Laura Shipp* and Jorge Blasco

How private is your period?: A systematic analysis of menstrual app privacy policies

Abstract: Menstruapps are mobile applications that can track a user's reproductive cycle, sex life and health in order to provide them with algorithmically derived insights into their body. These apps are now hugely popular, with the most favoured boasting over 100 million downloads. In this study, we investigate the privacy practices of a set of 30 Android menstruapps, a set which accounts for nearly 200 million downloads. We measured how the apps present information and behave on a number of privacy related topics, such as the complexity of the language used, the information collected by them, the involvement of third parties and how they describe user rights. Our results show that while common pieces of personal data such as name, email, etc. are treated appropriately by most applications, reproductive-related data is not covered by the privacy policies and in most cases, completely disregarded, even when it is required for the apps to work. We have informed app developers of our findings and have tried to engage them in dialogue around improving their privacy practices.

Keywords: menstruapps, privacy policy, GDPR, period-tracking

DOI 10.2478/popets-2020-0083
Received 2020-02-29; revised 2020-06-15; accepted 2020-06-16.

1 Introduction

Menstrual and fertility tracking apps, recently dubbed 'menstruapps' [19], are mobile applications that track various aspects of the health and menstrual cycles of their users. Menstrual cycle tracking is something that has been happening for decades [35], but menstruapps are a newer phenomenon. Different menstruapps offer different functionality and this is rarely just charting when a period occurs. As Levy states, "they do not solely represent digital versions of menstrual calendars but support observation, analysis and interpretation of a plethora of physical and mental states as well as behavioural patterns associated with menstrual cycles" [32]. Menstruapps are helping people get pregnant [40], are being used as a certified contraceptives [11], and being described as a "Fitbit for your period" [54].

As with many other applications based on user-generated data, period-tracking apps, and *femtech* more generally have become large sources of revenue. *Femtech* was valued at US$1 billion in 2018 [47], but is estimated to be worth US$50 billion by 2025 [22].

Serious questions have been raised around the extent to which it is possible to trust private corporations with this sensitive and intimate data [19]. There is the potential for the interests of developers to come into conflict with the wellbeing of their users, with the privacy of user data the likely collateral [30]. This is in part due to the unique combination of sensitive information that is entered into menstruapps. It is a mix of personally identifiable information, information pertaining to the health (both general and reproductive) of the user, along with information about the user's sexual practices. This can be as fine-grained as information about orgasms, use of contraceptives or the time of day sex took place. As this is user-entered data (and so not protected by system permissions) it raises specific privacy concerns. Yet, these concerns become more serious when considering the wider context and the scale to which these apps are collecting data. Some menstruapp companies are now boasting tens of millions of downloads worldwide, meaning if data collection occurs it will be widespread. Companies such as Glow Inc. claim to have one of the largest data sets on women's reproductive health in the world; data which is becoming commercially available in ways it was previously not [50]. If the user is ill-informed about what happens to their data, they may assume trust in the app and potentially what they share will increase [29]. Where women's bodies are under surveillance due to the political environment, bad practices within menstruapps could have severe consequences. For example, in Missouri where the state's only abortion clinic is under threat of government-ordered closure, it was found that the state department of health

***Corresponding Author: Laura Shipp:** Information Security Group, Royal Holloway, University of London, E-mail: Laura.Shipp@rhul.ac.uk
Jorge Blasco: Information Security Group, Royal Holloway, University of London, E-mail: Jorge.BlascoAlis@rhul.ac.uk

has been keeping spreadsheets of women's periods with the aim of determining if abortions took place [36].

Despite this serious context, the academic privacy research community have overlooked these kinds of apps. We set out to rectify this neglect by reviewing the privacy practices of a set of 30 popular menstruapps. In particular, the focus of our analysis is how the app developers deal with data collection and the information they provide users about this. With this in mind we aim to answer the following questions: (**Q1**) how well do developers inform menstruapp users about their privacy practices within their policies and privacy communications? (**Q2**) Is the information they provide clear and understandable? And most importantly, (**Q3**) do apps behave differently in reality in comparison to what is stated in their policy? To answer these questions we used a mixed-method approach.

The rest of the paper is structured as follows. Section 2 outlines the analysis that various organisations have completed on menstruapps to date, and where studies have used similar methods to test privacy practices. This will be developed further in section 3 where we explain how we carried out our investigation, comparing it to the work of others. Section 4 describes the obtained results and compares them with the mobile ecosystem. Section 5 discusses the meaning of our findings. Section 6 considers the limitations of our study, and finally, section 7 gathers our conclusions. Overall, we contribute a mixed-method and comparative study of the specific privacy concerns that menstruapps raise.

2 Related Work

Research in the menstruapp space has largely been done by a variety of organisations and individuals, ranging from rights organisations, private companies, consumer groups, and journalists. Each piece of previous work has focused on testing a particular dimension of privacy, rather than a trying to gain a big-picture view of the app privacy landscape. These include investigating the companies behind apps and their business models [19, 46], dynamic analysis of apps and their embedded libraries [2, 48] or evaluating the quality of clinical information they provide [23], among others. While relevant, these studies have addressed specific issues within privacy in isolation and with a disjointed set of apps being analysed. Comparing the results of multiple analyses together over a discreet and constant set of apps is needed to understand how this ecosystem deals with users' privacy. This is of particular relevance within the context of GDPR as some of the previous studies were completed before GDPR was in effect and/or lay outside of its jurisdiction.

Much research within the remit of *femtech* apps has been focused on pregnancy and maternal health rather than specifically around menstruation [18]. A close comparison can be made with work by Scott *et al.* which focused on apps for expecting mothers. A comparative analysis of security functions, usability and reliability of the apps was undertaken and the researchers identified a clear lack of security and privacy functions built into the apps [49]. The privacy of *femtech* has also caught the attention of legal scholars. Rosas emphasises the inadequacies of privacy measures within the sector, forecasting the consequences of a Yahoo-style data breach. They demonstrate that current health data protection laws in the US would not protect users from such a breach, something made more concerning by increased data sensitivity [47]. Where research was been undertaken specifically around menstruapps, the focus has not been privacy. For example, research published in an obstetrics and gynaecology journal evaluated the features offered by the most popular apps and whether the information they were providing was accurate [39]. Otherwise, within the Human-Computer Interaction community, the focus has been about understating why people track and whether the design of apps aided them or not [18], and designing alternative modes of tracking [20]. Value would be added to this work with in-depth analysis of period-tracking apps, allowing users to make comparisons across a broad range of criteria, and determine the best mode for their requirements.

Many works have demonstrated methods for the investigation of the quality of privacy practices of mobile apps. Some of these works focused on the accessibility and quality of the information provided to the user in relation to the app privacy practices. Reidenberg et al. [45] evaluated the discrepancies between how an organisation handles user data and what a variety of users (privacy experts, legal scholars and the general public) understand about those data handling practices. They found significant levels of disagreement about what different user groups understood from policies, suggesting that these can be hugely misleading. Jensen and Potts [31] found a similar mismatch between what the user understood versus what was said within the policy. Others have compared information within the policy with the reality of data sharing practices. For example, Cranor et al. [14] analysed policies from financial institutions and found discrepancies in practices within the same institutions. A later study by Sheng and Cranor [51] found that these practices barely changed when privacy legislation was introduced. Finally, a more recent study by Bowers et al. [7] analysed the language used in privacy policies of mobile money services. Their results show that there is still a lot of work to be done to produce privacy policies that are easy to read and accessible to users.

3 Methodology

Our three research questions address a variety of issues that range from the linguistic analysis of a privacy policy to the actual behaviour of the apps. Because of this, our methodology uses a mixed method approach, combining automatic quantitative analysis with more qualitative aspects when required. This section describes the different steps we used in our methodology to be able to answer each of our research questions.

Our first step was to select a set of apps to study (3.1) and to investigate the developers of those apps to provide some context to their design (3.2). We analysed the scopes of the privacy policies (3.3), the stated purpose of data collection (3.4) and what user rights were mentioned (3.5) to see how well developers inform their users about their privacy practices (Q1). We analyse the language used in the privacy policies (3.6) to see if this information is clear and understandable (Q2). Finally, we analyse the type of information collected, how this is reflected in the privacy policies (3.7), third parties that are embedded within the apps and (3.8) and the app network behaviours (3.9) to see if apps behave differently to what is stated in the privacy policy (Q3).

3.1 App Selection

Android apps were the basis for this research as they are easily downloadable, lower-bounds for download counts are publicly available and many tools exist for their static analysis. All apps were downloaded from the Google Play Store, using a new Google account on a new Android device. App downloads took place in June 2018. The search terms used to find the apps were: 'period tracker, period, menstrual cycle, menstruation tracker and menstrual calendar'. Search results were ranked in accordance with downloads and consumer ratings, filtering out any app with less than 100,000 downloads. This method was in line with Adhikari et al. [1], in which mobile health apps with the highest numbers of downloads and highest ratings of consumers were selected. The search was continued until 30 apps were reached (the limit of our study). This limit existed to enable us to carry out multiple forms of quantitative and qualitative analysis, particularly as some methods were undertaken manually.

Our study was initially restricted to free apps as they are likely to reach a wider audience [49]. During the app selection process, we identified two free apps #4 and #8 that had paid versions with more than 100,000 downloads. We decided to also analyse these two paid versions to verify their privacy practices. Often paid-for versions of apps contain the same poor privacy practices as a legacy from their free counterparts [26].

We added an additional exception to our free-app rule, *Natural Cycles*. This app is free to download but requires the user to pay a service fee after the month-long trial period expires. The app was included in the study as the trial period provides the same functionality as the paid version and this period allowed the researchers to perform all the required analysis for the study.

3.2 Background Analysis of Developers

Menstruapps are apps with the common goal of period tracking. Each app is created with distinct characteristics or features and each organisation developing them has a different business model and way of working. The business model, culture and ethos of each of the companies has impact on how the app gathers data, what they do with it and how this is explained to the user.

We conducted background research on each of the app developers. Rizk and Othman [46] performed a similar analysis on pregnancy and fertility apps, looking at the developers' origin, company business model, and information provided within the privacy policy. Our analysis differed from theirs as we included more apps but all were specifically period-trackers.

The business model of a company will generally dictate how the data collected by the app will be used. Menstruation data provides a highly revealing picture of a persons' reproductive health. If this data is being used for purposes such as targeted advertising, it should be clearly stated in the privacy policy. As part of their business model, we also looked into the other products and services offered by the developer and how they were being marketed. This provided insight into for how the data might be treated by the developers. With this, we verified if developers see menstruation-related data as sensitive and in need of safeguarding, or as just another data point that could be obtained from a non-health related application. The research was carried out by first accessing the Google Play Store to find any accessible information and links to developer websites, and then search engines were used.

