7-1-2013

Does This App Ask For Too Much Data? The Role Of Privacy Perceptions In User Behavior Towards Facebook Applications And Permission Dialogs

Hanna Krasnova
Humboldt-Universität zu Berlin, Berlin, Germany, krasnovh@wiwi.hu-berlin.de

Nicole Eling
Technische Universität Darmstadt, Darmstadt, Germany, eling@is.tu-darmstadt.de

Oleg Schneider
Humboldt-Universität zu Berlin, Berlin, Germany, schneideroleg@googlemail.com

Helena Wenninger
Technische Universität Darmstadt, Darmstadt, Germany, wenninger@is.tu-darmstadt.de

Thomas Widjaja
Technische Universität Darmstadt, Darmstadt, Germany, widjaja@is.tu-darmstadt.de

See next page for additional authors

Authors
Hanna Krasnova, Nicole Eling, Oleg Schneider, Helena Wenninger, Thomas Widjaja, and Peter Buxmann

DOES THIS APP ASK FOR TOO MUCH DATA? THE ROLE OF PRIVACY PERCEPTIONS IN USER BEHAVIOR TOWARDS FACEBOOK APPLICATIONS AND PERMISSION DIALOGS

Krasnova, Hanna, Humboldt-Universität zu Berlin, Spandauerstr. 1, 10178 Berlin, Germany, krasnovh@wiwi.hu-berlin.de

Eling, Nicole, Technische Universität Darmstadt, Hochschulstraße 1, 64289 Darmstadt, Germany, eling@is.tu-darmstadt.de

Schneider, Oleg, Humboldt-Universität zu Berlin, Spandauerstr. 1, 10178 Berlin, Germany, schneideroleg@googlemail.com

Wenninger, Helena, Technische Universität Darmstadt, Hochschulstraße 1, 64289 Darmstadt, Germany, wenninger@is.tu-darmstadt.de

Widjaja, Thomas, Technische Universität Darmstadt, Hochschulstraße 1, 64289 Darmstadt, Germany, widjaja@is.tu-darmstadt.de

Buxmann, Peter, Technische Universität Darmstadt, Hochschulstraße 1, 64289 Darmstadt, Germany, buxmann@is.tu-darmstadt.de

Abstract

Since Facebook has opened its platform to third-party developers, privacy concerns surrounding applications are mounting. By granting "permission" to an app request, users allow app providers to circumvent their privacy settings endangering their own privacy and that of their friends. Considering a rising use of Facebook apps and a paucity of research in this area, there is a pressing need to understand the rationale behind user acceptance of applications on Facebook and the role of different information items requested in this process. This study draws on experimental and survey findings gained on the basis of responses of 199 Facebook users. We find that users are affected by the number of information items included in a "permission" request, even though their concerns can be weakened by peer influence. Users tend to be particularly cautious when granting access to information of their friends, which rejects the assumption of "privacy egoism". Exploratory Factor Analysis reveals that in terms of privacy concerns users tend to categorize requested information items across five major clusters - friends', social, extended CV, basic CV and visual information. Our findings are relevant for Facebook page owners who employ applications to increase user engagement and learn more about their audience.

Keywords: Facebook, Privacy, Applications, Permissions.

1 Introduction

In 2007, Facebook (FB) undertook an unprecedented step by opening its platform to third-party application developers. The goal was to draw on their creativity to enrich FB with new applications and functionality. Social applications have mushroomed since then, providing users with significant functional and entertaining value (Russell-Bennett and Neale, 2009). Equally, businesses have grasped the unique value of apps viewing them as an opportunity to learn more about their customers, engage and reward them (Kalra and Shi, 2010). Despite this value potential, significant concerns are expressed regarding the widespread usage of FB apps (e.g. Perez, 2009; Steel and Fowler, 2010). In contrast to other third parties, FB applications are in a unique position to collect and process user data: While users may rely on privacy settings to limit accessibility of their information to undesirable others, applications can circumvent this barrier if a user *consents* to it. This consent is typically obtained in a *permission dialog* presented to a user, which she has to "*accept*" to be able to use the functionality of the application. This way, an application may "request" a user to grant it with access to his/her name, picture, birthday, gender, check-ins, relationship status, and even private messages. For example, analysis of permission dialogs of 9,411 FB applications conducted by Wang (2012) has revealed that 34.36% of applications request access to private email, 9.58% would like to access users' birthday, and more than 5% ask to access all the posts in the user's News Feed. In addition to user's personal details, applications can also ask to access information provided by user's *friends* as well as to *post in user's name* on the News Feed, thereby broadcasting user-related information to others. With FB users installing 20 million apps daily (Pring, 2012), privacy concerns surrounding FB apps are mounting. Already in 2008, the analysis of 150 most popular FB apps indicated that more than *90%* of apps have been granted more personal data than they really required, violating the principle of data minimization (Felt and Evans, 2008). Moreover, an investigation of the *Wall Street Journal* illustrated that top FB apps have been transferring personal data to more than 25 advertising companies, despite the fact that such practices are not allowed by FB (Steel and Fowler, 2010). Commenting on these practices, Soghoian (2008) argues that FB gives application developers too much access to users' data. While these developments are alarming, very little is known about users' decision-making process when accepting application "permissions". Indeed, do privacy considerations play a role in users' decisions to accept "permissions" of FB apps? Which information do users feel particularly sensitive about? Do users behave egoistically, when it comes to revealing information of their friends as opposed to their own information? To address these issues, we empirically investigate privacy dynamics of user interaction with applications' permissions on FB.

