interfaces that facilitate direct engagement. By focusing on the user's perspective, designers can create systems that are not only functional but also intuitive and enjoyable to use. This is achieved through techniques such as rapid prototyping, usability testing, and the incorporation of adaptive mechanisms that help users recover from errors gracefully.

**The C-HIP model** [19], which stands for Communication-Human Information Processing, provides a framework for designing effective visualizations. This model outlines the stages of information processing, from the source of the information to the behavior of the receiver. It emphasizes the importance of attention, comprehension, and memory in ensuring that users effectively process and act upon the information presented to them. The C-HIP model includes stages such as attention switch, attention maintenance, and comprehension, which are critical for designing visualizations that capture and retain users' attention. By ensuring that visualizations are salient and easily understandable, designers can improve users' ability to process complex information. This is particularly important in the context of a privacy dashboard, where users need to quickly grasp and act on privacy-related information.

**Gestalt principles** for visual design [10] can enhance a privacy dashboard by improving user comprehension and engagement. Principles such as proximity and similarity help group related information together, making it easier for users to compare the level of tracking of different apps. Moreover, the principle of closure can guide users' eyes to complete visual elements, ensuring they don't miss important information.

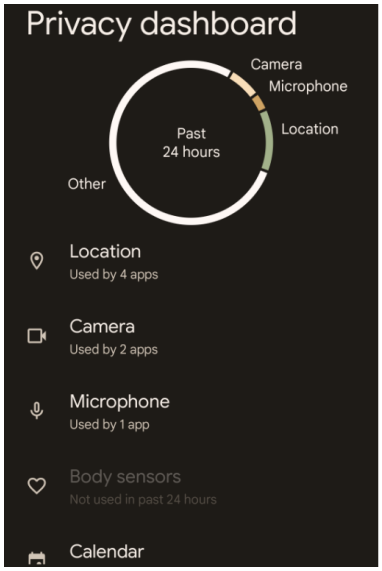



***Figure 1.*** *The dashboard shipped with Android has many flaws. The pie chart is a bad choice for showing usage of permissions. The 'other' section dominates. There is limited detail on demand.*

## Shortcomings of existing solutions

This section continues the state-of-the-art analysis by examining an existing privacy dashboard. This dashboard, the default in the open-source version of Android, is also used in the Pixel version sold by Google, ZenUI by Asus, ColorOS by OPPO and many more. Even in privacy-focused Android variants like CalyxOS, LineageOS, e/OS, etc. this dashboard remains largely unchanged. The consistency across these versions highlights the broad application of this version of this privacy dashboard and thus relevance of this analysis. Figure 1 illustrates the privacy dashboard from the Pixel Android version. The main difference between it and the other variants is the color palette.

### *The pie chart*

One of the primary issues with the dashboard is the use of a *pie chart* to represent data access distribution. This choice, while aiming for simplicity, **introduces ambiguity and complexity for end-users**. The pie chart of the stock dashboard can be seen on figure 1.

Many data categories are hidden under a "Show other permissions" button, complicating user understanding. Grouping many permissions under "Other" becomes problematic when this category is the largest, especially when hidden behind an extra button press. By hiding many data types initially, the dashboard creates an implicit hierarchy of data type permissions, potentially leading users to overlook important ones. The pie chart also lacks clear, glanceable symbols, making it difficult for users to understand what each slice represents. Aggregating unrelated permissions under vague labels violates the Gestalt law of similarity, misleading users about the similarity of data types. The attempt to simplify through a pie chart obscures vital information, reducing the dashboard's effectiveness.

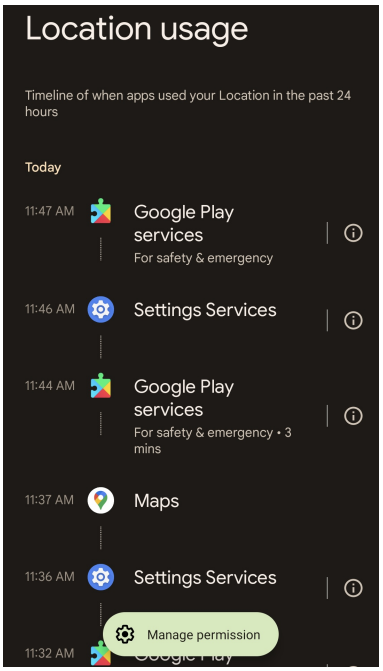



***Figure 2.*** *The location usage view on Android is disorganized*

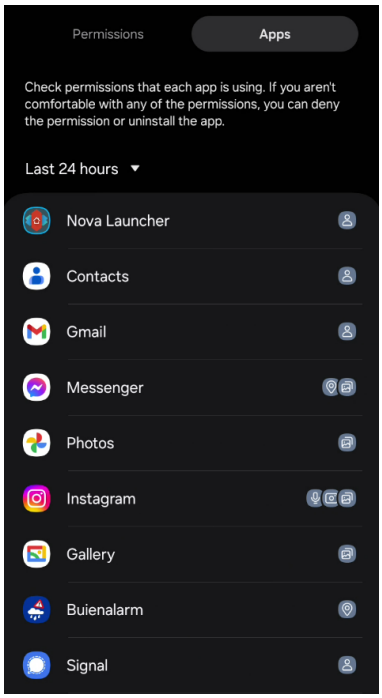



***Figure 3.*** *The only app-centric view on Android can be found in the permission manager. The goal is to inspect permissions per app.*

### *The short time window*

Both figure 1 and 2 are based on the last 24 hours only. The dashboard's 24-hour window is too short to provide meaningful insights into data access patterns. For example, a social media app might access precise location data to serve targeted ads based on your movements over several days. If only a single day's data is reviewed, this ongoing and potentially intrusive access might go unnoticed. Certain expected indicators, like microphone usage for video recording, are absent, raising questions about the dashboard's accuracy. System accesses are initially hidden, requiring extra steps to view, which might lead users to underestimate data access volumes.