3.3 Privacy Policy Location and Scope

It is imperative that privacy documentation is easy for the user to locate. For this reason, we noted where the privacy policy is available, whether within the app or within the Play Store, or both. Bowers et al. [7] investigated the availability of privacy policies and how often they were updated. They highlighted that if the policy is not readily available and up to date the user is not being properly informed about the current privacy practices of the organisation handling their data. This is particularly concerning considering the nature of the data held by menstruapps. Beyond this, we looked for other forms of privacy communication with the user within the app. We manually interacted with each app for 10 minutes, entering the necessary data needed to use the app. This largely included registering for an account and entering period data. In this, we looked for any notice that was given to the user about privacy or the privacy policy, or where the user was asked to enter into any kind of agreement. Given that many privacy policies are very difficult to understand, this may be an important form of information the user gains about the privacy practices of the app [45].

We categorise each privacy policy by their scope with the aim of investigating how effectively the policy communicates information to the user. Our motivation comes from the results obtained by Sunyaev et al. [53] whilst studying this same feature within a set mobile health applications. Broad or different scopes can easily mislead or confuse the user, as it can be unclear whether the policy is relevant for the app used or not. In our work we use the same categories:

- **Single app**: Policy covers only the app analysed.
- **Multiple apps**: Policy covers more than one app, including the one analysed.
- **Back-end services**: Policy covers only the services offered in the back-end of the application.
- **Developer homepage**: Policy covers the developer website only.
- **Company-based services**: Policy covers all the services offered by developer (apps, back-end, website, etc.).
- **No relation to app**: Policy does not mention the app analysed and bears no relation to app functionality.

In this study, company-based services, includes the homepage or website of the developers as some offered the same functionality as menstruapps on their websites. This analysis was undertaken manually given that the language used within each privacy policy differed.

3.4 Purpose of Collection

Since the introduction of GDPR in May 2018, data processors are required to describe the purpose for processing personal data within their privacy policies[1]. In this part of our analysis, we noted the justifications developers gave for information processing. In previous literature [53] these have been classified as: **app operation** (required for the app to work), **personalisation** (tailoring app experience based on collected information), and **secondary use**. In the context of menstruapps, we found that the *secondary use* category can be split in three new categories: **scientific use**, **aggregation** and **usage by third parties**. Scientific use refers to the usage of collected data for scientific and academic purposes; aggregation refers to the collection of data for analytics for the use of the developers, and third parties refers to any occasion where the data is being transferred to a third party for additional processing. This includes when data is being transferred to a third party that provides targeted advertising services. This analysis was undertaken by manually checking each policy for discussion of processing practices. Additional analysis performed on how third parties are included within the policies is described in section 3.8.

3.5 User Rights

Our framework also examined how each of the policies described the user rights to data access, deletion and portability as defined by GDPR. We choose these as they provide the most tangible and most common reference to GDPR within the privacy policies.

As the policies handled user rights differently, focusing on these three issues allowed for some consistency

[1] Note that all the apps are available in several European Union Google Play Stores and therefore GDPR applies to them.

across the policies. A similar approach was followed by Bowers et al. [7] and Wilson et al. [55].

We also looked for developer contact details and for mention of an appointed Data Protection Officer, both within the privacy policy and throughout the app. We then looked for where contact details may be provided within the app as part of feedback or support functions. All of these aspects are crucial requirements that a privacy policy should fulfil and therefore would contribute to answering Q1.

3.6 Language

Previous research shows that privacy policies are generally long [24, 53], ambiguous [44], difficult to understand [15, 38], and sometimes downplay the privacy implications of data collection [42]. This is despite the fact that privacy policies remain the key place where users may try and find out information about data handling and sharing practices [45].

We identified how good each of the policies were at clearly and specifically informing their users about their privacy practices. In research conducted on mHealth apps, Brunton found the apps had consistent, “obtuse linguistic constructions when describing their use of tacking, monitoring, and data collection” [8, p30]. This not only makes them more difficult to understand but directly contradicts privacy policy requirements as established in regulations such as GDPR: "...in a concise, transparent, intelligible and easily accessible form, using clear and plain language..." [13].

Our methods and rationale for conducting this analysis were inspired by previous works [5, 9, 41], which focused on how language flexibility can be used to obfuscate privacy policies. In our analysis we look into:

- **Qualitative adjectives and adverbs** that modify the emphasis of an element of the text.
- **Temporal adverbs** that may affect the frequency or introduce uncertainty in actions being mentioned in the text.
- **Conditional verbs** that may introduce uncertainty around whether something happens or not.

We used *spaCy* to split each privacy policy into tokens and extracted all adjectives, adverbs and verbs [28]. For each list of adjectives, adverbs and verbs we filtered out those that did not fit the previous description[2]. We then extracted all sentences that included a mention of any of the words, and analysed how that specific word is used in the context of that sentence.

For each privacy policy we also calculated their Flesch Reading Ease Score. This allows us to compare their complexity with apps in the mobile health space [31]. This, in turn, helps to address Q2 and the clarity of the messages given to the user through the policy.

3.7 Types of Information Collected

Menstruapps by nature require used-entered data, usually around a user’s reproductive cycle, to carry out their functionality.

The sensitivity of this data suggests that its processing would be included within the privacy policy, to inform users about how it is being used [29]. As apps can capture the same kind of data under different names, we first categorised all the possible types of data under a set of fixed classes. Our categorisation follows similar principles as Sunyaev et al. [53] but also includes reproductive health and sexual life related categories. These data classes fell into the following categories: **Personally identifiable information** (PII), **period data**, **reasons for using the app**, **sex logging**, **collection methods**, **test results** and other **health information**. Each of these categories had their own subcategories of more specific classification [3].

Information about collected data was then compared to what was mentioned by the privacy policy and what was transmitted via the network (section 3.9). This can help identify whether app behaviours follow the collection they state in privacy notices (Q3). Moreover, if one app collects significantly more data than others, with a similar functionality, this suggests that the data is not really required by the app.

3.8 Third Parties and Libraries

App developers use third party libraries and services to monetise their apps, to integrate other platforms or to provide additional services to their users. Often, the goal of third party code is collecting information about user interactions with apps [4]. When advertising libraries

[2] A list of the words we kept in each category can be found in Appendix C

[3] These are shown in Appendix B.

are present within an app they can 'leak' user information, such as the IMEI [33]. The presence of third party libraries can therefore come into conflict with user privacy. In a best case scenario the app should explain third party library behaviours, and at the very least, the third parties' privacy policies should be linked [56]. Moreover, under GDPR guidelines, data subjects now have the right to restrict the processing of their data by either developers or third parties. For this reason, we analyse how these third parties and their libraries are handled within menstruapp privacy policies. The aim of this was to investigate the clarity and explicitness of information about third party library behaviour given to the user. For this, we first used *LibScout* [3] to identify the libraries present in the Android binaries. We executed LibScout with library profiles extracted from March 2019. We discarded partial matches and checked for inter-library dependencies. Then, we identified which third party libraries were disclosed to the user in the privacy policy, checked whether information was given about what data was collected and checked whether the relevant third party privacy policy was linked. We finally compared the obtained list of libraries from the app binary against the results obtained from the privacy policy to check for omissions, therefore comparing what is claimed to the reality of the app behaviours.

Table 1. List of analysed apps and corresponding developers with their Google Play category.

ID	App. Name	Developer	Cat.
1	My Calendar	Simple Design Ltd.	♣
2	Clue	BioWink GmbH	♣
3	Flo	OWHEALTH, INC.	♣
4	P.Tracker (Deluxe)$^{\$}$	GP International LLC	♣
5	Maya	Plackal Tech	♣
6	My Calendar	SimpleInnovation	♡
7	Lilly Tracker	SMSROBOT LTD	♡
8	WomanLog (Pro)$^{\$}$	Pro Active App	♣
9	LadyTimer	Vipos Apps	♡
10	MyDaysX	Single developer*	△
11	M. Calendar	witiz.	♡
12	Petal	Blue Group	♣
13	Amila	Amila	♡
14	Glow	Glow Inc.	♣
15	Ovia	Ovia Health	♡
16	MC	Staywell	♣
17	My Tracker	Leap Fitness Group	♣
18	Once	Malang	♣
19	Eve	Glow Inc.	♣
20	Period Tracker	Leap Fitness Group	♠
21	MyPeriodTracker	Linchpin Health	♣
22	My Calendar	Living Better	♡
23	Luna	tinyChangeCompany	♣
24	Natural Cycles	NaturalCycles AG.	♣
25	Cherry	Rosa Care Group	♯
26	Period Tracker for Women	Global Tech Social	♣
27	P. Tracker	Sevenlogics, INC.	♣
28	PD	Nanobit Games	♣
29	Wonder Period	Wonder App Studio	♣
30	Menstrual Cycle	4Rice Group	♣

♣ = Health and Fitness. ♡ = Medical. △ = Lifestyle. ♠ = Beauty. ♯ = Tools. ()$^{\$}$ = Name addition for paid versions of the apps.

3.9 Network Traffic Analysis

To complement the analysis performed in sections 3.7 and 3.8, we performed a network traffic analysis on the apps. The goal of this analysis was to identify what kind of information was being exfiltrated from the device. In particular, we looked for personal data that would be introduced during app usage or other device data that would be sent via the network.

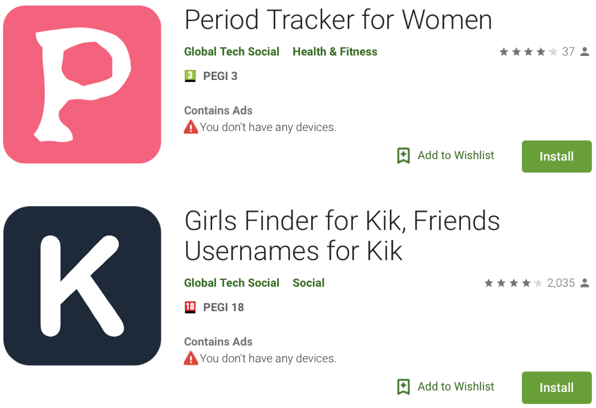
Privacy international [2] performed a similar analysis, investigating the data sharing practices of popular menstruapps across the world (as different apps are popular in different places). Our analysis followed a similar, and common, methodology for HTTP traffic interception. We used the same test device with Android 6.0 as in Section 3.3. This version allows for user-installed certificates that facilitate the analysis of network captures. We installed a self-signed certificate and redirected all phone traffic through a decryption proxy. After starting the packet capture, we performed a fresh install of each app and interacted with it for approximately 10 minutes. During this time, we signed up for an account (if required), entered menstruation and other user-related data and looked for sync and back up options and executed those. We then analysed the captured data files looking for specific data patterns that were introduced in the app to facilitate data matching. In three cases (apps #2, #14 and #19), we were not able to capture app-related data due to the use of certificate pinning. These are marked as CP in table 2. For these however, we were able to capture data that was sent to third parties via embedded libraries.