2 Research Background

FB differentiates between several categories of permission requests. *Basic Information* (*BI*) allows a default access to a limited set of user's data (e.g. Facebook ID, name, gender). To access more data, an app has to ask for additional permissions from *User and Friend (UF) permissions category* (*currently renamed into Extended Profile Properties (EPP)*). In total, over 40 information pieces can be accessed this way, ranging from user and friends' likes, interests, location, and photos (Facebook, 2012; 2013). Further, previously part of the *UF* permission category, *email* has recently been allocated into a separate "*Email Permissions*" (*EmP*) category (Facebook, 2013). Together, these categories represent commonly asked permissions on FB (Wang, 2012). Among the next category of *Extended Permissions* (*EP*) - "*publish_stream*" permission is often requested (recently renamed into *publish_actions*), which allows an app to post, comment, and like on a user's and friends' streams. Other categories of permissions are used relatively rarely (Wang, 2012). Overall, the scale of information FB apps can request calls for a better understanding of user behavior in this area.

An exhaustive literature review, resulting in 21 research papers, has revealed that existing research on FB apps so far has focused on three related yet distinct areas: (1) adoption and usage of FB apps on a

large scale (e.g. Nazir et al., 2008; Gjoka et al., 2008); (2) app usage on the individual level (e.g. King et al., 2011), and (3) frameworks for management of privacy configuration of FB apps (Wang et al., 2011). Particularly the first two lines of research are relevant for the purposes of our study. Exemplifying the *first* stream of research, Gjoka et al. (2008) show that 10% of apps account for 98% of total installations. Nazir et al. (2008) corroborate these results suggesting that the probability of a subscription of a new user to an app is proportional to the number of already existing users. Reasons include exposure to these apps on users' Wall and News Feed, which serve as primary advertising mechanisms for FB apps (Nazir et al., 2008). As a result, popular apps retain their strong positions, making it difficult for newcomers to become successful (Gjoka et al., 2008; Nazir et al., 2008). In addition to popularity, looking at individual motivations to find, add and remove apps (*second* stream of research), Besmer and Lipford (2010) identify *social interaction* as the major factor in app usage. Complementing this insight, Russell-Bennett and Neale (2009) find that an individual decision to recommend an app to one's friends is defined by either *social* and *functional* or *emotional* and *functiona*l value of an app. This is in line with the traditional studies on IT adoption, which suggest that both utilitarian and hedonic motives may play a role in user adoption decisions, especially when system use is voluntary (van der Heijden, 2004; Davis et al., 1992). In contrast, Besmer and Lipford (2010) show that *privacy concerns* regarding data (mis-)use hardly exist. In an attempt to explain this outcome, Kessler (2012) speculates that users may trade their personal data for the utility they receive from such an exchange. This argument corroborates with the 'Privacy Calculus' theory, according to which individual self-disclosure is a product of conflicting beliefs regarding expected benefits and privacy risks (Krasnova et al., 2010). Additionally, King et al. (2011) find that many users do not understand how apps work and which kind of information apps have access to. This is in line with Wang et al. (2011), who argue that it is hard for users to understand the permissions largely due to their chaotic display and the fact that permission dialogs fail to sufficiently inform users about the actual scope of these permissions. Taken together, even though traditional theories on IT adoption and "Privacy Calculus" as well as "app"-related research provide some initial insights, little is known about the cognitive dynamics behind user response to permission dialogs and their role in the application adoption process.

This scarcity, however, is partially compensated by research on *consent dialogs*. Studying a software-related context, Böhme and Köpsell (2010) confirm existence of habituation effects, allegedly as a result of the ubiquitous presence of End-User Licence Agreements (EULA). Investigating the effectiveness of phishing warnings, Egelman et al. (2008), however, show that user behavior can be dependent on the form of the warning, with less intrusive "pop-up" warnings being more likely to be ignored. The reasons for this may lie in users' willingness to trust the platform (Egelman et al., 2008), their reliance on "look and feel" of the websites (Fogg et al., 2001), limited knowledge (Böhme and Köpsell, 2010) and desire to complete the primary task (Egelman et al., 2008). While these studies offer additional insights into user information processing, FB permissions have unique particularities. Specifically, the exact listing of information to be accessed is unique for FB apps. Thus, which pieces of information "scare" users is unknown. In addition, the fact that some of the app permissions are revocable - a popular "*publish_stream*" permission allows users to control the audience to which an application can post - can give users a feeling of being in control, thereby stimulating them to accept an app (Xu et al., 2012). This and other fundamental differences call for a new in-depth study specifically concentrating on FB. To do so, in this study we adopt an empirical lens to explore the role and impact of FB privacy permissions in a decision to install an app.

3 Study Design and Analysis

3.1 Design and sampling

An exploratory scenario-based empirical study has been conducted, in which open and closed questions have been used. This approach is common in IS research to gain an in-depth understanding