### *Timeline*

The timeline view for specific permissions lacks depth and context. As can be seen in figure 2, users are left with many unanswered questions about the nature and implications of data accesses. The dashboard fails to clarify what specific location was accessed. There is no differentiation between precise and approximate location accesses, crucial for understanding privacy impact. Users are also left guessing about the accuracy of accessed location data and whether additional information like speed and bearing was included. Currently, the only app-specific information visible is the list of permissions granted to each app, demonstrated in figure 3.

We believe that offering both app-centric and data-centric visualizations, along with more customized options like displaying location access on a map, can improve the transparency of data collection and make it easier to identify overly intrusive apps.

## Requirements Analysis

This section first identifies users and describes a set of personas, following the design methodology of UCD. Employing the insights from the user identification, it then outlines the key requirements.

### *User Identification*

The users of this system are essentially everyone who uses a smartphone. To better understand their needs, we categorize users based on two main dimensions: technical knowledge and interest in privacy. This results in four distinct user groups.

- Low technical knowledge, low interest in privacy.
- Low technical knowledge, high interest in privacy.
- High technical knowledge, low interest in privacy.
- High technical knowledge, high interest in privacy.

These combinations map to three levels of data granularity, as users with high technical knowledge and low interest in privacy as well as those with low technical knowledge and high interest in privacy require similar levels of data granularity. According to the User-Centered Design framework, we model three personas to guide the design process.

### *Personas*

The following personas represent the different identified user groups and their respective needs and behaviors.

#### *Indifferent*

The indifferent persona shows little concern for privacy and is not inclined to invest time in understanding data privacy details. This persona is expected to engage minimally with the system, primarily using the glanceable dashboard and occasionally clicking through to secondary pages. The design for this persona is simple, providing a lot of information quickly with minimal interaction. The goal is to make the experience effortless and to deliver the most information as fast as possible.

The design for this persona should focus on simplifying privacy information extensively, requiring minimal to no effort to grasp it. All information has to be presented in a very glanceable way.

#### *Concerned*

The concerned persona is aware of privacy issues and uncomfortable with data tracking but feels powerless to make changes. This persona is inherently slightly demotivated. It is expected that the concerned persona will use the privacy dashboard, click through to secondary pages, and interact with the available options. The design focuses on motivating this persona to make use of existing privacy controls. The main goal is for this persona to see the impact of changing their privacy settings quickly and easily, with more granular options readily accessible.

The design should empower this persona with clear, educational visualizations that explain data privacy implications and offer actionable insights. The dashboard's primary objective is to show the impact of changing app permissions, using a glanceable design to lead them to higher granularity without overwhelming them.

#### *Engaged*

The engaged persona actively seeks comprehensive information and is willing to get down into privacy details. This persona is expected to quickly review basic information and then engage with more advanced privacy settings. The design for this persona focuses on providing detailed information and facilitating deep engagement. This persona is likely to bypass the initial dashboard and start interacting with secondary pages and specific access settings to obtain the maximum amount of information.

The design should provide this persona with detailed and customizable visualizations that allow for an in-depth understanding of how their data is handled. This persona requires the most granularity with much less focus on the glanceability of designs.

### Requirements

Using the insights from the user analysis and personas, we outline six usability and design requirements.

- R1 Users can get insights on whether an app is intrusive on their privacy.
- R2 Users get more insight into the details of app permissions and their inner workings such as foreground or background access.
- R3 The user is not overwhelmed with detail, but can find it if desired.
- R4 The intuitive design of the interface enables users to easily access detail-on-demand through clickable and zoomable elements.
- R5 Users know what action they can take against these intrusive apps.
- R6 The dashboard encourages users to regularly monitor and adjust their app permissions.

## Design of the Privacy Dashboard

This section outlines the design of the Privacy Dashboard. It starts with an overview of the visual components. First we elaborate on the design process. We describe how we addressed the requirements listed above, adopted a user-centered design, and refined our solution based on user feedback. Secondly, we elaborate on the logic behind the Privacy Score that is used to rank apps by intrusiveness on location privacy. Lastly, an overview of the system design and deployment architecture of the software implementation is provided.

### Overview

The general design philosophy for the dashboard is to show basic visualizations with hints to more detail ready to be displayed. This should require minimal interaction for a curious user. This detail-on-demand approach makes sure initial views are glanceable, with more granular views provided when required. This section introduces the overview of the different visual components used in the design. Figure 4 shows the high level flow, starting from the DASHBOARD OVERVIEW PAGE [4A], shown in **4A**. The design process led to the creation of three foundational widgets for the privacy dashboard, with potential for expansion. The MAP WIDGET [4A1] graphically represents location data points for all apps. The APP-CENTRIC WIDGET [4A2] shows which apps did the last three location requests. The DATA-CENTRIC WIDGET [4A3] shows the apps which have the worst impact on location privacy according to their Privacy Score.

The three subfigures in the middle of figure 4 are the MAP PAGE [4B], TIMELINE PAGE [4C] and APP-RANKING PAGE [4D] respectively. They are the next level of granularity, showing more detail for more curious users.

Finally, the SPECIFIC ACCESS PAGE [4E] shows the most detail and consolidates all available data regarding a specific privacy access, including time, location, and additional data sent to the requesting application.