4 Results

Table 1 shows the list of apps and corresponding developers included in the study. Overall, the apps studied represented 178 million downloads in the Play Store. For 16 of these apps we were able to find the corresponding iOS app and verify that they were governed by the same privacy policies. Download estimates for the corresponding iOS apps are not publicly available so we expect the actual download figure to be significantly greater. The two paid apps included in the study were also governed by the same privacy policy as their respective free apps.

Many of the apps were developed by larger companies that had a specific focus around women's digital health. These were apps #2, #3, #5, #13, #15, #19, #24. These are organisations that had websites that provided a variety of additional services, such as health and wellness content or assisted fertility services.

Most apps did fall under the category of health and fitness (20) and medical (7). However, we found some outliers with apps #10, #20 and #25 placed under lifestyle, beauty and tools respectively. Most of these, were created by developers who also made apps with other functionalities. These included weather and games and in one case (developer of app #26), an app that would allow users to "meet kik Girls and ... chat with your perfect match Girl" (Figure 1). This app did not have a privacy policy.



Period Tracker for Women

Global Tech Social Health & Fitness

PEGI 3

Contains Ads

You don't have any devices.

Girls Finder for Kik, Friends Usernames for Kik

Global Tech Social Social

PEGI 18

Contains Ads

You don't have any devices.

Fig. 1. Apps developed by Global Tech Social.

There were three companies that developed more than one app from our selection of apps (#14 and #19, #16 and #22 (same policy and contact details), #17 and #20).

Table 2 shows a summary of the results of executing our methodology. Our results show a great variability in how the developers treated and discussed the capture and process of user data. Within this diversity of results, we were able to identify several sets of practices that were common across different developers. In this section, we describe these around common themes. This provides insight on the state of the menstruapp ecosystem and allows us to draw conclusions about what apps are doing well and where there are areas for improvement.

4.1 Policy Availability

70% of apps provided a link to the privacy policy in both the Play Store and within the app allowing the user to easily access it. Beyond this, 3 apps provided it in the Play Store alone and 2 provided it only within the app. There were 4 apps with no privacy policies (in the Play Store or the app), despite being downloadable within the EU. These apps collectively account for at least 11.5 million Android downloads, meaning that a wide menstruapp user base is not given any information about the privacy practices of the app they use. None of these apps had any other form of privacy communication with the user, for example terms of service, or requesting agreement in the form of consent boxes to be ticked when opening the app. One of these apps (#26) did not send any information to the internet. In another case, app #9, there was a link under the privacy policy heading, but the link produced a *File Not Found* error message. This app sent all personal, period and other information entered into the app to the app servers.

Surprisingly, the free version of app #4 included the privacy policy within the app (in the Support screen) while the paid version included exactly the same screen without a link to the privacy policy or the terms of service. The links to the privacy policies in their respective Google Play pages pointed to the same privacy policy. Comparatively, menstruapps perform better than average in the market. Zimmeck et al. [57] studied the privacy policies of 1 million apps and found that only 50% had policies linked in the Play Store.

4.2 Accuracy and Completeness

Once accessed, we found the scope of the policies varied. In only 8 apps the respective privacy policy was directly related to the app functionality. 13 other apps had policies that included multiple apps (5) or services (8). This can be confusing in cases where the other services offered by the company have nothing to do with the app downloaded (for example, app #7 states that their privacy policy covers their website, games and apps developed by the company and other products or services). In some cases, such as #14, this does not affect the

Table 2. Summary of results obtained from the execution of our methodology

		Scope and Rationale						Language						3rd Parties				User Rights			Network Analysis	
App ID	Has P. Policy	Scope	App Operation	Personalisation	Agg. or Analytics	Scientific Use	Third Parties	Flesch Score	# Adverbs	# Adjectives	# Verbs	# Obf. Sentences	Text Length (sent.)	Embedded	Mentions	Data	Links	Contact	Deletion	Access	App	Libraries
1	Y	NA	Y		Y		Y	53.44	17	12	8	19	31	Y	S	N	1	G	N	N	-	DD
2	Y	CS	Y	Y	Y	Y		54.82	67	113	61	127	277	Y	S	S	2	S	Y	Y	CP	DD, PD
3	Y	SA	Y	Y	Y	Y	Y	45.54	88	156	65	156	309	Y	S	S	0	S, DPO	A	Y	ALL	-
4	Y	MA	Y				Y	42.42	11	19	18	25	51	Y	G	G	0	N	A	Y	ALL*	DD$^{\$}$
5	Y	SA	Y		Y		Y	53.60	12	33	28	40	86	Y	G	G	0	G	A	N	ALL	ALL
6	Y	MA	Y		Y		Y	45.31	20	34	32	49	134	Y	S	S	12	DPO	N	I	-	DD
7	Y	CS	Y				Y	53.98	10	23	20	32	88	Y	S	G	2	G	A	N	ALL*	DD
8	Y	MA	Y		Y			55.17	1	12	1	8	29	Y	N	N	0	G	N	N	ALL*	DD$^{\$}$
9	N	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	ALL	DD
10	Y	MA	Y		Y		Y	38.81	221	266	140	375	783	Y	S	G	21	DPO	Y	Y	ALL*	DD
11	N	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	-	-	-	-	-	DD
12	Y	NA	Y				Y	43.38	6	12	10	16	36	Y	N	N	0	N	N	N	ALL	DD
13	Y	SA	Y	Y	Y		Y	54.22	10	21	9	20	57	Y	S	S	0	G	Y	I	-	DD
14	Y	CS	Y	Y	Y	Y	Y	54.89	10	21	9	20	57	Y	G	S	0	SE	Y	Y	CP	-
15	Y	CS	Y	Y	Y	Y	Y	45.74	59	73	63	99	169	Y	G	G	0	G	Y	Y	ALL	DD
16	Y	CS	Y	Y	Y		Y	37.56	49	85	65	104	208	Y	G	G	0	SE	Y	Y	-	DD
17	Y	MA					Y	53.49	10	8	3	10	16	Y	S	N	1	G	N	N	ALL*	DD
18	Y	SA	Y	Y				39.69	10	28	11	30	115	Y	N	N	0	S	Y	Y	-	DD, PD
19	Y	CS	Y	Y	Y	Y	Y	54.89	10	21	9	20	57	Y	G	S	0	SE	Y	Y	CP	DD
20	Y	MA					Y	53.49	10	8	3	10	16	Y	S	N	1	G	N	N	-	DD
21	N	-	-	-	-	-	-	-	-	-	-	-	-	Y	-	-	0	-	-	-	-	DD
22	Y	SA	Y		Y		Y	45.31	20	34	32	49	134	Y	S	S	9	DPO	N	I	ALL*	DD
23	Y	NA	Y	Y	Y		Y	33.03	22	44	21	47	84	Y	S	N	1	DPO	Y	Y	PD, CD	DD
24	Y	SA	Y	Y	Y	Y	Y	45.33	53	86	33	97	226	Y	G	S	1	SE	C	Y	PII	DD
25	Y	CS	Y	Y	Y	Y	Y	47.44	29	36	39	65	137	Y	G	G	0	G	Y	N	ALL	DD, PD
26	N	-	-	-	-	-	-	-	-	-	-	-	-	N	-	-	-	-	-	-	-	-
27	Y	NA	Y	Y	Y		Y	50.16	12	30	18	34	79	Y	G	N	0	N	N	N	-	-
28	Y	NA	Y		Y		Y	45.87	19	29	36	53	119	Y	S	N	0	DPO	Y	Y	ALL	DD
29	Y	SA	Y	Y	Y		Y	41.21	10	10	14	19	57	Y	S	N	1	N	N	N	-	DD
30	Y		Y				Y	59.45	8	9	8	21	52	Y	S	N	0	N	N	N	-	-

Y= Yes. N= No. SA= Single app. MA= Multiple apps. BS= Back-end Services. CS=Company-based services. NA= No relation to the app. S= Specific. G = Generic. SE= Several. A= They keep aggregated data C= With conditions. I= Incomplete. DD = Device Data. PD = Period Data. ALL = All collected data. CP = Certificate pinning. CD = Codified data. * = only when performing backups. $\$$ = not present in paid version of the app

clarity of the policy as the company offers the same services across different platforms (web, mobile app, etc.). A similar analysis of health apps in 2015 found that only 66% of 600 apps did not have a policy that specifically addressed the app it was assigned to [53]. This could mean that menstruapps perform better, although this improved number could be an impact of GDPR legislation.

The most common purpose of collection was app operation (24 apps). Personalisation of services was less common with around half of this number of apps (13) collecting data under this rationale. 7 apps had policies that referred to internal and/or external scientific research (in menstruation-related fields) as a rationale for data collection. In this regard, app #2 provided the most detailed information, including a link to previous research carried out using data obtained from the app, stating that researchers undergo a vetting process and specifying that only anonymised data is shared with researchers. App #24 followed a similar approach in its policy.

23 apps mentioned they collected data to be shared with third parties for additional processing, including the provision of targeted advertisements or the embedding of third party links. 19 apps specified that they required personal data in some aggregated form and sometimes used this to perform usage analytics. As described in Section 4.5, most apps used third party services to perform this analysis. However, 2 of them (#8,#18) did not make any reference to these services in their privacy policies but used them for advertisement and analytics purposes.

Some of the analysed privacy policies correctly stated the specific data that was being collected by the app as well as the purposes for the collection of that data. The best app in this regard was #13. All the required or optional data items collected by the app were explicitly mentioned in the privacy policy (including three related to the menstrual cycle). The privacy policy of app #2 included the most comprehensive list of captured data mentioned within the privacy policy, justifying their need for each piece of data. In addition to this, they provided supplementary materials to their privacy policy to explain, in more accessible language, what happens to user data when it is entered into the app. This was done through a blog post linked in their privacy policy. Within this post, the company outlines how they treat a piece of user data, broken down into different categories, and demonstrating what it does with them. App #24 similarly had an additional resource for users, which they titled their Data Privacy Mission Statement which outlined what the company would and would not do with the user data. While the language of these texts is more accessible, the need of these extra resources, raises questions about the effectiveness of privacy policies to properly inform users about how their data is processed.

Throughout all of the apps, 7 had a Data Protection Officer named and contactable through their privacy policies. Within that #2 had a trust team. A further 2 apps (#6 and #22) mentioned a DPO being contactable through their general email address, and a third app (#28) had an email address with DPO in it. The company responsible for #14 and #19 had various contact details available for different areas of privacy and data protection, allowing a user to find the contact most relevant to their query. They also had an EU-specific representative. All but 3 apps provided contact details within the app, but 15 of these were in a feedback or bug reporting form. This was more effective in apps like #15 and #28 which had the added functionality of a chat forum with the support team, and apps #14 and #19 provided a submission form for things like data access requests. Apps #2, #3 and #4 also provided dedicated support zones. All of these systems allow users to raise concerns with the developer more easily, providing them with specialists to talk to or procedures for exercising their rights.