of user perceptions and behavior (e.g. Xu et al., 2012). The survey consisted of two parts: *First*, respondents were randomly assigned to one of the two conditions. In each condition, respondents were presented with the FB "permission" for the same type of application - a sweepstake, which was presented to respondents as "*Giveaway App: Win One Of Several iPads"* (see Figure 1). While our choice of a sweepstake places limitations on the generalizability of our findings, it makes our study relevant for practice: Sweepstakes have been found to lead to higher user engagement and conversion rates (Delo, 2011). Since we planned to recruit respondents via the mailing list of our University, the *logo of our reputable institution* was integrated into both scenarios to represent the "provider" of the app in each case. This choice approximated respondents to a typical scenario of a sweepstake: Most users participate in sweepstakes of pages they are fans of, implying certain degree of trust in an app provider (Gupta, 2011). The only *difference* between the two scenarios was *the level of information requested by the app*. To choose information items for the *"low"* information scenario, we analysed the permission structure for a set of comparable application types - "*promotions*", "*giveaways*" and "*quizzes*" - using Socialbakers (2012). This helped us find the reference point regarding the most common permission requests for such apps. For the category *"quizzes"/* "*promotions* the 30 / 20 most popular apps (were selected. For "*giveaways"* 20 randomly found apps have been identified by searching for the apps with a keyword *"giveaway"* and choosing only those apps with monthly active users (MAU) of more than or equal to 1000 users. This approach has allowed us to derive the most popular permission requests (besides *basic information*) within these categories. Thus, *email* was requested in 55%/40%/40%, *user likes* in 45%/25%/n.a. and *birthday* in 15%/25%/13.3% of "*promotions*", "*giveaways*" and "*quizzes*" apps in our sample respectively. Hence, in addition to *basic information*, these items were used for the *"low"* information scenario (*see Figure 1*). For the *"high"* information scenario our choice was based on the information items requested by a popular Yahoo! app with MAU of 9.6 million (Yahoo!App, 2012). In addition, request for *user_photos* was also included, since it belongs to one of the most frequently requested permissions (Wang, 2012). Overall, besides *basic information*, 14 user- and 6 friend-related items have been requested. Further, two *EP* permissions (already present in Yahoo! App) were included: "*publish_stream*" and "*read_stream*" - enabling an app to post and access user's News Feed respectively. Note that the terminology for permission category names was used following Facebook (2012) (e.g. *publish_stream* and not *publish_actions*).

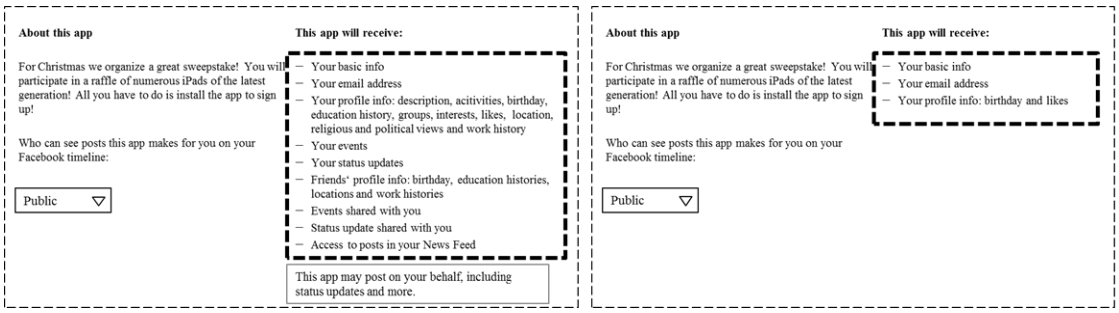

Figure 1. Schematic Representation of "high" (left) and "low" (right) information scenarios.[1]

Following the presentation of the permissions, respondents were asked, *among others*, to indicate their desire to install an app; and explain their rationale behind their decision (*see section 3.2*). In the *second part*, all respondents have answered an array of additional questions, which were the same for all participants (*see section 3.3*). The survey was advertised via the mailing list of our university, which mainly includes students and alumni. In total, 98 / 101 people responded to the *"low"/"high"* scenario respectively. Female FB users were slightly overrepresented in our sample - 57.7%. The

[1] Figure 1 reflects only schematic representations of the scenarios used in the study, shortened/adjusted for the publication in the proceedings.

mean age reached 24.7 (median=24.0; SD=5.1; min=18; max=60). Most respondents have spent most of their lifetime in Germany - 90.5% and were students - 87.6%. In terms of FB use, 11.4% of the sample used FB less than 5 minutes, 12.4% used it for 5-10 min., 15.4 - 11-20 min., 10.9 - 20-30 min., with the rest spending more than 30 min. *per day* on FB. As for FB apps, 62.2% stated to never have installed any apps, 26.4% installed between 1 and 5 apps and 4.5% installed between 5-10 apps.

3.2 Analysis: Intention to (not) install an app and rationale behind it

After permission dialogue was presented, respondents were asked to indicate how likely *they would install this app* as "*it is*" (Q1), and if it would be *recommended by a friend* (Q2). Respondents were generally turned off by a large number of information requests, which strengthened their desire to reject an app: In the "*high"* information scenario, only 4% of respondents considered installing the app (*Table 1*). Even when only the *target group* is considered - those who negated the statement "*In principle, I install no applications on Facebook*", respondents' interest remained low, even though visibly higher: 11.5%. Mann-Whitney U test has confirmed the presence of significant rank difference in the willingness to install an app for *"low"* and *"high"* information scenarios (Q1). Interestingly, a pair-wise Wilcoxon Signed Ranks Test has revealed that the likelihood of accepting an app was significantly higher when friend recommendation took place (p-value<0.01 for both scenarios for the overall sample and the target group), suggesting the effectiveness of viral distribution for apps.

Question	Scenario	Yes, absolutely	Rather, yes	Rather, no	Absolutely, not	Mean	Mann-Whitney U p-value
Q1. Would you install this app?	*low*	2.0%	12.0%	26.0%	60.0%	3.49	0.045
	high	1.0%	3.0%	24.8%	71.3%	3.66	
Q2. ..., if it is recommended by a friend?	*low*	2.1%	14.4%	37.1%	46.4%	3.28	0.086
	high	1.0%	5.0%	39.0%	55.0%	3.48	
Target group (n=71): (negative response to statement "*In principle, I install no applications on Facebook*")							
Q1. Would you install this app?	*low*	0.0%	25.0%	41.7%	33.3%	3.08	0.026
	high	2.9%	8.6%	28.6%	60.0%	3.46	
Q2. ..., if it is recommended by a friend?	*low*	5.7%	28.6%	40.0%	25.7%	2.86	0.047
	high	2.9%	11.4%	42.9%	42.9%	3.26	

Table 1. Willingness to install an app.