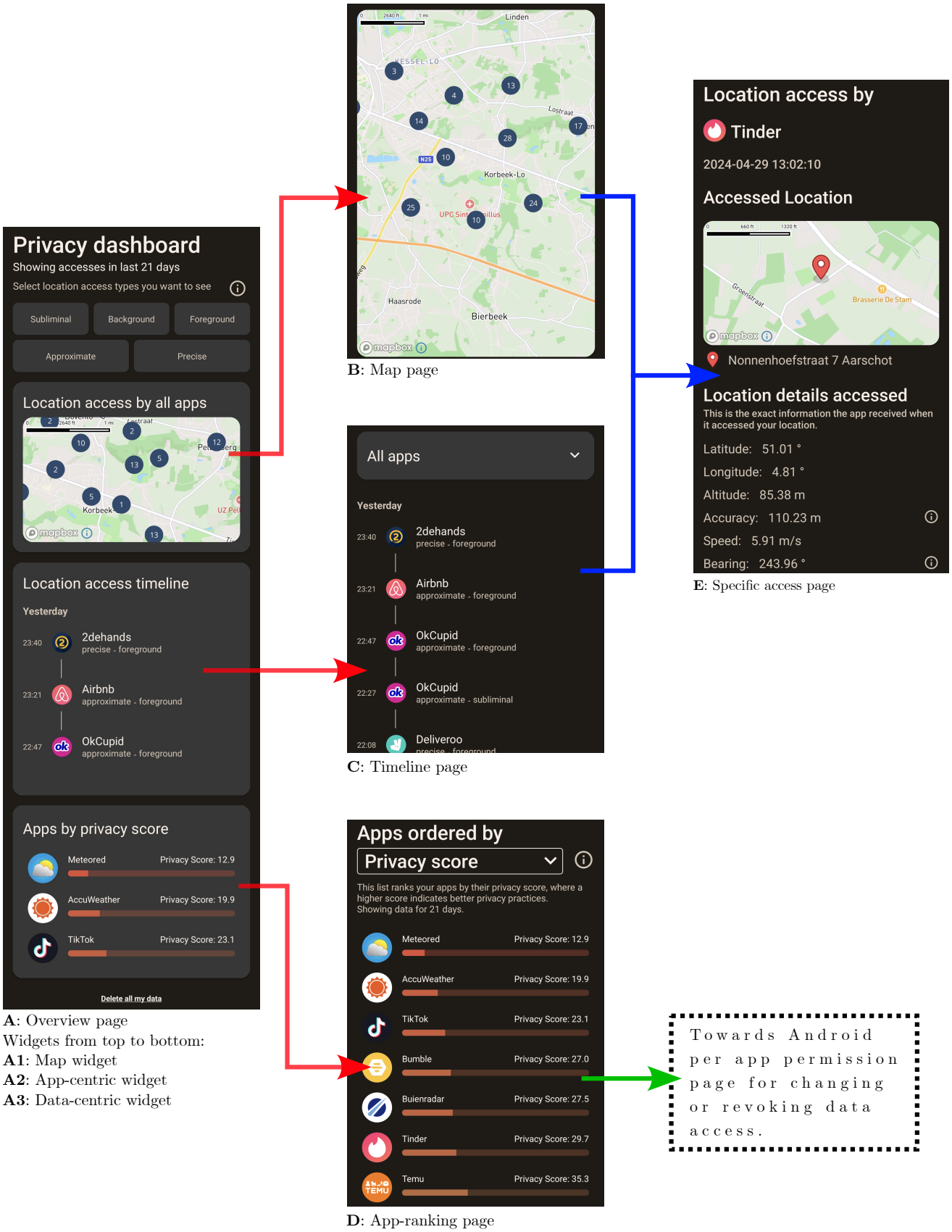
### Visual encoding

In this section, we assign information features to visual variables as described by [17]. They provide guidelines on which visual variables, i.e. position, shape hue, ... are to be prioritized based on whether the data is quantitative, ordinal or nominal.

In this step, we outline the specific data features that will be visualized within the Privacy Dashboard. This includes various elements that contribute to a comprehensive understanding of location privacy and app behavior. The data to visualize includes: active filters, the context of location requests which can be foreground, background, or when the device was off, longitude and latitude information, human-readable addresses, altitude, speed, accuracy, bearing information, time information, aggregated tracking information such as frequency, and a custom privacy score as described in section . Additionally, we include information about the app and guiding information for each page and widget, along with hints that indicate interactivity.

We encode each of these features in one or more of the aforementioned widgets. They are presented on the DASHBOARD OVERVIEW PAGE [4A] together with guiding information, active filters, and a 'delete data' button. We now consider how to use the most important visual variable: Position. The active filters and guiding information are placed at the top as necessary context for the rest of the overview page.

The widgets themselves have an inherently nominal relationship; they cannot be ranked based on the information they present.




**A**: Overview page
Widgets from top to bottom:
**A1**: Map widget
**A2**: App-centric widget
**A3**: Data-centric widget

**B**: Map page

**C**: Timeline page

**D**: App-ranking page

**E**: Specific access page

***Figure 4.*** *Overview of the interactive pages in the dashboard. Red arrows show how interactions with widgets lead to other pages. Blue arrows show how by tapping a location access in the* MAP PAGE [4B] *or the* TIMELINE PAGE [4C]*, the* SPECIFIC ACCESS PAGE [4E] *can be reached. The green arrow shows how the dashboard can automatically refer the user to the permission page of an app if they wish to take action against excessive tracking.*

However, we can subjectively rank them based on how well they entice uninterested users. The map widget stands out in this regard, as map visualizations are generally more appealing than timeline or temperature-like views. For this reason, we position the map widget at the top, making use of the positional visual variable. At the bottom follows the 'delete data' button. To differentiate between widgets and filters, we employ the Shape and Area visual variables. The filters have a natural ordering: subliminal $>$ background $>$ foreground. To indicate whether a filter is active, we use the visual variables of Hue and Lightness.

In the MAP PAGE [4B], the position is filled in with the geographical coordinates, while time information is not included in this view. Shape and hue are used generally by the map, ensuring that we choose contrasting shapes and hues for location access pins to enhance visibility and user comprehension.

