Supporting informed, idiosyncratic privacy preferences through personalised exposure profiles

Reuben Binns
Max Van Kleek
Jun Zhao
Ulrik Lyngs
Nigel Shadbolt
University of Oxford
Oxford, United Kingdom
reuben.binns@cs.ox.ac.uk
max.van.kleek@cs.ox.ac.uk
jun.zhao@cs.ox.ac.uk
ulrik.lyngs@cs.ox.ac.uk
nigel.shadbolt@cs.ox.ac.uk

Abstract
Are there ‘good’ privacy decisions that can be considered universally appropriate for end-users? In this paper, we draw together findings from previous work on privacy attitudes regarding tracking of end-users by third parties via smartphone applications. In particular, we highlight a range of highly idiosyncratic privacy preferences which are elicited only after people have had the chance to explore and reflect upon detailed, complex and personalised information about their data flows.

Author Keywords
privacy, tracking, preferences, individual differences

Introduction
Would you rather a small number of companies knew a lot about you, or a large number of companies knew a little about you? Would you prefer your data be kept in certain jurisdictions, such as the US, EU, or China? Would you trust a small company with your data if you found out it is a subsidiary of a large company? What would you think of your data being used exclusively for targeted online behavioural advertising of only certain kinds of products?

These are not choices that end-users are commonly faced with explicitly in their everyday lives. Most people have probably never thought about them, and it is difficult to
know what they would choose if they had.

However, over the course of several studies in which participants have been presented with an overview of their information exposure and given a chance to reflect, we have observed complex preferences such as these frequently being expressed.

**Background**

Even if they are vaguely aware that their data may be being collected via their interaction with websites and mobile apps, most people are probably unaware of the specifics. In order to address this lack of knowledge, previous work has explored the design of ‘privacy indicators’, which aim to present important privacy-related aspects of services by displaying information in a summary form [4, 5, 6, 8, 7, 1]. Privacy indicators aim help users to understand the privacy implications of a service when making a choice about whether to use it, by showing what kinds of data are collected, for what purposes.

However, showing people information about the privacy-relevant features of apps so that they can make informed privacy choices involves significant end-user burden; each app must be evaluated and action taken accordingly. In response to this problem, various proposals aim to automatically infer people’s privacy preferences from their behaviour, and then have those preferences automatically enforced by software agents (for instance, Liu et al.’s *Privacy Assistant* [9].

While such approaches show significant promise in reducing the user burden of dealing with existing permission models, we believe they do not encompass the full picture; there may be a range of latent complex and idiosyncratic privacy preferences that are not elicited in the context of users’ existing everyday privacy choices.

Such attitudes may only come to light after people have been exposed to richer, more holistic and personalised privacy-related information.

Recent work has explored eliciting such attitudes through a variety of transparency-enhancing interfaces. In particular, unlike previous studies which focused primarily on first-party data collection by apps, recent work by the authors [12, 3, 11], inspired by Balebako et al [2], focuses on revealing the specific identities of multiple third parties associated with an app, and including contextual information about the companies, their locations and business models. Over the course of 3 qualitative studies, featuring interviews with 59 participants and an online survey, we have observed that reflecting on such information about their exposure to multiple third parties can bring different kinds of privacy concerns to the surface. While many of the broad concerns and goals—such as a feeling of creepiness at being tracked, and a desire to limit exposure—were common across all users, participants often came up with very different priorities and strategies when asked to explain and explore in detail.

In each of these studies, participants explored their existing exposure to data tracking via smartphone apps, using prototype transparency tools (see figure 1, described in [12, 11]).

**Familiarity breeds contempt?**

One common feature participants raised in response to different trackers was their level of familiarity with the company. As in previous studies [10, 2], we found that trust in companies or entities depended on users’ familiarity with them. Companies producing end-user goods and services were more likely to be deemed
trustworthy than those that did not, including those only providing app back-end services [10].

For example, one participant discussed the importance of recognising a specific brand “there’s some comfort in Google being a brand that you recognise, as opposed to The [company name] Company. Like, who are they?” P9.

Familiar names also proved important when participants considered contextual information about parent-subsidiary relationship. In one case it was revealed that one of the third party advertising networks associated with a service (AdMob) was a subsidiary of Google. For many participants, this changed the way they felt about AdMob, both positively and negatively. Some felt reassured upon learning that AdMob was owned by a more established company, because the latter’s “data policy is scrutinized internally and externally, so I’d feel more comfortable with them having data than an independent company”.

For others, familiarity bred contempt; finding out that AdMob is owned by Google was a reason to avoid the service, to prevent their data being “sucked into the all encompassing data mountain of Google and attached to my very persistent Google identity” (P18).