To gain an in-depth understanding of users' decisions about (not) installing the app, respondents had to justify their choice in an open *"Why?"* question. By not priming respondents about possible factors, we assured that really important decision categories were reported. Data corpus involved *2928 words* supplied by *191 respondents* and was analysed in two steps: First, relevant categories were identified by three authors. As a result, *five* main categories (with "*privacy concern*" category involving three subcategories) have been derived, as shown in Table 2. *In the second step*, two independent coders coded the data. With Cohen's Kappa – a measure of Inter-Coder Reliability – for all variables in our coding scheme being in the range of 0.702 - 0.941 (p=0.000), the quality of the coding procedure was assured. In case of disagreement, authors reached a decision via a compromise. 63.3% of the respondents mentioned only *one category*, 25.1% mentioned *two* and 7.5% mentioned *three or more categories (hence, percentages in columns of Table 2 do not add up to 100%)*. Importantly, responses reported in Table 2 describe the factors affecting the decision, but not the directionality of the effect.

According to our data, *privacy concern* emerged as the major factor affecting decision to (not) install an app for 64.3% of respondents in our sample. In this category, especially the *collection of data* by FB apps was of high relevance (55.8%). For instance, one of the respondents in the *"high"* information scenario complained: *"I would not allow the invasion of my privacy and in addition of my friends without limitations. This is completely unacceptable."* Potential *use of data* for spam,

advertising or commercial purposes as well as the *loss of control* were other privacy-related factors, which emerged from our analysis. With respect to the *loss of control*, respondents were particularly worried about the "*publish_stream*" permission, which allows the app to post in user's name. For example one respondent argued: *"And the worst: The app is allowed to post in MY name??!!! There is only one person who is allowed to do that and this is me."* Overall, *privacy concern* was the most mentioned factor in both scenarios. However, due to the fact that in the *"high"* information scenario the app asked for a lot more information items, there was a significantly higher share of respondents emphasizing their *privacy concern* in this group (p-value=0.008). By and large, such high emphasis on *privacy concerns* in both scenarios contradicts the findings of Besmer and Lipford (2010), who show that privacy concerns hardly play a role for app usage. The reason behind this discrepancy may be rooted in the fact that authors studied general app usage, so that respondents possibly had a low recall of their drivers and impediments when accepting an app. Indeed, considering that after installation, app permissions are not made transparent anymore, it is likely that users are little concerned about privacy at this stage. In contrast, in our study, respondents were put into a scenario where they actually had to decide on whether to install an app or not - a scenario highly approximated to reality.

Category	Share of respondents "low" Inf. Scenario	Share of respondents "high" Inf. Scenario	Mann-Whitney U test p-value	Share of resp. (pooled)
Privacy Concern	55.1%	73.3%	**0.008**	64.3%
• Collection of Data	43.9%	67.3%	**0.001**	55.8%
• Use of Data	10.2%	6.9%	0.410	8.5%
• Loss of Control	7.1%	20.8%	**0.006**	14.1%
General Attitude Towards FB Apps	32.7%	23.8%	0.164	28.1%
Value of the App	23.5%	12.9%	***0.053***	18.1%
Trust	16.3%	17.8%	0.780	17.1%
Others	8.2%	11.9%	0.384	10.1%

Table 2. Factors affecting the decision to (not) install an app.

In addition to *privacy concerns*, *general attitude towards FB apps* - the category reflecting users' positive or negative perceptions about FB apps in general (e.g. interesting/uninteresting, like/dislike, etc.) - was often mentioned (28.1% of respondents). For example: *"...Because I hardly install apps on Facebook"* or *"...Because I am critical of such Facebook apps"*. Further, v*alue of the app* category was mentioned to explain a decision to (not) accept an app, with most statements revolving around the incentive (iPads) or the probability of winning, which reflects the focus of our sweepstake scenario. Noteworthy is that this category was mentioned significantly more often in the *"low*" information scenario (p=0.053). This might be caused by respondents in the *"high"* information scenario focusing more on privacy and thus, ignoring other factors. Overall, the fact that, in total, *value of the app* category was mentioned by only *36* respondents can be caused by our choice of a sweepstake, since it is only incentivized by the prize and the probability of winning. In contrast, other apps (e.g. utility applications) may provide significantly more functionality and thus, result in greater value for users. Finally, *trust* towards either provider or the app emerged as another category of relevance: "*I would trust my university, given it is truly the university*". Interestingly, we did not find any significant difference in the number of mentions across scenarios (p=0.780). However, there were differences in the relative importance attached to *trust* when compared to other categories across scenarios: In the *"low"* information scenario respondents were more likely to stress the value of the apps (23.5% mentioned "*value*" vs. 16.3% mentioned "*trust*"). The opposite was true for the *"high"* information scenario: 12.9% vs. 17.8%. This suggests that when significant risks are involved, users rather rely on trust-mechanisms to rationalize their choices.