In the TIMELINE PAGE [4C], position is used to encode the most important information, which is time. Shape and hue are employed to differentiate between filters, guiding information, and core information. Filter and guiding information are placed at the top because they provide essential context for the rest of the page. App icons encode the nominal data of the app that made the request, occupying hue, lightness, saturation, and shape.

In the APP-RANKING PAGE [4D], filters and guiding information are again positioned at the top, while app icons are used to differentiate between apps. The filter options provide an alternative way of measuring the "amount of tracking" per app. The visual variable of position is used to encode the relative ranking of apps from most tracking to least tracking, as the main purpose is to compare apps to each other. Additionally, hue and length are used to encode the absolute value of the tracking metric. A sequential red-yellow-green color scale is employed to represent the quantitative value, with red indicating a higher level of intrusiveness on privacy.

### *Visual design addressing requirements*

Following the research conducted by Wilkinson et al. [18], the design framework incorporates both app-centric and data-centric user preferences to improve insights on privacy impacts of apps. Two data-centric widgets are used to show the "where" and "when" of data access. An APP-CENTRIC WIDGET [4A2] ranks the top three applications based on the amount of location data accessed, allowing for a comparison of the tracking behaviors of different apps. In contrast, the DATA-CENTRIC WIDGET [4A3] displays a timeline of the three most recent instances of data access, distinguishing between precise and approximate location data as well as different access types. This approach enables users to categorize the various types of location accesses, emphasizing their unique characteristics and differences rather than treating them as the same. The APP-RANKING PAGE [4D] displays all apps ranked by their privacy impact in descending order, allowing users to easily compare the tracking behaviors of various applications. For example, the two weather apps shown at the top of figure 4D were simulated to frequently request precise location data, which makes them more prominent in the visualization compared to apps with less significant privacy impacts. The counting of accesses, the clear visualization of different categories, and the ability to compare apps by Privacy Score align with the requirements outlined in R1.

Requirement R2 stipulates that users should also get meta information about the different permissions. For example, users might not know that an app can either request permissions in the foreground or the background. Therefore location accuracy and access types are clearly defined and explained using expandable info buttons. One such info button can be seen in the top right of DASHBOARD OVERVIEW PAGE [4A]. As another example of meta information, each entry in the TIMELINE PAGE [4C] has an info button for further details, ensuring users can find further information if desired.

Next, we check the compliance with R3. The modular design enables users to customize the arrangement of widgets, allowing them to highlight the data visualizations they find most significant. Additionally, the design includes well-considered default settings aimed at users who may be indifferent, as they are less likely to engage with customization options. The MAP WIDGET [4A1] is optimally zoomed to strike a balance between detail and coverage, and it groups data points into clusters when multiple accesses occur in close proximity to reduce visual clutter. Users can inspect each data point through a SPECIFIC ACCESS PAGE [4E], providing detailed explanations. The TIMELINE PAGE [4C] presents a complete history with timestamps for each access, featuring toggles that allow users to filter by specific apps or view all apps collectively. The modular design, zoom functionality, detailed click-through options, and app-specific filtering exemplify the principle of detail-on-demand, in accordance with R3.

The following design choices contribute to an intuitive user experience, aligning with R4. The MAP PAGE [4B] allows users to zoom in and navigate a larger map for a detailed examination of location data, with the level of detail adjustable by the user. Users can easily switch between app-centric and data-centric views. They can also toggle between viewing data from all apps or filtering for a single app, by means of straightforward controls. A set of such toggles can be seen at the top of the DASHBOARD OVERVIEW PAGE [4A]. All widgets are interactive, leading to their respective detailed pages when clicked. The design of the DASHBOARD OVERVIEW PAGE [4A] leverages the Gestalt principle of figure-ground to clearly distinguish the widgets as separate entities. It helps people perceive visual elements in relation to one another, specifically how to distinguish an object from its background. By incorporating material design and adhering to standard Android UI conventions, the interface provides a familiar experience that aids users in navigating the controls.

The APP-RANKING PAGE [4D] lists apps ordered by their location access within the selected time span. Users can navigate to user control options for each app, guiding the user towards taking actions against overly curious apps as required by R5.

An options button allows users to change the time frame over which data is shown, with text indicating the current time frame. It shows how app behavior can change depending on the selected window, thereby encouraging regular monitoring in accordance with R6.

The design was continually refined through consultations and user feedback, ensuring its effectiveness and appeal. Adding to realism, a data set of real location data was used. A subset of locations from the area where the users live or study, was used as simulated data in the interview sessions.

A respondent raised a new privacy concern about the security of location data stored in the app, fearing potential hacking and unauthorized access. Users also expressed worries that if the

phone were stolen, the data could be directly accessed from the device. To address this, a "Delete My Data" button was added to enhance user privacy control, which can be seen on the bottom of the DASHBOARD OVERVIEW PAGE [4A]. Again, prompted by user feedback, general UI changes were made. Information bubbles were streamlined for better usability. Timeline page bubbles were consolidated into a single central info button. Some buttons, like those for longitude and latitude, were removed due to user understanding. Others were changed to always-visible text to minimize excessive tapping. All explanations were shortened to include only essential information.

### The Privacy Score

One addition that originated from feedback of every single participant, is the Privacy Score shown in **4D**. The goal is aiding users in quickly evaluating the privacy risks associated with each app. For this we design a score which serves as a *heuristic* for how privacy friendly the location data collection behavior is for a given app. The score is a glanceable way to get information and allows users to very quickly see if they want to take action against certain apps.

Prior to discussing the methodology for calculating this heuristic, we will first outline the various types of location data access. Location access types can be categorized into four distinct categories based on user engagement and device activity. *Sanctioned Access* refers to location data explicitly requested by the user during a specific action, such as initiating navigation through a mapping application. In contrast, *Foreground Access* occurs when the user actively interacts with an application, allowing it to update location information without being a direct request. For instance, a dating app may share the user's location with others while they are swiping through profiles. *Background Access* is characterized by location tracking that happens without the user's active engagement, such as when a weather app provides updates based on the user's location while they are using other functions on their device. Lastly, *Subliminal Access* describes location data collection that occurs when the device is inactive but not powered down. An example includes a dating app updating the user's location even when the phone's screen is off. Each type of access highlights the varying degrees of user awareness and control over location data.