None of the apps were able to provide the necessary information on all privacy rights, as determined by GDPR. The apps that came closest to this were #2 and #10, #13. #2 had full deletion and a time frame for completion. #10 had full deletion, but also provided data access in a portable format. Only a third of the apps provided a button within the app for deleting data or deleting a user account (in which data was held). Surprisingly, our network traces showed that app #22 would not send any notification to the server to delete backups when this button was pressed. App #13 provided data access, and full deletion within a specified time frame. Beyond these 3, apps #14, #16, #18 and #19, provided full deletion, and data access. Considering that some of the developers had specific teams devoted to privacy, this issue demonstrates how difficult it can be to appropriately cover all the available rights under GDPR. As has been found in other studies of self-tracking services, the place where these apps were most likely to fail was in providing access to data held about the user [30]. When looking into the opt-out options for targeted advertisement (often described as ‘personalisation’), only 33% of the apps provided these within the app.

Overall, apps developed by companies focused on women’s digital health demonstrated a better understanding of the sensitivity of the data handled by menstruapps and responded accordingly with their policies and practices. Developers without this focus seemed to consider menstruapps as another marketplace to develop just another app, with no special privacy considerations taken.

4.3 Language Accessibility

It can be challenging to provide information on unexpected privacy practices in a way that is not overly long and ultimately impairs transparency [24]. A privacy policy should be easy to read for a wide range of users and have sentences with clear meanings that do not obfuscate important information.

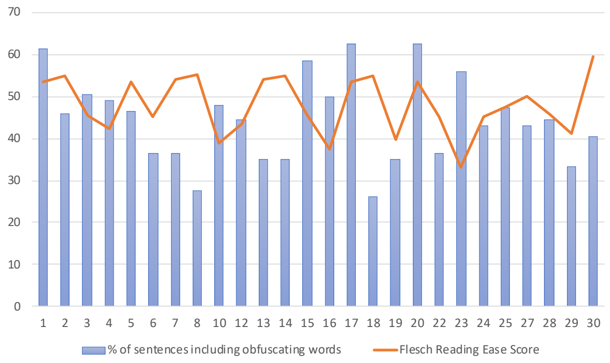


Fig. 2. Obfuscating language statistics and Flesch Reading Ease scores obtained for the privacy policies.

Figure 2 shows the Flesch Reading Ease scores obtained for each of the privacy policies. The maximum score obtained was 59.45 for app #30 - an app that did not exfiltrate any personal or device data. Apps #2, #7, #13, #14 and #19 had the next best results (54-55), but were still some way from values where considered to be written in plain English (60-70) [21].

The fact that #2 had one of the best readability scores is significant given that it is one of the longer policies of the set. It emphasises that a policy can be long in order to cover the appropriate points, which could be more complex in collection-heavy apps. This, however, does not have to come at the expense of the accessible language. Policies from #7, #13, #14 and #19 were shorter than #2 and so potentially easier for the user to get through. Comparatively, apps #2, #14 and #19 had fewer sentences with obfuscating terms than other policies.

Many, if not all policies contained more than one sentence in which the language used resulted in ambiguity, or were inaccessible or were misleading in some way. Figure 2 also shows the percentage of sentences where obfuscating words were found (see Appendix C). The median and average Flesch reading ease score obtained for the policies was 45 and 47 respectively. This indicates that the language in these policies are often difficult to understand by people without a university level education [21].

One particular issue was the length of the policy. In the case of #10, the policy consisted of 783 sentences, with 48% of those sentences including terms that introduced uncertainty. This, in conjunction with its Flesch reading score of 38, makes it the most complex policy in our study. A key redeeming feature of this policy was the layered approach that it took. It provided a small summary at the beginning of the policy so the user could quickly access relevant information. However, given the length and complexity of the policy it is more likely that users will not read those details, resulting in them being as uninformed as users of apps with no privacy policies [24]. Although the readability of menstruapps seems to be slightly easier than the average [31], there was greater variability among them. Our research shows, in line with previous research, that there is still a lot of work to be done to make privacy policies more accessible.

On many occasions, obfuscation occurred due to the usage of the word *may*, linked to the actions of either the developers or the user. A detailed analysis of these cases showed that the usage of such verbs increased uncertainty when developer actions were the subjects of the sentences. Examples of this can be seen in the policies of app #3: "We may share certain Personal Data [...]" or app #10: "We may share only location and other automatic collected device information [...]". These statements are particularly uncertain as they fail to provide clarity about the action. This is because of the use of *may*, and how that action is carried out because of the use of words like *certain* or *other*. In contrast, the use of *may* in relation to the user generally offered the option to act on their privacy rights. For instance, app #2 includes a statement describing how the user can withdraw their consent: "You may withdraw your consent to this use of your data at any time by deleting your Clue account".

Unfortunately, examples of the usage of words to introduce uncertainty can be drawn from many privacy policies. In #8, the privacy policy introduces the modifier of *potentially* to refer to sensitive information: "In cases where partners are receiving potentially sensitive information, we require them [...]". This introduces uncertainty about both, the data that will be shared (the user is not informed about what information is considered sensitive) and the partners it will be shared with. In #6 and #22, the policies state: "Please note, however, that the information we have about you might still qualify as 'personal data' under GDPR". In this case the privacy policy fails to specify again what kind of the collected data is actually considered as personal data.

Some policies were too short to provide the user with the necessary information to understand how their data was being processed. In some cases this was exacerbated with obfuscating terms, making it difficult to understand the little information provided in the privacy policy. During our research, we found apps fitting into both of these policy behaviours. The shortest policies were #17 and #20 with 16 sentences, #8 with 29 sen-

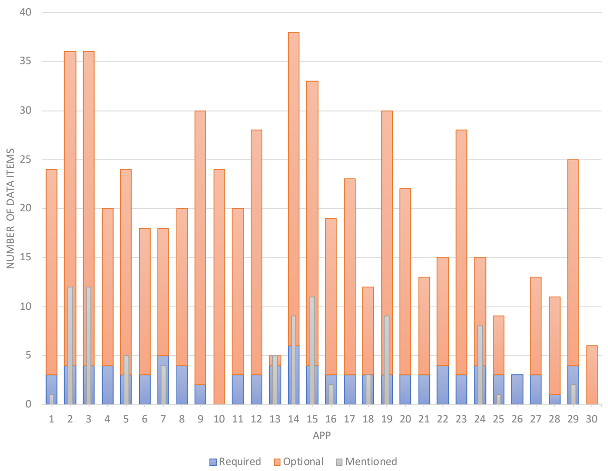


Fig. 3. Number of data items collected by each menstruapp compared to the number of data items mentioned in the corresponding privacy policy.

tences and, #1 with 31 sentences. Within this last policy, 61% of its sentences included terms that increased uncertainty. This is significant as the app has 100 million Android downloads.

In all these cases, the privacy policies shared several sentences with policies from other mobile apps, indicating that they are either based on templates or were taken directly from other services. This makes the meaning of the policies even more unclear as if the policies are not written by the developer there is less assurance that the specific app behaviours are reflected in that policy.

4.4 Data Handling

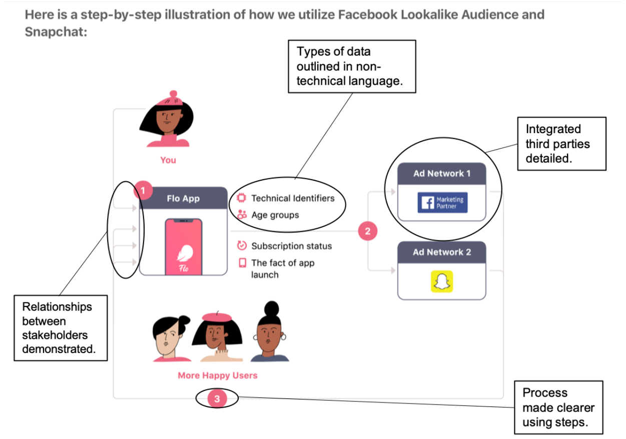
Where apps did have policies, it was sometimes difficult to ascertain whether the policy covered the app it was supposed to. We found that 40% of the apps had policies that covered, at best, all products and services provided by the developer and, at worst, had no relation to the app at all. This introduces uncertainty about whether the practices you are reading about relate to your app and data as a user. The generic nature of these apps resulted in them failing to mention the specific sensitivities of some of the data collected by the app. For instance, the privacy policy of app #28, mentions collecting personal data such as email, IP address, and other PII to operate the user's game account. The app, which does not include any game-related features, sends all the collected information, including period related data, to the app servers.

There remained further ambiguity when searching for how information collected by the app was handled by app developers. Figure 3 demonstrates the levels of data collected vary from app to app. A notable trend was that the number of data points required by each app does not match up to the data points mentioned within the privacy policy. The data required was greater than the data mentioned in 56% of apps. Surprisingly, the most common category of data collected, but not mentioned were those pertaining to a user's menstrual cycle. Most of the apps (28) required period data to work and almost half of them (14) sent menstrual data to the app servers. However, only 6 explicitly mentioned this data in their privacy policy. The amount of detail given by each privacy policy in comparison to the amounts of data needed for the app varied also. Out of these 6 apps, only 3 specified about the dates of the user's last period, whilst all 6 apps needed more data items than this (Appendix B).

#13 was the app that provided the most comprehensive list of the required or optional data items within its policy. The privacy policy of app #2 included the most comprehensive list of captured data mentioned within the privacy policy. App #3 explicitly mentioned all except two required data points. These were included within the more broad term of period data. Whilst the research was being undertaken, its developers updated its privacy policy (July 2019) [4]. In their new version they now provide diagrams to visually demonstrate what happens with user's data. This matched our results from the network analysis of the app traffic. An example of this can be found in Figure 4, which has been amended to accentuate the clarity of the message now provided to users.

6 of the analysed apps transmitted all collected data to the app servers for backups (#4, #7, #8, #10, #17 and #22). Only 2 of them (#8 and #10) mentioned this fact within the privacy policy, in both cases requiring the users to create an account to enable this service. In one case, app #17, the developers even stated in their privacy policy that they did not collect period data which was not the case: "The only situation we may get access to your PII is when you personally decide to email us your feedback or to provide us with a bug log report. The PII we may get from you in that situation are strictly limited to your name, email address, device

[4] For consistency, the research used the version of the policy that was originally downloaded with the rest of the privacy policies



Here is a step-by-step illustration of how we utilize Facebook Lookalike Audience and Snapchat:

Types of data outlined in non-technical language.
Integrated third parties detailed.
You
1
Flo App
Technical Identifiers
Age groups
Subscription status
The fact of app launch
2
Ad Network 1
Marketing Partner
Ad Network 2
Relationships between stakeholders demonstrated.
More Happy Users
3
Process made clearer using steps.

Fig. 4. Adapted diagram from app #3 privacy policy highlighting where the company has made things more transparent

information, location Information and your survey response only."