3.3 Exploring the magnitude of privacy concerns over information items

To deepen our understanding of the magnitude of *privacy concerns* different information items evoke, in the next step all respondents were presented with the same application scenario as before, but the area typically listing different information items was blanked. This way, *all respondents* were asked a question with regard to *the same scenario*: "*How concerned would you be if this app requested access to the following information items*?" Participants had to give their judgement with regard to 20 user- and 17 friend-related items from BI and UF/(EPP+EmP) permission categories (see Table 3, columns 1, 2, 3). Additionally, *publish_stream* and *read_stream* items were asked for. The answer options ranged: "*1=absolutely unconcerned; 2=slightly concerned; 3=concerned; 4=quite concerned; 5=very concerned*". The option "*cannot judge*" was also available, but was chosen very rarely. No significant differences have been registered between respondents who were previously assigned to *"high"* and *"low"* information conditions, which allowed us to pool the data. Overall, this part of the survey allowed us to gain a comprehensive view of user attitudes towards different information items typical for FB apps in particular and for Social Media (SM) platforms in general.

Table 3 lists the means of responses for all information items we asked for, sorted in the descending order of the magnitude of concern attached to *user* information (column 2 and 3). We notice that when it comes to their own information, users are mainly concerned about *losing control* over their identity: *publish_stream* item was attributed the highest mean level of concern, with the difference to the next most "disturbing" information item "*photos*" being statistically significant (p=0.000). In terms of their information, respondents were further particularly concerned about *photos*, ability of an app to access their News Feed (*read stream*), their *location, videos, former employer*, and *relationship details*. Of least importance were *home town, birthday*, *likes*, and *basic information*. Next, we examined whether respondents` concern regarding information items was a function of the *availability of this information (column 5 of Table 3)*. Indeed, it is plausible to assume that respondents who *did not* provide certain data would also be *less concerned* about an application requesting access to it. Interestingly, however, either we found *no difference* for specific information items, or respondents who have *actually provided* the data expressed *lower* concern regarding an app getting access to it *(column 6 of Table 3*). It is possible, that respondents who did not provide data on their *location*, *relationship details*, *about me*, *relationship status*, *education*, *interests*, *birthday* and *home town*, were guided by their privacy concerns to begin with. As a result, they took a more privacy-oriented stance with regard to information items - independent of their availability.

Interestingly, respondents were also highly concerned about an application accessing information of their *friend*s, with information regarding friends' *photos, location, videos, about me, religious and political views, status updates,* and *former employer* being most sensitive. In fact, users were more concerned, for example, about an app accessing the *location* data of *their friends* compared to their *own religious and political views* (p=0.000), *status updates* (p=0.000) and *events* (p=0.000). Overall, our data does *not* provide evidence for respondents being *"egoistic"* with respect to information of others. On the contrary, for *11 information items* (starting downwards from *religious and political views* in column 1 of Table 3), respondents were significantly more concerned about giving out data of their friends than their own (column 4 of Table 3). As concern regarding their own information grew, respondents were equally concerned about giving out their own and their friends' data regarding *location, videos, former employer*, and *relationship details*. The only exception were "*photos*", where respondents have shown a sign of "privacy egoism" - they were significantly more concerned about giving out their own vs. others' photos (column 4 of Table 3).

In our ad hoc analysis, we have conducted a Hierarchical Agglomerative Cluster Analysis with the Ward's linkage on respondents' *concerns regarding their own data*. By applying the elbow rule on the coefficients rendered by the agglomeration schedule, we derived that differentiating between 2 groups represents the optimal cluster number for our sample. Next, a non-hierarchical K-means clustering method has been applied to separate our dataset into 2 clusters. As a result, two groups of users emerged: 73% who were highly concerned about their data (mean of concern over all user information

items=4.39) and 27% who showed a moderate level of concern (mean=2.82). For all information items the difference between cluster 1 and 2 was statistically significant (p=0.000), but, notably, the ranking of privacy concerns for information items (most sensitive vs. least sensitive) remained largely unchanged for both clusters. Most remarkable was the fact that even for users in cluster 1 - those who were *particularly concerned about their own information* - we discovered the same pattern of attitudes towards friend`s information (vs. their own information) as registered for the whole sample (p-values similar to those reported in column 4 of Table 3). We conclude that, largely independent of concerns regarding their own information, users feel responsible regarding information of others.

Types of Information	Mean of concern		p-value (pair-wise)	Share of resp. provi-ding this data	p-value: (info provided/ not provided)	***Highest** Factor Loadings from EFA; Eigenvalues: F1=18.1; F2=3.9; F3=1.6; F4=1.4; F5=1.2; **F1= relate to friend information; F2-F5 = relate to user information***				
	user info	friends' info								
Column: 1	**2**	**3**	**4**	**5**	**6**	***F1***	**F2**	**F3**	**F4**	**F5**
publish_stream	4.77	**n/a**	**n/a**	**n/a**	**n/a**	***n/a***	**n/a**	**n/a**	**n/a**	**n/a**
photos	4.59	4.50	0.09	91%	0.801	*.688*				**.717**
read stream	4.46	**n/a**	n/a	n/a	n/a	*n/a*	**.472**			
location	4.33	4.34	0.95	41%	**0.016**	*.797*	**.331**			
videos	4.26	4.34	0.24	30%	0.287	*.651*				**.716**
former employer	4.23	4.33	0.15	30%	0.287	*.766*		**.569**		
relationship details	4.19	4.21	0.72	18%	**0.005**	*.814*		**.744**		
about me	4.03	4.34	0.00	46%	**0.094**	*.770*	**.477**			
email	4.03	**n/a**	**n/a**	**n/a**	**n/a**	***n/a***			**.730**	
rel. and pol. views	4.02	4.34	0.00	15%	0.287	*.729*		**.622**		
status updates	4.00	4.34	0.00	81%	0.525	*.783*	**.566**			
activities	3.92	4.23	0.00	52%	0.268	*.820*	**.713**			
events	3.78	4.13	0.00	64%	0.567	*.737*	**.728**			
relationship status	3.76	4.11	0.00	34%	**0.016**	*.829*		**.669**		
education	3.70	4.13	0.00	63%	**0.005**	*.819*		**.556**		
groups	3.66	3.99	0.00	73%	0.797	*.863*	**.744**			
interests	3.60	4.10	0.00	53%	**0.013**	*.888*	**.658**			
basic information	3.55	**n/a**	**n/a**	**n/a**	**n/a**	***n/a***			**.654**	
likes	3.52	3.97	0.00	90%	0.549	*.862*	**.765**			
birthday	3.32	3.91	0.00	77%	**0.000**	*.778*			**.589**	
home town	3.15	3.91	0.00	64%	**0.000**	*.825*			**.610**	