Furthermore, we make the distinction between precise and approximate location requests. *Precise Location Access* is defined as the use of GPS or similar technologies to pinpoint the device's exact location with high accuracy, enabling applications to provide tailored services based on the user's specific whereabouts. In contrast, *Approximate Location Access* involves determining the general area of the device using broader methods, such as Wi-Fi networks and cell towers, which can provide a less accurate but still useful estimation of the user's location.

The Privacy Score for an app is calculated by evaluating location access statistics in three categories: foreground, background, and subliminal. Sanctioned location accesses do not impact the score. We make the basic assumption that location requests from an app in the background while the screen is turned off are more suspicious than an app which is currently being interacted with by the user. Daily assessments track the frequency of precise and approximate accesses, assigning penalty weights based on privacy implications. These penalties are summed and deducted from a base score of 100, along with reductions for significant location access clusters, which may represent points of interest like home or school. Points of interest are identified by analyzing clusters of location data points currently in memory, laying within a 30-meter radius. Final Privacy Scores are normalized on a scale from 0 to 100, adjusting outlier scores as needed. Weightings for each access type are derived from average weights suggested by survey respondents, ensuring the scoring reflects user concerns about privacy risks.

### System design

Figure 5 illustrates how the PrivacyBuddy app gets the input data for its dashboard. We have designed a *security architecture* that allows controlled querying of location data access, enabling standard apps to make use of this information. This facilitates designers in creating a variety of efficient visualizations without requiring System API permissions. To ensure secure data release, we build upon the ContentProvider component from the Android Framework. ContentProviders manage access to structured data, encapsulating it and providing a standard interface for other applications. To secure these providers, we create a custom permission that other apps can declare in their manifest, ensuring users are informed when an app requests data from the provider's database. To add to security, we apply the principle of least privilege, and filter out any data that is not a location access request. Furthermore, our architecture allows for additional fine grained access control, query control and aggregated release.

#### System API, privileged apps and standard apps

Figure 5 shows the Android Framework in red. It runs with elevated permissions, which allows if for instance to log all data requests done by other apps. Consequently, it logs all location data requests from all apps installed and running on the phone. This is the information we want to visualize. Due to the sensitive nature of this information, the Android architecture restricts standard apps from accessing such information.

One app that is allowed to access this information is the Settings App, which is how it gets data for the stock dashboard. Privileged apps and API calls are shown in blue on figure 5. The privileged interfaces are part of the Android System API. Privileged apps must be built and packaged when the OS is compiled and the custom ROM is packaged. Privileged System API interfaces are not exported for use by apps once the OS is installed on the device.

We extended the privileged Settings App with a dedicated ContentProvider, such that we can output location data access information to the PrivacyBuddy app. The extensions are shown in green on figure 5. The Location Usage ContentProvider (LUCP) fetches Data Request Logs through a privileged API call and filters the response to only include information pertaining to location accesses.

The PrivacyBuddy app is in all sense a standard app. It only has to inform the user that it needs read permission on the LUCP by declaring that in its manifest. It can then query the LUCP for data, and upon response use the data to visualize tracking behavior. For more information on the implementation, we refer to the source code [1].