In some cases, there were examples where developers invited the user to give over further information without clear explanation of what they gained for doing so. This is demonstrated with the sometimes excessive number of data collection options and how this varied across the menstruapp app space we studied (Appendix B). These included, adding the collection method used to manage period blood, relationship status, zipcode or a profile photo, which arguably do not meaningfully add to the functionality of monitoring a reproductive cycle. An example of such behaviour can be seen in app #15: "[...] Because our Services get more fine-tuned and better with data, you may choose to tell us more about yourself [...] to experience these benefits, including [...] adding a photo of yourself to your home screen, [...]". Our network traffic analysis showed all these data points being sent to the app servers. Even with the continuous advancements in machine learning today, it is highly difficult to see how a profile picture may improve the period tracking services offered by these kinds of apps.

4.5 Use of Third Parties Libraries

The execution of LibScout showed that all free apps except one included at least one library from a third party capable of collecting personal data. Facebook and Google Ad Mob were the most prominent libraries, appearing in at least 29 and 13 apps for Facebook and Google respectively. Our network analysis showed that all of these apps sent device data (including identifiers) to Facebook servers. However, only 7 of them included a reference to the collection of device identifiers, which was the main piece of data collected by the library. Of these, 5 privacy policies (#2, #3, #6, #13 and #22) explicitly mentioned the third party libraries that were embedded within the app and the specific kind of data that would be sent to them. Although this is a relatively low number (16% of the analysed policies covering 26.5 million downloads) it shows that privacy policies can properly describe how third parties collect user data. In some cases, such as #2, each third party would be included under a specific section within the privacy policy. Although we could not capture traffic sent to developer-owned servers, network analysis of #2 revealed that embedded third parties also received period-related data in the form of in-app events. These are messages that are generated whenever the user performs an action within the app and include all the information that was typed-in or selected by the user (e.g. new period entry). The privacy policy of #2 explicitly mentioned that the third party would be receiving personal data, but did not specify which kind of personal data. In this particular case, opting out could not be done via the app, but could be done via an email to an address specified within the privacy policy. This kind of behaviour makes it difficult for the users to identify which third parties are collecting their data and how they can exercise their rights regarding that data.

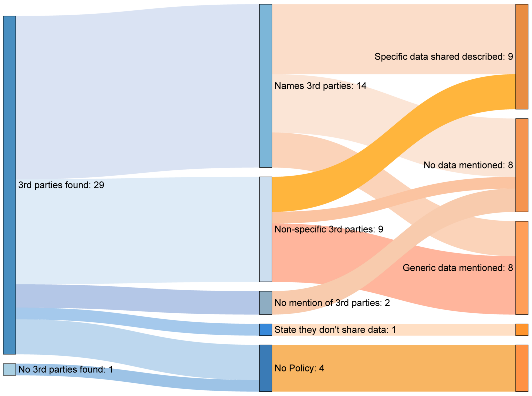


Fig. 5. Sankey diagram representing privacy policies from free apps that embed third parties (1st column), mention them on their privacy policies (2nd column) and describe the data that is shared with them (3rd column).

The rest of the apps (Figure 5) either failed to specify which data would be collected (19), failed to name the specific third parties being involved in data collection (9), failed to recognise that the apps included third parties collecting personal data from the user (6) or even decided to embed these third parties within apps that lacked a privacy policy (3). Most of these apps provided general statements about third parties using language such as "trusted providers" or specifying the services they share the data for (e.g payments, data storage or analytics). In these cases, they just refer the reader to the privacy policies of those service providers. As an example, within the policy for app #5 generic statements are given about “business partners” and “sponsors” as possible recipients of user information. The analysis of its network traffic revealed that the app was sending, in the form of in-app events, all collected data to third party services.

The two paid versions of apps #4 and #8 did not include any of the advertisement libraries that were in their respective free versions. They are therefore within the minority of apps that remove third-party tracking from their paid versions [26]. This could be because the developers are more aware of the implications these libraries have on the privacy of their users. On the other hand, app #8 did not include any mention of third parties in their privacy policy which is the same for both versions. It remains unclear if this is because they are not aware of the data these collect (device identifiers, etc.) or lack awareness of privacy regulations. We contacted the developers in October 2019 to share our results with them. At the time of this writing we have had no response. [5].

When compared to other popular apps [12], menstruapps embedded fewer third party libraries (4 vs 20 on average) but collected potentially more sensitive personal data. This confirms a trend where developers of mobile apps routinely embed data tracking libraries without considering the privacy consequences of those libraries [27].

5 Discussion

Overall, our results show that menstruapps employ slightly better privacy practices than those in health-related apps or the general app ecosystem [12, 31, 53, 57]. Yet, these apps still have a significant number of issues that could lead to serious consequences given the data they handle consists of information like sexual orientation, menstruation and pregnancy.

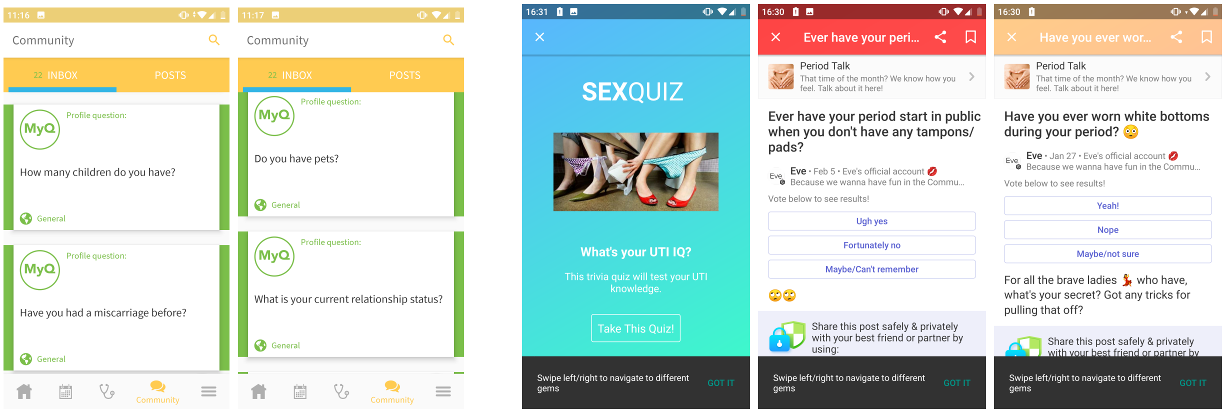
Data Transparency vs Data Greediness

Companies that featured under the *femtech* umbrella – with a focus on the provision of services related to women’s health – tended to have a better understanding of the sensitivity of the data entered into the apps and responded accordingly with their policies and practices. These companies often had more extensive resources behind the design and maintenance of their privacy policies, for example, trust focused teams which were able to aid users in exercising their rights over their data. Apps (#2, #13, #24) went a step further, explaining their practices to their users well, for example in a blog post by deliberately outlining of what happens to period data. Comparatively, apps developed by companies in the same *femtech* space (#14, #15 and #19), often invited their users to share information beyond what was functionally needed. They were also more likely to overreach in their data collection than other apps that did not fit into this category. This was sometimes done by encouraging the users to enter optional data into their profiles (as shown in section 4.4 for #15). In other occasions, this was done in a more subtle way through polls either in the community sections of the apps and often made by employees of the app companies, or in the form of quizzes (Figure 6). This information was sometimes included within the policies as polls or surveys without specifying the kind of data that would be collected through them. Again, it is difficult to see what added benefit the user gains from filling in the information, whereas there is clear benefit and incentive from the developers in terms of further opportunities for data collection. This becomes increasingly worrying given that apps like #15 state in their privacy policy "We will not delete the posts or comments you’ve written and shared publicly, including on social media or in any Community or chat features". The result of this sentence makes it unclear whether data within polls is included in this bracket of irremovable data.

Policies Should Include All App Behaviours

These apps are above average in comparison to other app categories, [31], menstruapps, however, still include complex wording and obfuscating terms within their policies. As happens with most privacy policies, it is not

[5] At the time of camera-ready submission in May 2020



11:16

Community

22 INBOX | POSTS

Profile question:

MyQ

How many children do you have?

General

Profile question:

MyQ

Have you had a miscarriage before?

General

Community

11:17

Community

22 INBOX | POSTS

Profile question:

MyQ

Do you have pets?

General

Profile question:

MyQ

What is your current relationship status?

General

Community

(a) App #15

16:31

SEXQUIZ

What's your UTI IQ?

This trivia quiz will test your UTI knowledge.

Take This Quiz!

Swipe left/right to navigate to different gems GOT IT

16:30

Ever have your peri...

Period Talk

That time of the month? We know how you feel. Talk about it here!

Ever have your period start in public when you don't have any tampons/ pads?

Eve • Feb 5 • Eve's official account

Because we wanna have fun in the Commu...

Vote below to see results!

Ugh yes

Fortunately no

Maybe/Can't remember

🙄🙄

Share this post safely & privately with your best friend or partner by using:

Swipe left/right to navigate to different gems GOT IT

16:30

Have you ever wor...

Period Talk

That time of the month? We know how you feel. Talk about it here!

Have you ever worn white bottoms during your period? 😳

Eve • Jan 27 • Eve's official account

Because we wanna have fun in the Commu...

Vote below to see results!

Yeah!

Nope

Maybe/not sure

For all the brave ladies 💃 who have, what's your secret? Got any tricks for pulling that off?

Share this post safely & privately with your best friend or partner by

Swipe left/right to navigate to different gems GOT IT

(b) App #19

Fig. 6. Screenshots of apps requesting additional information from the user in the form of quizzes, surveys and community posts.

surprising to see the ambiguous term ’may’ being the most frequently used [5, 44]. On the other hand, some policies were too short to provide useful information to the user or potentially employed inapplicable policies replicated from elsewhere. Most of the analysed privacy policies were unable to hit a good balance between providing all the necessary details of their privacy practices whilst doing this in accessible and understandable language that is not oversimplified. This demonstrates the difficulty of interpreting what is needed to fulfil GDPR requirements while providing accessible privacy policies for users [37, 55]. Our research therefore corroborates with work done by others in highlighting the ineffectiveness of privacy policies, as they are developed today, as a method of communicating these practices [10].

Despite these difficulties, apps like #2 demonstrate that, when spending the necessary resources and effort, companies can provide reasonably good privacy policies. The best intentions, however, can still result in some aspects being overlooked. App developers can communicate with their legal team to produce a privacy policy that communicates what the app is capturing and its purpose. Describing what third party libraries capture is more difficult. Previous studies on the usage of third party libraries in Android show that most developers do not verify the security of third party libraries before embedding them on their apps [17]. On a similar note, trends on third party library usage by developers show that, overall, developers do not take into account the amount of data a library gathers when embedding them into their apps [6]. In contrast, an analysis of apps that embedded the Facebook SDK revealed that, understanding and controlling the amount of data these libraries collect from their users is difficult, even when developers want to control how these libraries behave and collect data [43]. This, which could be expected from ad libraries [6], and affects analytics libraries [34]. In the particular case of app #2, it was unclear if this happened because they were not aware of the data collection or they favoured the functionality the library provided. Still, developers should be aware of the amount of data third party libraries collect when embedding them into their apps, and reflect this in their privacy policies.