Table 3. User privacy concern regarding requested information items.

3.4 Which information items influence users' decisions to accept an app?

Next, we explored the role of privacy concerns over specific information items in user decisions to install an app. Indeed, concerns over information items included in our *"low"* and *"high"* information scenarios were also measured (highlighted in grey in columns 2 and 3 of Table 3). For this purpose, correlation analysis has been conducted considering the exploratory nature of our study. Specifically, users' willingness to install an app "as is" and under "friend recommendation" (Q1 and Q2 in Table 1)

were correlated with user privacy concerns for the information items included in respective scenarios, as summarized in Table 4 (columns: Q1 and Q2). Spearman's Rho method was used since our data was not normally distributed (p-value of the Kolmogorov-Smirnov test for all "concern" items equal to 0.000). Moreover, our approach to measurement rather suggests ordinal data (Lehman et al., 2005). We find that for the *"low*" information scenario, especially concerns about such personal items as *basic information, email and birthday* had a strong and significant correlation with user decision for both Q1 and Q2. Correlation with the concerns about the app accessing user "*likes*" was also significant, but of smaller magnitude. Overall, the fact that concerns over all information items were significantly positively linked with user decision to reject an app, shows that users aligned their behavior to their privacy attitudes. One of the reasons could be the small amount of time respondents needed to actually study the information items requested in the *"low"* information scenario. While a whopping 42% in the *"high"* information scenario stated to *not read* or only *briefly read* the information in the permission request, this share was significantly lower for the *"low"* information scenario - 34.7% (p-value=0.074). We find slightly different dynamics with regard to "*high*" information scenario: Here, concerns about user information items such as *basic information, likes, "about me", groups, interests, religious and political views, status updates (Q1), events (Q2)* and, interestingly, *photos*, were not linked to user decision. Two reasons can explain these inconsistencies. On the one hand, users might be overwhelmed by the length of the privacy permission: Only 20% of respondents in the *"high"* information scenario stated to have studied the permissions "*in detail*" before making a decision. This effect has been well-described in the EULA-related research (e.g. Good et al., 2007). Second, their attention may have gotten fixed on other information items: Specifically, concerns about user email, birthday, activities, education, location, former employer, events (Q1), status updates (Q2) and all *friends-related* items exhibit a strong positive correlation with willingness to reject an app. Overall, the relevance of *all* friend-related information items in user decision is astounding.

***"Low"* Information Scenario**	**Q1**	**Q2**	***"High"* Information Scenario (continued)**	**Q1**	**Q2**
u_basic information	**.504***	**.518***	u_location	**.330***	**.234***
u_email	**.443***	**.409***	u_ rel. and pol. views	.200	.216
u_birthday	**.443***	**.381***	u_ former employer	**.311***	**.343***
u_likes	**.239***	**.225***	u_events	**.237***	.169
***"High"* Information Scenario**	**Q1**	**Q2**	u_status updates	.213	**.240***
basic information	.144	.208	u_photos	.101	.043
email	**.193***	**.177***	f_birthday	**.405***	**.254***
u_birthday	**.278***	**.332***	f_education	**.385***	**.336***
u_likes	.083	.167	f_location	**.336***	**.238***
u_about_me	.138	.156	f_ former employer	**.403***	**.340***
u_activities	**.252***	**.305***	f_events	**.296***	**.224***
u_education	**.365***	**.288***	f_status updates	**.283***	.175
u_groups	.104	.098	u_read_stream	**.331***	**.257***
u_interests	.129	.139	u_publish_stream	**.385***	**.346***

*Table 4. Correlations between Q1/Q2 (1=absolutely "accept"; 4= absolutely "reject") and Privacy Concerns regarding a specific item (1=not concerned at all; 5=very concerned). * p < .05 ("u_" refers to user's and "f_" to friends' information.).*

Further, using correlation analysis we found no link between availability of any information item requested by a permission and users' willingness to install an app "as is" (Q1) and via "friend recommendation" (Q2) in both scenarios we tested. This suggests that user decisions are independent of the actual information availability, but are rather a function of other factors (see Table 2).