[1] The adaptations to the source code for the custom ROM and the Pri-

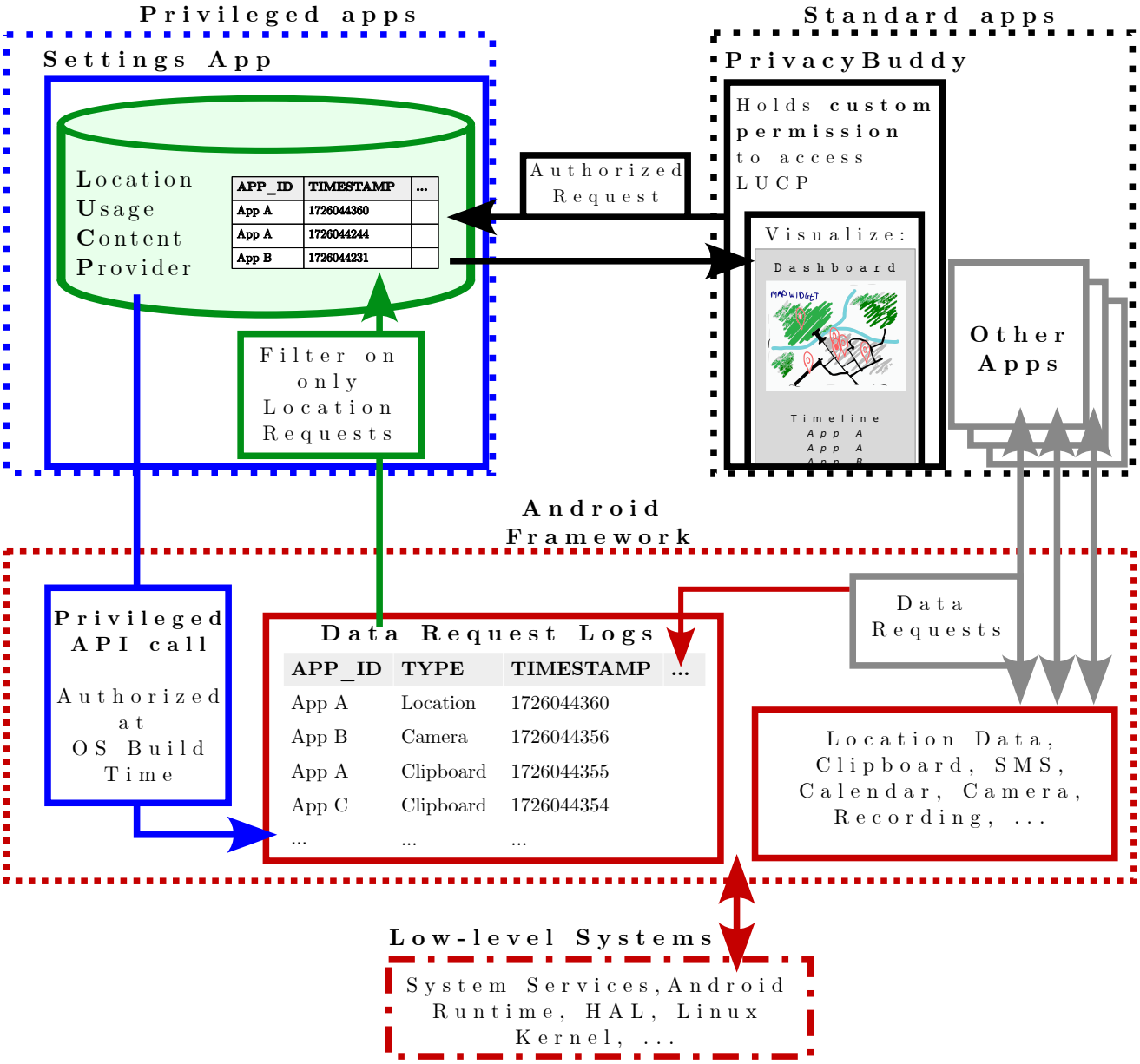



***Figure 5.*** *The System Diagram shows how information flows from Android Framework components (red), to the extensions we made (green) to the privileged Settings App (blue), and finally to the dashboard inside the PrivacyBuddy app (black).*

## Evaluation

We conducted a user study to evaluate our solution. This section outlines the methodology, discusses the findings with user quotes, and assesses which shortcomings of the existing dashboard have been addressed.

### *User Study Methodology*

We used a qualitative research approach to gather detailed insights into user interactions with the Privacy Dashboard App, focusing on individual experiences, perceptions, and suggestions. This feedback is crucial for evaluating the app's usability, design, and features. We prompted for the following objectives.

**Usability.** How easily can users navigate and use the app? Are there any obstacles?

**User Interface Design.** Does the design aid or hinder usability? How does the layout affect the user experience?

**Feature Effectiveness.** Are the features adequate for understanding and managing location data? Which features are most and least useful?

**Potential Enhancements.** What additional features or improvements do users suggest?

#### *Participants*

Six participants were selected to represent a diverse user base, categorized by age (18-30 and 31-60) and technical expertise related to privacy concerns. Aside from the diverse age, each persona, as described in section *Personas*, is represented by two participants.

#### *Format*

Participants had four minutes to explore the app independently, sharing their thoughts on navigation, interface design, and initial impressions. Three tasks assessed the app's effectiveness in conveying location data, such as identifying frequently accessed locations and details of the last known location. Participants then answered open-ended questions on:

- Understanding of the app's purpose.
- Effectiveness in achieving its goals.
- Usability and design.
- Personal engagement and likelihood of continued use.
- Intuitiveness and learning curve.
- Suggestions for improvements.

### *Results*

All participants demonstrated a clear understanding of the app's purpose without the need for prior explanation. Respondent 1 articulated, "The purpose is to keep an eye on how much location data various apps are sharing with the app vendor or owner." Similarly, Respondent 3 added, "Just to always show how many apps are accessing your location," while Respondent 5 emphasized, "To make people aware of how much data you actually send through to companies." Despite collective understanding of the intent to inform users about location data sharing, none of the respondents mentioned taking action against overly curious apps.

Feedback on the app's effectiveness was mixed. Respondent 1 felt it clearly achieved its goal, making it easy to see which apps were frequently checking their location. Respondent 2 agreed, saying it was done in an attractive way. However, Respondent 3 noted some difficulty finding certain features. Respondent 4 found the app a bit overwhelming but appreciated the amount of data it showed.

When it came to usability and design, Respondents 1 and 3 found the widget that displays the most location accesses per app extremely useful. Respondent 2 praised the map as the most visually engaging and informative feature. Respondent 4 appreciated the general information but felt it lacked actionable insights, while Respondent 5 valued the map for its ability to quickly show locations and desired more detailed visualizations. However, some aspects of the app felt unnecessary or confusing. Respondent 1 had difficulties understanding background and subliminal location accesses, and Respondent 3 experienced initial confusion due to the separation between the map and timeline. Respondent 4 felt overwhelmed by the large amount of data presented.

Most participants indicated that they would engage with the app primarily when they first noticed or downloaded it, but their usage would likely decrease over time. Respondents 1, 2, 5, and 6 mentioned they would mainly use the app to limit the number of location accesses by various apps. In contrast, Respondents 3 and 4 expressed skepticism about the app's value with repeated use, as they did not foresee gaining new insights from frequent checks.

Regarding intuitiveness and the learning curve, some features of the app were found to be confusing. Respondents 2 and 6 took longer to understand the full range of interactions available within the map widget, while Respondent 3 found the separation between the map and timeline to be confusing

With regard to suggested improvements, respondents expressed a desire for more actionable insights and simplified data analysis, along with clearer guidance on privacy risks and recommended actions. Improved data visualization was emphasized, particularly the need to distinguish between precise and approximate location accesses. Additionally, better integration and clearer navigation between the map and timeline were recommended to enhance the user experience. Participants also suggested incorporating app-wide filtering options to view location data by app name, access type, or time period. A Privacy Score feature to rate each app based on the frequency and type of location data it accesses was proposed as well. Finally, addressing security concerns regarding the location data stored within the app was deemed essential. This feedback has already been processed for the final version of the dashboard. Overall, participants agreed that they never paid much attention to the stock dashboard, but they expressed enthusiasm for the new design and indicated a strong desire for improvements. This reflects a general sense of optimism about the app's potential after the redesign. Many participants stated that they would be willing to use the app themselves, highlighting its relevance and importance in managing location data privacy.

