



# Trusting the Big Brother Inside My Pocket: User-Oriented Requirements for Contact Tracing

Claudia Negri-Ribalta, Nicolas Herbaut, Camille Salinesi

## ► To cite this version:

Claudia Negri-Ribalta, Nicolas Herbaut, Camille Salinesi. Trusting the Big Brother Inside My Pocket: User-Oriented Requirements for Contact Tracing. Jochen De Weerd; Artem Polyvyanyy. Intelligent Information Systems. CAiSE Forum 2022, Leuven, Belgium, June 6–10, 2022, Proceedings, 452, Springer International Publishing, pp.82-91, 2022, Lecture Notes in Business Information Processing, 978-3-031-07480-6. 10.1007/978-3-031-07481-3\_10 . hal-03789989

**HAL Id: hal-03789989**

**<https://hal.science/hal-03789989>**

Submitted on 27 Sep 2022

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Trusting the Big Brother Inside my Pocket: User-Oriented Requirements for Contact Tracing

Claudia Negri-Ribalta<sup>\*1</sup>, Nicolas Herbaut<sup><sup>1</sup></sup>, and Camille Salinesi<sup><sup>1</sup></sup>

Centre de Recherche en Informatique, Paris I, France

`claudia-sofia.negri-ribalta@univ-paris1.fr`, `{first}.{last}@univ-paris1.fr`

**Abstract.** Contact tracing (CT) apps have been rolled out as part of combined efforts to control the COVID-19 pandemic. However, these apps haven't been download by the totality of the population, with users raising concerns over the usage of data and data protection. This article analyzes how different app providers and data protection levels can affect the willingness to download a COVID-19 contact tracing app between university students, and if there is a difference between students who downloaded and didn't download the app. Through the usage of a factorial survey experiment (FSE) the paper statistically analyses the data protection and trust (DPT) requirements, using the French COVID-19 CT as a use case. The results show that universities have a positive impact on the trust levels of app providers, while private companies tend to have a negative impact on trust. Subjects also have high expectations on data protection. Our results highlight the importance of data protection and app providers for building trust with users of CT COVID-19 apps.

**Keywords:** contact tracing · trust · data protection · requirements · HCI

## 1 Introduction

COVID-19 related apps were rolled out as part of national plans to control the pandemic by contact tracing. If governments plan on using these apps as part of their efforts for future pandemics, it's pivotal to comprehend data protection and trust (DPT) requirements.

Existing research [15,10,14,2] indicates that users are concerned that these apps might be used for surveillance, hence there are trust conflicts. Yet, some of these studies haven't necessarily worked with statistical analysis or used hypothetical scenarios.

This is an exploratory research, part of a series of DPT articles in the context of information system engineering (ISE). We study the willingness to download COVID-19 contact tracing (CT) apps by university students, and compared between users who did and didn't download the French CT COVID-19 app. The main research question is : How do app providers and data protection affect the willingness to download a COVID-19 contact tracing app?

---

<sup>\*</sup> Corresponding author

We expect both app providers and data protection factors to be significant on the willingness to download. App providers which include universities will have higher willingness to download, while higher data protection levels will also have higher willingness to download. We also hypothesize that there is a difference in the coefficient between subjects who downloaded the apps and those who didn't. To gather the data, we used a factorial survey experiment (FSE) which was analyzed with a random intercept mixed multilevel model. FSE is a research tool used in other disciplines, such as sociology [4] and has been occasionally used in RE like in [6,13]. FSE - if correctly designed - can have high internal and external validity. Its high internal validity is a consequence of the experimental variation of the data gathering, which resembles an experiment. Consequently, it can draw conclusions from non-random samples, as different stimuli were presented to the same subject [4]. Its external validity is a consequence of its survey-like characteristics.

This article is a validation and implementation of previous research presented by [13], whose purpose is to help the RE process when dealing with requirements that might have a social bias or ethical requirements. This specific framework proposes the usage of FSE as a tool for the RE activity and highlights that quantitative data can be gathered from different stakeholders - such as users - without necessarily having a random sampling. There has been some history of using FSE in RE, as seen in [6,8]. However, [13] proposed the usage of FSE for RE as a framework, for specific types of requirements. The framework allows the RE practitioner to compare results between different stakeholders. In this case study of CT COVID-19 apps, we compared subjects who download and didn't download the french CT COVID-19 app. Our results suggest these groups have different coefficients in the subjects of interest.

By using the framework proposed by [13], this article not only provides validation and implementation for the specifically cited framework but also provides statistical evidence and analysis on how app providers and data protection can affect the willingness of users to download an app. It also provides evidence that open-source seems to be not that significant in the willingness to download an app.

The paper is divided as follows: Section "Preliminaries and related work" familiarizes the reader with CT apps and DPT elements. Section "Research method" discusses the research methods used. In "Results and analysis", the data is presented, analyzed and discussed. Section "Threats to validity and future work" analyzes the limitations of this research. Finally, "Conclusion" summarizes the research.

## 2 Preliminaries and related work

**Trust.** Trust is a concept that spans across multiple areas of our society, and is complex and multidimensional [12]. Due to page limitations, we simply observe that in the literature [1,7,11,12], three elements always appear as decisive for trust: risk, expected behavior and a relationship between two agents. In this paper, we use the definition of trust given by [1]: "trust is the willingness of the trustor to rely on a trustee to do what's promised in a given context, irrespective of the ability to monitor or control the trustee, and even though negative consequences may occur".

**Data protection.** For this research, data protection is understood as compliance with the GDPR (CIA principles, data protection by design and default, among others articles).

**Privacy and trust issues with contact tracing apps.** France second CT app, TousAntiCovid, was launched in October 2020 after (among others) criticisms on the privacy of the first version<sup>1</sup>. By June 2021 (we carried our survey between May and June 2021), the app had been downloaded 19 million times<sup>2</sup>.