3.5 Towards a new topology of user privacy perceptions of information

Responses across 37 information items summarized in Table 3 allowed us to derive a topology of user privacy concerns with respect to specific information items. To do so, Exploratory Factor Analysis (EFA) was employed: EFA is a widespread technique used to study and cluster variables in a sample with high complexity (Hair et al., 1998). We used a Principal Components Analysis with Varimax rotation to examine the emerging category structure since one of the goals was complexity reduction. A cross-check using Principal Axis Factoring with Direct Oblimin rotation has rendered equivalent results. As a result, 5 factors with Eigenvalues higher than 1 were extracted (see Table 3). Together these factors explained 71% of the variance in the data structure. Factor loadings of all, but one item, exceeded the threshold of 0.4 (Hair et al., 1998). Additionally, 65% of items fulfilled the narrow definition of "factor purity" Saucier (1994). The internal consistency of factor components measured by Cronbach's Alpha was high, reaching 0.77 – 0.98. *Factor 1* "*concern over Friends' Information*", encompassed *all* items supplied by friends. *Factor 2* - referred to as "*concern over Social Information*" subsumed such items as *user likes, groups, events, activities, status updates, information on users' News Feed*, information *"about a user"* and *location*. *Factor 3* - referred to as "*concern over Extended Curriculum Vitae (CV) Information*" included such information pieces as *religious and political views, former employer, education, relationship status* and *details*. *Factor 4* - referred to as "*concern over Basic CV Information*" encompassed user *email, basic info, hometown* and *birthday*. Finally, in *Factor 5* - called "*concern over Visual Information*" - information about *photos* and *videos* was included. In line with our previous findings (see Table 3), we find that users are most concerned about their *Visual Information* (mean across all items=4.42), followed by their *Friend's Information* (mean=4.18). *Social and Extended CV Information* is a little less concerning, with a mean score of 3.91 and 3.97 respectively. Overall, participants appear least threatened about their *Basic CV Information* (mean=3.52). Using a T-test for a dependent sample, we found that the means differ significantly for all pairs of factors, with only one exception, i.e. means of *Social Information* and *Extended CV* were not significantly different from each other. Overall, our topology is useful in interpreting significance of some and not-significance of other information items reported for the *"high*" information scenario in section 3.4. It appears that when faced with lengthy information requests, users may focus on specific information "chunks" and disregard others. Thus, concerns over user *likes, groups, interests, location, "about me"* were found to be not important in user decision regarding the app. Since all of these items belong to the "social information" category identified in our topology, it can be concluded that users tend to omit it when processing large permission requests. In contrast, "friends" data is highly influential (see Table 4). While these conclusions are tentative, our topology represents the first step towards the new taxonomy of information sensitivity in the context of social media. Indeed, while in the past financial, health and sexual life data have typically been mentioned as major examples of sensitive information, social media users attach a different value to their private data.

4 Discussion, Implications and Concluding Remarks

The *theoretical* contribution of this study is five-fold. *First*, we show that privacy concerns are a major factor affecting users' decision-making process: With increasing amount of personal data requested by an app, the willingness to install an app decreases. Noteworthy, however, is the fact that even in *"low"* information scenarios a whopping 55.1% claimed to have privacy concerns, suggesting a pivotal role of these attitudes in the decision-making process. This effect is, however, weakened when the app is recommended by a friend. *Second*, in terms of specific permission items, we find that users are mainly concerned about *losing control* over their identity - a revocable "*publish_stream"* item often requested by apps on FB (Wang, 2012). Other most "disturbing" user-related information items are represented by "*photos*", ability of an app to access users' *News Feed* (*read_stream*), *user location*, *videos*, *former employer,* and *relationship details*. To our surprise, respondents expressed high concern about applications accessing information of their *friend*s, with friends' *photos, location, videos, about me, religious and political views,* and *status updates* being most sensitive. For a large share of information

items respondents either attached a higher or equal value to *friends'* information as to their own. Hence, "privacy egoism" assumption was rejected. *Third*, we find that users' attention to information permissions is contingent on its length. While in a *"low"* information scenario respondents acted on their concerns when deciding on the app, in a *"high"* information scenario answers were less consistent. This suggests that when facing a lengthy request, users may focus on specific information "chunks" or adopt other heuristics. *Fourth*, we did not find any evidence confirming a significant link between the presence of a particular information item on a user profile and user decision to (not) install an app. Apparently users rather act on their (privacy) attitudes, rather than evaluate the actual risk contained in the request. *Fifth*, this study represents the first step towards a new topology of information sensitivity in the context of social media. Indeed, while in the past financial, health and sexual life data have typically been mentioned as major examples of sensitive information, social media use implies disclosure of more and essentially different information pieces with long longevity. However, the value users attach to these information items in terms of privacy has so far remained unknown. Our findings contribute to this research gap, suggesting that in terms of their privacy concerns users cognitively differentiate between *such categories* of personal data items such as *friends', social information, extended CV, basic CV* and *visual information*.

Our results provide an array of *practical* implications: Since *peer* recommendations have a strong influence on users' willingness to accept an app, viral distribution of apps emerges as a promising approach. Further, our results call for a greater caution on the part of app developers when deciding on which information items to request. Specifically, items associated with the loss of *control* - like "*publish_stream"* - emerge as an impediment on the road to widespread adoption. Further, *visual* information items are treated with suspicion. Particularly in *"low"* information requests, users appear to consistently integrate concerns about individual information items into their decision-making process. Beyond the context of FB, our findings offer rich insights for other emerging areas, such as Android-Apps or location-based services, which equally depend on users' willingness to disclose specific information pieces in the installation process. For the purposes of our study, our sole reliance on FB has, however, allowed us for a close-to-reality experimental design since prominent and detailed display of the permission request is available on FB. Taken together, this study represents an initial step in exploring the effects of privacy permissions on FB application adoption and the weight attributed to different information items in this process. Insights regarding the cognitive calculus of users and the developed topology of information sensitivity in the context of social media can serve as a springboard for further investigations.

References

Besmer, A. and Lipford, R.H. (2010). Users' (Mis)Conceptions of Social Applications. Graphics Interface Conference 2010. Ottawa, Ontario, Canada.