### *Improvement compared to state-of-the-art*

The final goal is to improve on the existing dashboard, whose shortcomings we enumerated in section *Shortcomings of existing solutions*. Table 1 demonstrates how our privacy dashboard addresses the specific issues found in stock Android designs, using location data accesses as a representative use case.

vacyBuddy app are available at
`https://github.com/toondehaeneKUL/customlineageos`
`https://github.com/toondehaeneKUL/privacybuddy`

| Issue in stock dashboard | Improvement in new design |
|---|---|
| The pie chart illustrates the percentage distribution of each data access type, with a significant 'other' category that offers no meaningful insights. | For the use case of location data, we provide type-specific visualizations such as the MAP WIDGET [4A1] and the timeline. Type-specific visualizations are more efficient according to the user study. |
| The stock dashboard has an unchangeable time window of 24 hours, which is too short for recognizing long-term data access patterns. | This work's privacy dashboard has a configurable time window. For example, users can see visualizations based on data of the last three weeks. |
| The dashboard does not visualize the accuracy of the requested location, even though it is an important feature for determining privacy impact. | Accuracy information of location requests is available. Filters are included to see for example only requests for inaccurate locations. |
| The stock dashboard does not indicate whether the request originated from a foreground or background app, despite this being a key feature for assessing privacy impact. | Information about foreground or background access of location requests is available. Filters are included to see for example only requests from the background. |
| The stock dashboard fails to display the position shared during a location request, despite this being a crucial feature for assessing privacy implications. For instance, many users may be uncomfortable sharing their location while at home, but may have no issues reporting their position when at work. | Information about latitude, longitude, elevation and bearing can be viewed in the SPECIFIC ACCESS PAGE [4E] for each request by every app. |
| The stock dashboard includes little visualization to compare the tracking behavior of apps. The logic of the dashboard to sort, and therefore prioritize, the apps in the list is not clear. | By filtering on specific apps, apps can be compared on the MAP WIDGET [4A1] and on the TIMELINE PAGE [4C]. Apps can also be directly compared on their Privacy Score in the APP-CENTRIC WIDGET [4A2]. |
| The stock dashboard makes it harder to see tracking activity from system apps. It hides away system app information behind an opt-in configuration. | System apps and user apps are treated equally. |
| The apps in the list are not organized in any specific order. This oversight misses the chance to highlight the most significant apps, particularly those that are least privacy-friendly, at the top. | We designed a Privacy Score for the APP-CENTRIC WIDGET [4A2] to offer an easy, glanceable method for comparing tracking behavior. This feature prioritizes potential misuse by sorting apps based on the worst Privacy Scores. |

**Table 1: The privacy dashboard design addresses the specific issues found in stock Android designs.**

### *Limitations*

In this section we briefly discuss limitations and potential threats to validity of our research.

**Small Survey Size.** The user study involved only six participants, yet considerable effort was made to ensure it was representative of the larger user base. A larger sample size could provide more comprehensive insights and validate the findings more robustly.

**Limited Design Iterations.** Due to time constraints, only one round of processing user feedback was done. One more design iteration could have addressed additional user feedback and further refined the app's usability and features.

**Limited Authorization for ContentProvider Access.** Although the permission is included in the PrivacyBuddy app at installation time, the app does not require explicit user interaction to access certain data through ContentProviders. This limitation is inherent to the OS design.

**API Limitations.** Although user tests were conducted with a fully implemented version of the app, the current API does not support fetching altitude data. This limitation affects the app's ability to provide precise location information in real-time. However, this did not impact the user study as we did simulate all levels of granularity.

## Future work

Similar approaches can be used for making other dashboards, or widgets in the same dashboard, but for different types of data requests. For example, it might be interesting to visualize accesses to the clipboard. Data collection by suspicious apps will show patterns, where normal use is rather random when the user is copying small text. Many other data types are possible. To implement such extensions to our work, the system level components of this work can be reused and adapted. Additionally on system level, we believe that the Android Open-Source Project should introduce the ability to create custom ContentProvider permissions that require user interaction before any other app can access them.

## Conclusion

This work aims to help users improve their privacy by providing visualizations that identify excessive tracking by apps. It highlights the shortcomings of the privacy dashboard in both stock Android and open-source variants, such as bad chart choices, hidden data and lack of visualizations that allow to compare apps.

Building on lessons from previous research and leveraging established visual design frameworks, we address the shortcomings of the stock Android privacy dashboard. The User-Centered Design framework was employed to tailor the system to the needs and preferences of target users. This involved iterative design processes incorporating user feedback at every stage, ensuring the final dashboard is both usable and effective.

The Communication-Human Information Processing model proved useful for designing visualizations that capture and retain users' attention, ensuring that the information presented is easily understandable and actionable. This is particularly important in the context of a privacy dashboard, where users need to quickly grasp and act on privacy-related information. Additionally, Gestalt principles were applied to enhance user comprehension and engagement by grouping related information together

and guiding users' eyes to complete visual elements, ensuring they don't miss important information. These frameworks collectively contributed to creating a dashboard that effectively communicates complex privacy data in an accessible and intuitive manner.

The design has simple yet effective interfaces and uses detail-on-demand to cater to a wide range of users. The Privacy Score accounts for which features best visualize tracking behavior and allows users to detect overly curious apps at a glance. We contribute a trustworthy and secure deployment architecture of our implementation components, for both data collection at the Android system level as well as for the privacy dashboard in the form of a user-level Android app.

Using the specific case of location data requests, we demonstrate that visualizations designed specifically for location data access requests are preferred. This preference is confirmed through a user study with participants of varying technical knowledge.