**Exposure minimisation**

Relatedly, in response to observing their overall exposure, some participants formulated what might be called an ‘exposure minimisation’ strategy. Under this strategy, their choices were informed not just by the privacy practices of a particular app, but also mediated by beliefs about which entities might have access to the data in question already. In this sense, many data flows may be ‘overdetermined’; the same organisation may have access to the same data about an individual through multiple different channels. A consequence of this strategy is that a user might rationally prefer an app that shares data with a greater number of third parties, if they believe that their data has already been shared with those third parties by some other app or in some other context.

For example P1 stated “I feel that the more new companies [there are] involved in collecting my data, the more I feel like I’m increasing my exposure to the world and to risks [...] increasing the danger of it being leaked”. Similarly P10 remarked: “it’s kind of in for a penny, in for a pound with Google”, implying that since Google already had his information, there would be no danger in sharing it with them again, while P27 argued “all else being equal I’d prefer to give my data to fewer parties”. This strategy was often borne out of protective instincts; P17 reported that “it instinctively makes me feel less threatened, knowing that my information will be shared with fewer people [companies]”.

However, some felt that thinking about marginal exposure in this way might encourage complacency in the face of over-zealous trackers: textbfP24 claimed that (“just because another app is doing it doesn’t make it ok. [that] is basically saying two wrongs make a right - that’s completely silly!”). P4 noted that there was a danger of “a ratcheting up effect, where you think well because my phone is already doing this, then I’m ok with it happening [...] so it is a double-edged sword”.

Participants talked at length about the love-hate relationships they felt they had with oligopolistic platforms. Remarks reflected a variety of internal conflicts from individual participants. For example, P19 mentioned:

I suppose I expected Google to have so much because I do use a lot of Google Apps, but
looking at it like this is simply alarming! The fact that Google can get so much more than the others alone... textbf{P19}

And yet, the deal that ubiquitous cloud services offered could also be characterised in much more positive terms.

'I'll just sit here and look after this information—your photos and stuff. I know that I need you sometimes, but you need me! And it all works out in the end.' textbf{P19}

The (geo)political economy of data

In some cases, participants reasoned on the basis of their perception of the (geo)political and economic factors affecting the possible use and abuse of their data. While such reasoning might reflect the demographic skew of our samples - most had university degrees - it nonetheless featured as an important and recurrent class of considerations influencing their preferences.

Perceptions of Regulatory Scrutiny

When faced with an unfamiliar entity, several participants described the importance of companies' size, jurisdiction and location, because such factors were important clues as to the level of regulatory scrutiny that would be applied to both private and government activities. Another participant believed that a bigger company would be 'more thoroughly monitored and restricted' by regulators than a small start-up (P22).

Specific geopolitical concerns

One participant discussed her preference for having her data held in jurisdictions with protective and/or well-known data protection laws. The same participant particularly wanted to ensure her data remained inaccessible to the Chinese government.

While such concerns are niche, they are clearly of enormous importance to certain users - and irrelevant to others. We posit that such an awareness of information geopolitics indicates the need for tools to support viewing and understanding the jurisdictional contexts where people's data are held.

Conclusion

In drawing out these themes, we aim to show that there are a wide variety of considerations and strategies that can come into play once end-users have time to get acquainted with the complex reality of today's smartphone apps and the ubiquitous data tracking networks associated with them. While there may be some role for automated enforcement of privacy preferences, there appears to be a significant class of subjective preferences which are unlikely to be predictable from the current range of behavioural data that would be available to an automated privacy assistant. Rather than focusing only on e.g. what permissions people grant to apps, it is therefore also important to develop privacy tools that provide the information and scaffolding people may need to develop and articulate a range of nuanced and idiosyncratic privacy attitudes in response to the complex social worlds in which they find themselves. To the extent that the differences lie in the details, we need elicitation methods which allow such detailed privacy preferences to emerge.

Acknowledgements

All authors were funded by the Engineering & Physical Sciences Research Council (EPSRC), under SOCIAM: The Theory and Practice of Social Machines EP/J017728/2, and under the PETRAS IoT Hub Strategic Fund,
References


Figure 1: X-Ray Refine. The exposure profile view consists of a stacked bar chart showing the impact contribution of each app, with destinations along the x-axis (sorted from most to least). Hovering over an app (in the legend) highlights the contribution of that app to the profile. Clicking on an app produces a focus display containing app details; clicking on a company name invokes a focus that displays destination details, if known. Hovering over app types highlight destinations of that type (e.g. advertisement, analytics, app functionality, etc). Supporting views show contributions by type, and location.