Böhme, R. and Köpsell, S. (2010). Trained to Accept? A Field Experiment on Consent Dialogs. SIGCHI Conference on Human Factors in Computing Systems. Atlanta, GA, USA: ACM.

Davis, F.D., Bagozzi, R.P. and Warshaw, P.R. (1992). Extrinsic and Intrinsic Motivation to Use Computers in the Workplace. Journal of Applied Social Psychology, 22 (14), 1111-1132.

Delo, C. (2011). Why Brands Still Need Facebook 'Fans' [Online]. AdAge digital. Available: http://adage.com/article/digital/study-facebook-fan-worth-10-average-brands/231128/ (Accessed 4.12.2012).

Egelman, S., Cranor, L.F. and Hong, J.I. (2008). You've Been Warned: An Empirical Study of the Effectiveness of Web Browser Phishing Warnings. Proceedings of the SIGCHI Conference on Human Factors in Computing Systems. Florence, Italy.

Facebook. (2012). User and Friend Permissions [Online]. Facebook developers. Available: http://developers.facebook.com/docs/reference/login/user-friend-permissions (Accessed 04.11.2012).

Facebook. (2013). Permissions [Online]. Facebook developers. Available: http://developers.facebook.com/docs/reference/login/ (Accessed 28.03.2013).

Felt, A. and Evans, D. (2008). Privacy Protection for Social Networking Platforms. Workshop on Web 2.0 Security and Privacy. Oakland, CA, USA.

Gjoka, M., Sirivianos, M., Markopoulou, A. and Yang, X. (2008). Poking Facebook: Characterization of OSN Applications. First Workshop on Online Social Networks. NY, USA

Good, N.S., Grossklags, J., Mulligan, D.K. and Konstan, J.A. (2007). Noticing Notice: A Large Scale Experiment on the Timing of Software License Agreements. Proceedings of the SIGCHI conference on human factors in computing systems. San Jose, CA, USA.

Gupta, A. (2011). How Brands Are Using Facebook Apps for Contests and Campaigns [Online]. Available: http://socialmediatoday.com/achintya-gupta/304412/how-brands-are-using-facebook-apps-contests-and-campaigns (Accessed 07.12.2012).

Hair, J.F., Anderson, R.E., Tatham, R.L. and Black, W.C. (1998). Multivariate Data Analysis with Readings, Prentice-Hall.

Kalra, A. and Shi, M. (2010). Consumer Value-Maximizing Sweepstakes and Contests. Journal of Marketing Research, 47 (2)**,** 287-300.

Kessler, S. (2012). How Much Do Your Friends' Facebook Apps Know About You? [Online]. Available: http://mashable.com/2012/03/30/facebook-friends-apps-privacy/ (Accessed 07.12.2012).

King, J., Lampinen, A. and Smolen, A. (2011). Privacy: Is There An App for That? Symposium On Usable Privacy and Security (SOUPS) 2011. Pittsburgh, PA, USA.

Lehman, A., O'Rourke, N., Hatcher, L. and Stepanski, E. (2005). JMP for Basic Univariate and Multivariate Statistics: A Step-by-step Guide, Cary, NC, SAS Institute.

Nazir, A., Raza, S. and Chuah, C.-N. (2008). Unveiling Facebook: a Measurement Study of Social Network Based Applications. IMC'08. Vouliagmeni, Greece.

Perez, S. (2009). Does that Facebook App Have a Privacy Policy? Probably Not. [Online]. Available: http://www.readwriteweb.com/archives/does_that_facebook_app_have_a_privacy_policy_probably_not.php, (Accessed 08.10.2012).

Pring, C. (2012). 100 social media statistics for 2012 [Online]. Available: http://thesocialskinny.com/100-social-media-statistics-for-2012/ (Accessed 07.12.2012).

Russell-Bennett, R. and Neale, L. (2009). Social Networking: Investigating the Features of Facebook Application. Academy of Marketing Annual Conference 2009. Leeds: Leeds Metropolitan University.

Saucier, G. (1994). A Brief Version of Goldberg's Unipolar Big-Five Markers. Journal of Personality Assessment, 63 (3)**,** 506-516.

Socialbakers. (2012). Applications on Facebook [Online]. Available: http://www.socialbakers.com/facebook-applications (Accessed 04.12.2012).

Soghoian, C. (2008). Exclusive: The Next Facebook Privacy Scandal [Online]. CNET. Available: http://news.cnet.com/8301-13739_3-9854409-46.html (Accessed 07.12.2012).

Steel, E. and Fowler, G.A. (2010). Facebook in Privacy Breach [Online]. The Wall Street Journal. Available: http://online.wsj.com/article/SB10001424052702304772804575558484075236968.html# (Accessed 07.12.2012).

van der Heijden, H. (2004). User Acceptance of Hedonic Information Systems. MIS Quarterly, 28 (4)**,** 695-704.

Wang, N. (2012). 'Third-Party Applications' Data Practices on Facebook. 2012 ACM Annual Conference. New York, NY, USA.

Wang, N., Xu, H. and Grossklags, J. (2011). Third-Party Apps on Facebook: Privacy and the Illusion of Control. ACM Symposium on Computer Human Interaction for Management of Information Technology (CHIMIT). Boston, MA, USA.

Xu, H., Teo, H.-H., Tan, B.C. and Agarwal, R. (2012). Effects of Individual Self-Protection, Industry Self-Regulation, and Government Regulation on Privacy Concerns: A Study of Location-Based Services. Information Systems Research**,** pp. 1–22.

Yahoo!App. (2012). Available: http://www.facebook.com/apps/application.php?id=90376669494 (Accessed 04.12.2012).