## References


[1] */e/OS product description - a pro-privacy mobile operating system and cloud services — doc.e.foundation*. https://doc.e.foundation/what-s-e.

[2] Talal Haj Bakry and Tommy Mysk. *Precise location information leaking through system pasteboard*. 2020.

[3] *CalyxOS Features — calyxos.org*. https://calyxos.org/features/.

[4] Weicheng Cao et al. "A large scale study of user behavior, expectations and engagement with android permissions". In: *30th USENIX Security Symposium (USENIX Security 21)*. 2021, pp. 803–820.

[5] European Union. *General Data Protection Regulation (GDPR)*. Regulation (EU) 2016/679, Article 5(1)(c). 2016. URL: https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=celex:32016R0679.

[6] Federal Trade Commission. *Complaint: In the Matter of Goldenshores Technologies, LLC, and Erik M. Geidl*. Accessed: 2024-09-12. 2013. URL: https://www.ftc.gov/sites/default/files/documents/cases/131205goldenshorescmpt.pdf.

[7] Adrienne Porter Felt et al. "Android permissions demystified". In: *Proceedings of the 18th ACM conference on Computer and communications security*. 2011, pp. 627–638.

[8] *GrapheneOS features overview — grapheneos.org*. https://grapheneos.org/features#privacy-by-default.

[9] Quinn Grundy, Kellia Chiu, and Lisa Bero. "Commercialization of user data by developers of medicines-related apps: A content analysis". In: *Journal of General Internal Medicine* 34 (2019), pp. 2833–2841.

[10] Kurt Koffka. *Principles of Gestalt psychology*. routledge, 2013.

[11] Jacob Leon Kröger, Milagros Miceli, and Florian Müller. "How data can be used against people: A classification of personal data misuses". In: *Available at SSRN 3887097* (2021).

[12] Michael E. Kummer and Patrick Schulte. "When Private Information Settles the Bill: Money and Privacy in Google's Market for Smartphone Applications". In: *Management Science* 65.8 (2019), pp. 3470–3494. ISSN: 0025-1909. DOI: https://doi.org/10.1287/mnsc.2018.3132.

[13] LineageOS. *LineageOS – LineageOS Android Distribution — lineageos.org*. https://lineageos.org/.

[14] Tara Matthews. "Designing and evaluating glanceable peripheral displays". In: *Proceedings of the 6th conference on Designing Interactive systems*. 2006, pp. 343–345.

[15] Roy D Pea. "User centered system design: new perspectives on human-computer interaction". In: *Journal educational computing research* 3 (1987), pp. 129–134.

[16] Eva-Maria Schomakers et al. "Internet users' perceptions of information sensitivity – insights from Germany". In: *International Journal of Information Management* 46 (2018), pp. 142–150. ISSN: 0268-4012. DOI: https://doi.org/10.1016/j.ijinfomgt.2018.11.018.

[17] Christian Tominski and Heidrun Schumann. *Interactive visual data analysis*. AK Peters/CRC Press, 2020.

[18] Daricia Wilkinson et al. "Privacy at a Glance: The User-Centric Design of Glanceable Data Exposure Visualizations". In: *Proceedings on Privacy Enhancing Technologies* 2020.2 (2020), pp. 416–435. ISSN: 2299-0984. DOI: https://doi.org/10.2478/popets-2020-0034.

[19] Michael S Wogalter. "Communication-human information processing (C-HIP) model". In: *Forensic human factors and ergonomics*. CRC Press, 2018, pp. 33–49.

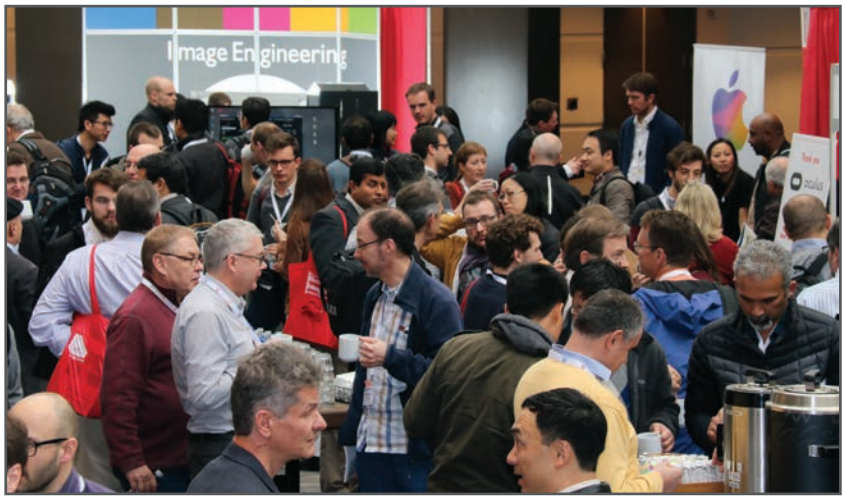

## Author Biography

*Toon Dehaene received his MS in Engineering Technology from the Catholic University of Leuven (2023). His Master's Thesis proposed a novel technique for anonymizing location data on mobile platforms. Following that, he began his PhD candidacy at the DistriNet Research Group under the supervision of Prof. Bert Lagaisse and Prof. Vincent Naessens, focusing on Privacy Engineering. His work centers on algorithms and architectures for enhancing privacy-friendly collection of location data on mobile devices.*

# JOIN US AT THE NEXT EI!

# electronic IMAGING

*Imaging across applications . . . Where industry and academia meet!*





**• SHORT COURSES • EXHIBITS • DEMONSTRATION SESSION • PLENARY TALKS •**
**• INTERACTIVE PAPER SESSION • SPECIAL EVENTS • TECHNICAL SESSIONS •**

## www.electronicimaging.org

ADVANCING IMAGING SCIENCE 75 1947 IS&T 2022 imaging.org