Past research in the effects of new legislation such as GDPR, shows privacy policies now provide more information to the user. Yet, that does not result in them being more accessible, giving the user more control mechanisms or describing all the behaviours of the service provider accurately [16, 25]. Our study demonstrates this is also true for the menstruapp space, even in those cases were companies have dedicated privacy teams.

Period Data is not Considered Sensitive

Overall 66% of apps failed to specify that they collected period data. Of these, 14 sent this data to their servers and 3 sent this to third parties as in-app events. Although 7 of them discuss this in a general light, they fail to outline the specific data that they require the user to provide while setting up the app. One of the most prominent cases was app #7. This app required five different types of menstruation data (period, cycle, luteal phase and ovulation lengths as well as dates of previous periods), that were all uploaded to their servers, yet they did not mention period data of any kind in its policy. In fact, only 6 of the apps explicitly mentioned that they required period dates within their respective privacy policies. This points to a worrying trend where developers consider period and sexual data as simply another piece of data rather than health data that is sensitive in nature. There are however, examples of pharmaceutical companies and health insurers who have registered interest in acquiring this data, suggesting it is far from being another piece of PII [52].

6 Limitations

Part of our methodology includes the execution of qualitative analysis that have to be executed manually. Whilst we used these methods to ensure accuracy in our results, we also acknowledge that this limits the scalability of our analysis if it were to be applied to the whole menstruapp ecosystem. Our app selection mechanism tries to tackle this by prioritising those with most downloads. With this, our 30 app analysis is able to cover more than 90% of the ecosystem in terms of downloads, even when there are more than 200 apps in the store related to period tracking.

Our library analysis uses *LibScout* [3] which relies on a pre-obtained library database. We use a library database from March 2019. This could result in recent libraries not being detected. We use network captures to identify domains of libraries that may have not been detected via *LibScout*. Although we were able to identify all domains in these traces, there were some of the contents being sent to some servers that used a non-standard encoding, so we were not able to inspect them. In all cases except one, the apps were already sending all the data out via other means. We marked the case we were not able to inspect with *CD* in table 2.

Apps in the mobile ecosystem are frequently updated with new features, etc. These changes also affect privacy policies in general. For this purpose, we have uploaded the privacy policies hashes of apps used to perform this analysis to an online repository so these can be viewed and used in further research (https://github.com/guizos/HowPrivateIsYourPeriodData).

7 Conclusion

In this paper, we have presented a comprehensive study of the privacy and information practices of menstruapps. The apps included within this study account for at least 178 million Android downloads. We used a mixed-methods approach to perform a detailed review of the privacy practices of 30 menstruapps. Our results show that, while the industry follows slightly better practices than the general app ecosystem, it also has several areas where their practices can improve. As it stands, the information provided to users about these practices is not always straightforward and, in general, they lack clear instructions on how users can exercise their privacy rights. None of the privacy policies could be considered easy to understand. While this is in line with previous research [31], these policies should be straightforward, informative and transparent. This is to ensure that users understand the implications of sharing sensitive data, such as period cycles or sex data with the apps, fully consent to how it is used, and know how to exercise any of their rights should they want to. In particular, our research demonstrates that, in many cases, app developers fail to consider required menstruation and sex data as specially sensitive, mentioning only standard PII within their privacy policies. We found that despite this, companies that solely focus on developing women-centred services tend to be better at depicting and explaining their privacy practices to users. Third party libraries are still a problem in terms of data collection. The best example for this is app #2. This application was above the best apps in most metrics. However, they embed an analytics library, that transmitted sensitive information when collecting in-app events.

Our results show that there are two main ways of improving the privacy policies of menstruapps in general. First, as with other kinds of apps [31], developers should try to simplify the language of their privacy policies. Second, and more in line with this category of app, developers should pay special attention at the data they collect and how this is used within their apps and their third party libraries. Menstruapps privacy policies lack the necessary level of detail to accurately describe how the app behaves in relation to their data. To encourage this, we have already contacted each developer and provided them with these results and specific recommendations to improve their privacy practices.

Acknowledgements

This research has been partially sponsored by the Engineering and Physical Sciences Research Council (EPSRC) and the UK government as part of the Centre for Doctoral Training in Cyber Security at Royal Holloway, University of London (EP/P009301/1). We thank Katherine Brickell for her comments and suggestions an earlier version of this paper. We thank Pallavi Sivakumaran for her support during the app traffic analysis.

References

[1] Adhikari, R., Richards, D., and Scott, K. (2014). Security and privacy issues related to the use of mobile health apps. ACIS.

[2] Ahmed, E. (2019). No body's business but mine: How menstruation apps are sharing your data. https://privacyinternational.org/long-read/3196/no-bodys-business-mine-how-menstruation-apps-are-sharing-your-data. Accessed on March 2020.

[3] Backes, M., Bugiel, S., and Derr, E. (2016). Reliable third-party library detection in android and its security applications. In *Proceedings of the 2016 ACM SIGSAC Conference on Computer and Communications Security*, pages 356–367. ACM.

[4] Balebako, R., Marsh, A., Lin, J., Hong, J. I., and Cranor, L. F. (2014). The privacy and security behaviors of smartphone app developers.

[5] Bhatia, J., Breaux, T. D., Reidenberg, J. R., and Norton, T. B. (2016). A theory of vagueness and privacy risk perception. In *2016 IEEE 24th International Requirements Engineering Conference (RE)*, pages 26–35. IEEE.

[6] Book, T., Pridgen, A., and Wallach, D. S. (2013). Longitudinal analysis of android ad library permissions. *arXiv preprint arXiv:1303.0857*.

[7] Bowers, J., Reaves, B., Sherman, I. N., Traynor, P., and Butler, K. (2017). Regulators, mount up! analysis of privacy policies for mobile money services. In *Thirteenth Symposium on Usable Privacy and Security (SOUPS 2017)*, pages 97–114.

[8] Brunton, F. and Nissenbaum, H. (2015). *Obfuscation: A user's guide for privacy and protest*. Mit Press.

[9] Burkell, J. and Fortier, A. (2013). Privacy policy disclosures of behavioural tracking on consumer health websites. In *Proceedings of the 76th ASIS&T Annual Meeting: Beyond the Cloud: Rethinking Information Boundaries*, page 56. American Society for Information Science.

[10] Cate, F. H. (2010). The limits of notice and choice. *IEEE Security & Privacy*, 8(2):59–62.

[11] Center for Devices and Radiological Health (2018). FDA allows marketing of first direct-to-consumer app for contraceptive use to prevent pregnancy. https://www.fda.gov/news-events/press-announcements/fda-allows-marketing-first-direct-consumer-app-contraceptive-use-prevent-pregnancy. Accessed on August 2019.

[12] Claesson, A. and Bjørstad, T. E. (2020). Out of control - a review of data sharing by popular mobile apps. https://fil.forbrukerradet.no/wp-content/uploads/2020/01/mnemonic-security-test-report-v1.0.pdf. Accessed on January 2020.

[13] Council of European Union (2016). 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). https://eur-lex.europa.eu/legal-content/EN/TXT/?uri=CELEX%3A32016R0679.

[14] Cranor, L. F., Leon, P. G., and Ur, B. (2016). A large-scale evaluation of us financial institutions' standardized privacy notices. *ACM Transactions on the Web (TWEB)*, 10(3):1–33.

[15] Das, G., Cheung, C., Nebeker, C., Bietz, M., and Bloss, C. (2018). Privacy policies for apps targeted toward youth: Descriptive analysis of readability. *JMIR Mhealth Uhealth*, 6(1):e3.

[16] Degeling, M., Utz, C., Lentzsch, C., Hosseini, H., Schaub, F., and Holz, T. (2019). We value your privacy ... now take some cookies: Measuring the gdpr's impact on web privacy. *Proceedings 2019 Network and Distributed System Security Symposium*.

[17] Derr, E., Bugiel, S., Fahl, S., Acar, Y., and Backes, M. (2017). Keep me updated: An empirical study of third-party library updatability on android. In *Proceedings of the 2017 ACM SIGSAC Conference on Computer and Communications Security*, CCS '17, page 2187–2200, New York, NY, USA. Association for Computing Machinery.

[18] Epstein, D. A., Lee, N. B., Kang, J. H., Agapie, E., Schroeder, J., Pina, L. R., Fogarty, J., Kientz, J. A., and Munson, S. (2017). Examining menstrual tracking to inform the design of personal informatics tools. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems*, pages 6876–6888. ACM.

[19] Felizi, N. and Varon, J. (2017). Menstruapps - how to turn your period into money (for others). https://chupadados.codingrights.org/en/menstruapps-como-transformar-sua-menstruacao-em-dinheiro-para-os-outros/. Accessed on August 2019.

[20] Flemings, M., Kazmi, S., Pak, R., and Shaer, O. (2018). Crimson wave: Shedding light on menstrual health. In *Proceedings of the Twelfth International Conference on Tangible, Embedded, and Embodied Interaction*, TEI '18, page 343–348, New York, NY, USA. Association for Computing Machinery.

[21] Flesch, R. (1979). How to write plain english: Let's start with the formula. *University of Canterbury*.

[22] Frost & Sullivan (2018). Femtech - time for a digital revolution in the women's health market. https://ww2.frost.com/frost-perspectives/femtechtime-digital-revolution-womens-health-market/ Accessed on May 2020.

[23] Gilding, K. (2020). Which femtech apps can you trust? https://www.medicalplasticsnews.com/news/which-femtech-apps-can-you-trust/ Accessed on April 2020.

[24] Gluck, J., Schaub, F., Friedman, A., Habib, H., Sadeh, N., Cranor, L. F., and Agarwal, Y. (2016). How short is too short? implications of length and framing on the effectiveness of privacy notices. In *Twelfth Symposium on Usable Privacy and Security (SOUPS 2016)*, pages 321–340.

[25] Habib, H., Zou, Y., Jannu, A., Sridhar, N., Swoopes, C., Acquisti, A., Cranor, L. F., Sadeh, N., and Schaub, F. (2019). An empirical analysis of data deletion and opt-out choices on 150 websites. In *Fifteenth Symposium on Usable Privacy and Security (SOUPS 2019)*.

[26] Han, C., Reyes, I., Elazari Bar On, A., Reardon, J., Feal, Á., Bamberger, K. A., Egelman, S., and Vallina-Rodriguez, N. (2019). Do you get what you pay for? comparing the privacy behaviors of free vs. paid apps. In *The Workshop on Technology and Consumer Protection (ConPro'19)*.

[27] HM Government (2019). Online harms white paper. https://www.gov.uk/government/consultations/online-harms-white-paper. Accessed on August 2019.

[28] Honnibal, M. and Johnson, M. (2015). An improved non-monotonic transition system for dependency parsing. In *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing*, pages 1373–1378, Lisbon, Portugal. Association for Computational Linguistics.

[29] Huckvale, K., Prieto, J. T., Tilney, M., Benghozi, P.-J., and Car, J. (2015). Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC medicine*, 13(1):214.

[30] Hutton, L., Price, B. A., Kelly, R., McCormick, C., Bandara, A. K., Hatzakis, T., Meadows, M., and Nuseibeh, B. (2018). Assessing the privacy of mhealth apps for self-tracking: heuristic evaluation approach. *JMIR mHealth and uHealth*, 6(10):e185.

[31] Jensen, C. and Potts, C. (2004). Privacy policies as decision-making tools: an evaluation of online privacy notices. In *Proceedings of the SIGCHI conference on Human Factors in Computing Systems*, pages 471–478. ACM.

[32] Levy, J. (2018). Of mobiles and menses: Researching period tracking apps and issues of response-ability. *Studies on Home and Community Science*, 11(2):108–115.

[33] Li, L., Bissyandé, T. F., Klein, J., and Le Traon, Y. (2016). An investigation into the use of common libraries in android apps. In *2016 IEEE 23rd International Conference on Software Analysis, Evolution, and Reengineering (SANER)*, volume 1, pages 403–414. IEEE.

[34] Liu, X., Liu, J., Zhu, S., Wang, W., and Zhang, X. (2019). Privacy risk analysis and mitigation of analytics libraries in the android ecosystem. *IEEE Transactions on Mobile Computing*.

[35] Lupton, D. (2015). 'mastering your fertility': The digitised reproductive citizen. *Chapter for Negotiating Digital Citizenship: Control, Contest and Culture, edited by Anthony McCosker, Sonja Vivienne and Amelia Johns. To be published by Rowman and Littlefield, London. Forthcoming*.

[36] Mahdawi, A. (2019). If the government tracks women's periods, why not track male ejaculation, too? https://fortune.com/2014/08/27/how-max-levchins-glow-app-got-25000-women-pregnant/ Accessed on March 2020.

[37] McDonald, A. M. and Cranor, L. F. (2008). The cost of reading privacy policies. *Isjlp*, 4:543.

[38] Mcdonald, A. M., Reeder, R. W., Kelley, P. G., and Cranor, L. F. (2009). A comparative study of online privacy policies and formats. In *International Symposium on Privacy Enhancing Technologies Symposium*, pages 37–55. Springer.

[39] Moglia, M. L., Nguyen, H. V., Chyjek, K., Chen, K. T., and Castaño, P. M. (2016). Evaluation of smartphone menstrual cycle tracking applications using an adapted applications scoring system. *Obstetrics & Gynecology*, 127(6):1153–1160.

[40] Morrissey, J. (2018). Women struggling to get pregnant turn to fertility apps. https://www.nytimes.com/2018/08/27/business/women-fertility-apps-pregnancy.html. Accessed on August 2019.

[41] Pollach, I. (2005). A typology of communicative strategies in online privacy policies: Ethics, power and informed consent. *Journal of Business Ethics*, 62(3):221.

[42] Pollach, I. (2007). What's wrong with online privacy policies? *Communications of the ACM*, 50(9):103–108.

[43] Privacy International (2018). How apps on android share data with facebook (even if you don't have a facebook account). https://privacyinternational.org/report/2647/how-apps-android-share-data-facebook-report. Accessed on March 2020.

[44] Reidenberg, J. R., Bhatia, J., Breaux, T. D., and Norton, T. B. (2016). Ambiguity in privacy policies and the impact of regulation. *The Journal of Legal Studies*, 45(S2):S163–S190.

[45] Reidenberg, J. R., Breaux, T., Cranor, L. F., French, B., Grannis, A., Graves, J. T., Liu, F., McDonald, A., Norton, T. B., and Ramanath, R. (2015). Disagreeable privacy policies: Mismatches between meaning and users' understanding. *Berkeley Tech. LJ*, 30:39.

[46] Rizk, V. and Othman, D. (2016). Quantifying fertility and reproduction through mobile apps: A critical overview. *Arrow for change*, 22(1):13–21.

[47] Rosas, C. (2019). The future is femtech: Privacy and data security issues surrounding femtech applications. *Hastings Business Law Journal*, 15(2):319.

[48] Schechner, S. (2019). You give apps sensitive personal information. then they tell facebook. https://www.wsj.com/articles/you-give-apps-sensitive-personal-information-then-they-tell-facebook-11550851636. Accessed on August 2019.

[49] Scott, K. M., Gome, G. A., Richards, D., and Caldwell, P. H. (2015). How trustworthy are apps for maternal and child health? *Health and Technology*, 4(4):329–336.

[50] Sen, P. (2014). How max levchin's glow app got 25,000 women pregnant. https://fortune.com/2014/08/27/how-max-levchins-glow-app-got-25000-women-pregnant/ Accessed on March 2020.

[51] Sheng, X. and Cranor, L. F. (2005). An evaluation of the effect of us financial privacy legislation through the analysis of privacy policies. *ISJLP*, 2:943.

[52] Steel, E. and Dembosky, A. (2013). Health apps run into privacy snags. https://www.ft.com/content/b709cf4a-12dd-11e3-a05e-00144feabdc0/ Accessed on May 2020.

[53] Sunyaev, A., Dehling, T., Taylor, P. L., and Mandl, K. D. (2014). Availability and quality of mobile health app privacy policies. *Journal of the American Medical Informatics Association*, 22(e1):e28–e33.

[54] Weigel, M. (2016). 'Fitbit for your period': the rise of fertility tracking. https://www.theguardian.com/technology/2016/mar/23/fitbit-for-your-period-the-rise-of-fertility-tracking. Accessed on August 2019.

[55] Wilson, S., Schaub, F., Liu, F., Sathyendra, K. M., Smullen, D., Zimmeck, S., Ramanath, R., Story, P., Liu, F., Sadeh, N., et al. (2018). Analyzing privacy policies at scale: From crowdsourcing to automated annotations. *ACM Transactions on the Web (TWEB)*, 13(1):1–29.

[56] Yu, L., Luo, X., Liu, X., and Zhang, T. (2016). Can we trust the privacy policies of android apps? In *2016 46th Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN)*, pages 538–549. IEEE.

[57] Zimmeck, S., Story, P., Smullen, D., Ravichander, A., Wang, Z., Reidenberg, J., Russell, N. C., and Sadeh, N. (2019). Maps: Scaling privacy compliance analysis to a million apps. *Proceedings on Privacy Enhancing Technologies*, 2019(3):66–86.

Appendices

Appendix A: Downloads and origin of developer by app

ID	App. Name	Package name	Downl.	Developer	Country
1	My Calendar	com.lbrc.PeriodCalendar	100 mill.	Simple Design Ltd.	Hong Kong
2	Clue	com.clue.android	10 mill.	BioWink GmbH	Germany
3	Flo	org.iggymedia.periodtracker	10 mill.	OWHEALTH, INC.	USA
4	P.Tracker	com.period.tracker.lite	10 mill.	GP International LLC	USA
5	Maya	in.plackal.lovecyclesfree	5 mill.	Plackal Tech	India
6	My Calendar	com.popularapp.periodcalendar	5 mill.	SimpleInnovation	USA
7	Lilly Tracker	com.smsrobot.period	5 mill.	SMSROBOT LTD	Ireland
8	WomanLog	com.womanlog	5 mill.	Pro Active App	Latvia
9	LadyTimer	com.ladytimer.ovulationcalendar	5 mill.	Vipos Apps	USA
10	MyDaysX	com.chris.android.mydaysfree	5 mill.	Single developer*	Germany
11	M. Calendar	com.g*r.mc	5 mill.	witiz.	China
12	Petal	com.go.flo	1 mill.	Blue Group	China
13	Amila	com.periodapp.period	1 mill.	Amila	Canada
14	Glow	com.glow.android	1 mill.	Glow Inc.	USA
15	Ovia	com.ovuline.fertility	1 mill.	Ovia Health	USA
16	MC	org.medhelp.mc	1 mill.	Staywell	USA
17	My Tracker	periodtracker.pregnancy.ovulationtracker	1 mill.	Leap Fitness Group	Hong Kong
18	Once	net.android.wzworks.magicday	1 mill.	Malang	South Korea
19	Eve	com.glow.andriod.eve	1 mill.	Glow Inc.	USA
20	Period Tracker	com.northpark.periodtracker	1 mill.	Leap Fitness Group	Hong Kong
21	MyPeriodTracker	com.linchpin.myperiodtracker	1 mill.	Linchpin Health	India
22	My Calendar	com.brc.PeriodTrackerDiary	0.5 mill.	Living Better	USA
23	Luna	com.period.cal	0.5 mill.	tinyChangeCompany	No country found
24	Natural Cycles	com.naturalcycles.cordova	0.5 mill.	NaturalCycles AG.	Sweden
25	Cherry	com.period.tracker.menstrual.cycle.cherry	0.5 mill.	Rosa Care Group	USA
26	P. Tracker for Women	women.periods.periods_for_women	0.5 mill.	Global Tech Social	No country found
27	P. Tracker	com.cg.android.ptracker	0.1 mill.	Sevenlogics, INC.	USA
28	PD	org.nanobit.perioddiary	0.1 mill.	Nanobit Games	Croatia
29	Wonder Period	wonder.period.tracker.ovulation.calculator	0.1 mill.	Wonder App Studio	Hong Kong
30	Menstrual Cycle	ccom.fourricegroup.mc	0.1 mill.	4Rice Group	Vietnam

*Corresponds or includes to a person's full name. This has been omitted for privacy reasons.

Appendix B: Types of information collected by the app and mentioned by the privacy policy

	App ID	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20	21	22	23	24	25	26	27	28	29	30
PII	Name	N ■	N ■	O ■	N	O ■	O	N ■	N	O	O	N	N	N	R ■	O ■	O ■	O	N ■	O ■	N	N	O	N	R ■	N ■	N	N	N	R ■	N
	DOB	N	R	R ■	N	O	N	N ■	N	O	N	N	O	N	O ■	O ■	N	N	O ■	O ■	N	N	N	O	R ■	N	N	N	N	N	N
	Photos	N	N	O	N	O	N	O ■	N	O	O	N	N	N	O ■	O ■	N	O	N	O ■	N	N	N	N	N	N	N	N	N	O	N
	Sexual Orientation	N	N	N	N	N	N	N	N	N	N	N	N	N	O	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	Gender	N	N	N ■	N	N	N	N ■	N	N	N	N	N	N	O ■	N	N	N	O ■	O ■	N	N	N	N	N	N	N	N	N	N	N
	Bio	N	N	N	N	O	N	N	N	O	N	N	N	N	N	N	N	N	N	O	N	N	N	N	N	N	N	N	N	N	N
	Location	N	N	N ■	N	O	N ■	N	N	O	N	N	N	N	O	N	N	N	N	O	N	N	N ■	N	N	N	N	N	N	N ■	N
	Relationship Status	N	N	N	N	N	N	N	N	N	N	N	N	N	O	N	N	N	N	O	N	N	N	N	N	N	N	N	N	N	N
	Insurance	N	N	N	N	N	N	N	N	N	N	N	N	N	O	O ■	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	Occupation	N	N	N ■	N	N	N	N	N	N	N	N	N	N	O	O ■	N ■	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	Zipcode	N	N	N ■	N	N	N	N	N	N	O	N	N	N	O ■	O	N	N	N	N ■	N	N	N	N	O	N	N	N	N	N	N
	Ethnicity	N	N	N	N	N	N	N	N	N	O	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Period Data	Period Dates	R	R ■	R ■	R	R	R	R	R	O	O	O	R	R ■	R ■	R ■	R	R	R	R ■	R	R	R	R	R	R	R	R	R	R	O
	Cycle Length	R	R	R	R	R	R	R	R	R	O	R	R	R ■	R	R ■	R	R	R	R	R	R	R	R	N	R	R	R	O	R	O
	Period Length	R	R ■	R	R	R	R	R	R	R	O	R	R	R ■	R	R	R	R	R	R	R	R	R	R	N	R	R	R	O	R	O
	Luteal Phase Length	O	N	O	N	N	O	R	R	O	O	R	N	N	N	N	N	O	N	N	O	O	R	N	N	N	N	N	O	O	N
	Ovualation Length	O	O	N	R	N	O	R	N	O	O	N	N	N ■	N	N	N	N	N	N	N	O	N	N	N	N	N	N	O	N	O
	Flow	O	O ■	O	O	O	O	O	O	O	O	O	O	N	O	O	O	O	N	O	O	N	O	O	O	O	N	O	N	O	N
	Fluid	O	O	O	N	O	O	N	O	O	O	N	O	N	O	O	O	O	N	O	O	N	N	O	N	N	N	N	N	N	N
	Cycle Regularity	O	N	N	N	N	N	N	N	N	N	N	N	N	O	R	N	N	N	O	N	N	N	N	N	N	N	N	N	N	N
	Cervix	O	N	N	N	N	O	N	N	N	O	O	N	N	O	O	O	N	N	N	O	N	N	N	N	N	N	N	N	N	N
	Track Cycles	N	O	O	N	O	N	N	N	O	N	N	O	N	N	O	N	N	N	N	N	N	N	O	N	N	N	N	N	O	N
	Trying to Conceive	N	O	O	O	O	N	N	N	O	N	N	O	N	O	O ■	O	N	O	N	N	N	N	O	O ■	N	N	O	N	O	N
	Avoiding Pregnancy	N	N	N	N	O	N	N	N	O	N	N	N	N	O	N	N	N	N	N	N	N	N	N	O	N	N	N	N	N	N
	Tracking Pregnancy	O	O	O	O	O	N	N	O	O	O	O	O	N	O	O	N	O	N	N	O	O	N	O	O	N	N	N	O	N	N
	Symptoms	O	O ■	O ■	O	O ■	O	O	O	O	O	O	O	N	O	O	O	O	O	O	O	O	O	O	N	N	N	O	O	O	N
	Moods	O	O ■	O ■	O	O ■	O	O	O	O	O	O	O	N	O	O ■	O	O	N	O	O	O	O	O	N	O	N	O	O	O	N
Sex Logging	When	O	O ■	O	O	O	O	O	O	O	O	O	O	O ■	O	O	O	O	O	O	O	O	O	O	O	O	N	O	O	O	O
	Abstinence	N	N	O	O	N	N	N	N	N	N	N	N	N	O	O	N	N	N	O	N	N	N	N	N	O	N	N	N	N	N
	Protected	O	O	O	O	O	O	O	O	O	O	O	O	N	N	N	N	O	O	O	O	N	O	O	O	O	N	O	N	O	N
	Unprotected	O	O	O	O	O	O	O	O	O	N	O	O	N	N	N	N	N	N	O	O	N	O	O	O	O	N	O	N	O	N
	Masturbation	N	N	O	N	N	N	N	O	N	N	N	N	N	N	N	N	N	N	O	N	N	N	N	N	N	N	N	N	N	N
	Times	O	N	N	N	N	N	N	O	O	N	N	O	N	O	N	N	O	N	O	O	N	N	O	N	N	N	N	N	N	N
	Orgasm	O	N	N	N	N	N	N	O	N	O	N	O	N	N	N	N	O	N	O	O	N	N	O	N	N	N	N	N	N	N
	Sex drive	N	O	O	N	N	N	N	O	O	N	N	N	N	O	O	N	N	N	O	N	N	N	N	N	N	N	N	N	N	N
	Withdraw	N	O	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	O	N
	Contraceptive Pill	N	O	O	N	O	O	O	O	N	O	O	O	N	R	N	O	O	O	O	O	N	O	O	N	N	N	N	N	N	O
	Other contraceptive	O	O	N	N	O	N	O	N	N	O	O	N	N	R	N	O	O	N	O	N	N	N	N	N	N	N	N	N	N	N
Collection	Tampon	N	O	O	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	Sanitary Towel	N	O	O	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	Panty Liner	N	O	O	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	Menstrual Cup	N	O	O	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	Several	N	N	O	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
Test	Pregnancy	N	N	N ■	N	N	N	O	N	N	O	O	O	N	O	O	N	O	N	N	O	N	N	O	N ■	N	N	N	N	O	N
	Ovulation	O	N	N	N	N	N	O	N	O	O	O	O	N	O	O	O	O	N	N	O	N	N	O	R ■	N	N	N	N	O	N
Health Information	Sleep	N	O	O	O	N	N	N	N	O	N	N	O	N	O	O	O	O	N	N	N	N	N	O	N	N	N	N	N	O	N
	Water	N	N	O	O	N	N	N	N	N	N	N	O	N	N	O	N	O	N	N	N	N	N	O	N	N	N	N	N	O	N
	Exercise	N	O	O	O	N	N	N	N	O	N	N	O	N	O	O	O	N	N	O	N	N	N	O	N	N	N	N	N	O	N
	Energy	N	O	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	Conditions	N	O	O	N	N	N	N	N	N	N	N	N	N	O ■	O	N	N	N	N ■	N	N	N	N	N	N	N	N	N	N	N
	Blood pressure	N	N	N	N	N	N	O	N	O	N	N	N	N	N	O	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	Food	N	O ■	O	O	N	N	N	N	N	N	N	O	N	N	O	N	N	N	N	N	N	N	O	N	N	N	N	N	O	N
	Stool/Digestion	N	O	N	N	N	N	N	N	N	N	N	O	N	N	N	N	N	N	N	N	N	N	O	N	N	N	N	N	O	N
	Skin/Hair	N	O ■	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N
	Appointments	N	O	N	N	O	N	N	N	O	N	O	N	N	N	N	O	N	O	N	N	N	N	N	N	N	N	N	N	N	N
	Medicines	O	O	N	N	N	O	N	O	N	N	O	N	N	O	O	O	N	N	N	O	O	O	N	N	N	N	O	N	N	N
	Bad Habits	N	N	O	N	N	N	N	N	N	N	N	O	N	O	N	N	N	N	O	N	N	N	O	N	N	N	N	N	N	N
	More	O	O ■	O	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	N	O	N
	Weight	O	O ■	O ■	O	O ■	O	N	O	O	O	O	O	N	O ■	O ■	O	O	O	O ■	O	O	O	O	O ■	N	N	O	O	O	N
	Temperature	O	O ■	O ■	O	O ■	O	O	O	O	O	O	O	N	O ■	O	N	O	N	O ■	O	O	O	O	R ■	N	N	O	O	O	N
	Height	O	O	O	O	N	N	N	N	N	N	N	O	N	O	O ■	N	N	N	O	O	O	N	O	O ■	N	N	N	N	N	N

N = Not required. R = Required. O = Optional. ■ = Explicitly mentioned in the privacy policy.

Appendix C: Words enabling language obfuscation

The following list of words was used to identify sentences trying to downplay or modify the emphasis of certain statements within the privacy policies, reducing their clarity and making them difficult to interpret (see Section 3.6). This kinds of words have been previously used in [9, 41]. The words in this list were first extracted from the privacy policies with *spaCy*, mapped to its corresponding category (adjective, adverb or verb) using *spaCy's* model and then manually reviewed. For verbs, only those sentences having the user, the developers, the app name, the company name, *we* or *you* as subject were selected.

- **Adjectives**: absolute, acceptable, accurate, adequate, affirmative, aggregate, aggregated, alternative, anonymized, appropriate, approximate, associated, authorized, automated, automatic, automattic, better, broad, broader, certain, clear, comfortable, committed, compelling, competent, complete, compliant, comprehensive, correct, current, customized, daily, designated, desirable, direct, disproportionate, easier, effective, efficient, eligible, enforceable, equivalent, essential, explicit, friendly, fundamental, general, good, great, hard, identified, immediate, important, impossible, impractical, improved, inaccurate, inappropriate, incompatible, incomplete, inconsistent, informational, intended, interested, interesting, legible, legitimate, mandatory, meaningful, mindful, minimum, most, natural, necessary, occasional, only, optimal, optional, original, particular, perfect, periodic, periodical, permissible, persistent, possible, potential, precise, predictive, preferred, prominent, proper, prospective, protective, qualified, real, reasonable, regular, related, relative, reliable, required, respective, responsive, safer, satisfied, seamless, sensitive, separate, significant, similar, single, sole, special, specific, specified, standard, statutory, strict, sufficient, suitable, supervisory, supplementary, systematic, targeted, timely, transparent, unaccepted, unambiguous, unauthorized, uncertain, understandable, unique, unlawful, unlikely, unnecessary, unsupervised, useful, usual, vague, valid, verifiable, visible, vital, voluntary.
- **Adverbs:** above, absolutely, accordingly, additionally , affirmatively, already, alternatively, always, anytime, automatically, back, before, better, carefully, commonly, completely, contextually, continually, correctly, currently, directly, effectively, efficiently, entirely, even, ever, exclusively, explicitly, expresslly, extremely, faster, first, formerly, forward, frequently, fully, further, generally, globally, historically, immediately, importantly, incredibly, indefinitely, independently, indirectly, individually, intentionaly, just, knowingly, last, later, least, legally, less, locally, long, longer, mainly, manually, maybe, might, more, most, mostly, much, never, newly, normally, occasionally, often, once, only, onward, optionally, otherwise, particularly, partly, periodically, permanently, potentially, precisely , predominantly, previously, primarily, prior, probably, prominently, promptly, properly, quickly, really, reasonably, regardless, regularly, remotely, routinely, safely, satisfactorily, separately, significantly, similarly, simply, solely, sometimes, soon, specially, strictly, strongly, subsequently, substantially, substantialy, successfully, sufficiently, tipically, typically, uniquely, unreasonably, usually, very, voluntarily, well, whenever, wherever.
- **Verbs:** could, would may, might, can.