Previous research established the existence of a link between trust in the app provider and willingness to download a Covid-19 CT app [2,5]. [5] suggested that licensing the app with open-source<sup>3</sup> could help with the trust issues.

From a user perspective, data protection, privacy and surveillance concerns have been highlighted across the literature for [14,2,15]. Indeed surveillance and data leaks are specifically mentioned in different studies as a reason for not downloading this app [2,15]. [2] concluded that subjects who have less trust in the government associate the CT apps with surveillance. Furthermore, [2] concludes that even if these apps have perfect privacy and data practices, some subjects might never download them.

### 3 Research method

For the study, we followed factorial survey experiment (FSE) research method. FSE has been regarded as a suitable tool to deal with topics that might have social bias [3,13]. FSE has been used to investigate beliefs, attitudes, judgments, or requirements, in a variety of disciplines, ranging from social science to requirement engineering [4,3,6].

#### 3.1 Introduction to Factorial Survey Experiment

In brief, this method consists of presenting the respondents with several scenarios, called vignettes, to be rated them with a defined scale. Each vignette consists on a combination of different factors [4] (elements of interest), each declined into several levels. For example, for 5 factors with 3 levels each, the universe of vignettes is  $3 \times 3 \times 3 \times 3 \times 3 = 3^5$ .

Respondents must rate all or a fraction of vignettes (to avoid fatigue) depending on the design: randomize, decking or design algorithms such as D-efficiency [3]. Since no question is directly prompted, social bias is limited. Because the same respondent rated several related scenarios, causal relationships between variables can be investigated [3].

<sup>1</sup> <https://reuters.com/article/us-health-coronavirus-france-apps-idUSKBN27A0AZ>

<sup>2</sup> <https://data.gouv.fr/fr/datasets/metriques-dutilisation-de-lapplication-tousanticovid/>

<sup>3</sup> As defined in <https://opensource.org/docs/osd>

L1 analysis investigates relations between vignettes, while L2 analysis focuses on how different groups rate similar vignettes. Different statistical model can be used: OLS, mixed multilevel with random intercept (R.I.) or random slope (R.S.), etc.

Due to the intentional variability of the variables inside the survey, FSE resemble an experiment, having high internal validity if properly designed [3]. Although random sampling is desirable due to the external validity, FSE allows generalizing its conclusions as they are a reaction to the experimental stimuli rather than the sampling [4]. In other words, it acts as an experiment too, and conclusion can be made even without random sampling. “The experimental logic of an FS has the advantage of permitting general conclusion about causal mechanism using non-random” [4, pg.11].

Factor	Levels	Detail of factor levels
App provider (\$AP)	7	Any combination between government, university, private company
Data Protection (\$DP)	3	High, basic, little
Open source (\$OS)	2	Open source code, proprietary code

**Table 1.** Vignette factors and their levels in the survey

### 3.2 Design of the study

**Factors definition and their levels** The research focuses on DPT and open-source. Table 1 shows the factors and the levels. The factors were carefully worded to avoid including societal values. For example, privacy is usually perceived as an important value, therefore it was defined using the GDPR, rather than privacy. A vignette was of the form ‘How comfortable would you feel while downloading a COVID-19 contact tracing application that is \$OS, with \$DP data protection and was created by \$AP?’

**Survey presentation and testing.** The vignette universe size of this study is 42, a number too high for respondents to evaluate. The vignette universe was divided in two decks, based on open-source and each subject was presented 21 vignettes, that could be rated in an 11-Likert scale.

The survey was pre-tested with 80 participants. Feedback made us adapt the presentation of the scenarios into a “table like” form. Our universe of interest is university students in France (2.7 million students). Using Cochran’s formula, with a 95% confidence and a 5% margin of error, the desired sample size was of 385 answers.

Given the pandemic, it was difficult to gather a random sample from around France, as mobility was limited. Furthermore, clustering different cities would be difficult due to movement restrictions and universities shutdown. Thus it was decided to grab the sample from students from Paris, understanding the impacts on the external validity. As explained in Section 3, due to the internal validity of FSE, results of this study are still significant and interesting, even without a random sampling and even possible biases.

## 4 Results and analysis

434 persons answered the survey between May and June 2021, with 414 valid answers (12 rushers, 10 failed the attention test, 7 not studying in Paris were discarded). Only 3 students in Agricultural science answered, and they were deleted as it was not statistically significant. The final dataset has answers from 210 females and 204 males. 231 (56%) respondees answered they had downloaded TousAnti-Covid, while 183 (44%) indicated they hadn't. 163 students were from engineering or technology, 89 from social science, 75 from natural science, 56 from humanities and 32 from medical science.

After cleaning the data, we ran a  $\chi^2$  test of independence with an alpha of 0.05 and a Bonferroni correction, between the following variables: download, area of study, knowledge of open-source, gender, and last usage of the application. We fail to reject all the  $H_0$  of association (independence) between variables.

### 4.1 Multilevel Model

Following the FSE literature, the data is analyzed using a random intercept multilevel mixed model, based on the reference book [4]. The model is presented in Equation (1), which is a formalization of the model used to analyze our data.

In Equation (1), each vignette  $i$  is rated by subject  $t$ . The factors \$DP and \$AP are evaluated through the beta-coefficient, as they are within-subject variables. \$OS and \$dwn are evaluated through the gamma coefficient, as they are between-subject variables. Citing [4], "X represents the p vignette dimension, Z represents the q response variables".  $\mu$  is the error added by the fact that subjects rate several vignettes, and  $\varepsilon$  is the random component. The random intercept is "the sum of the intercept and this error component [i.e.,  $\beta_0 + \mu_t$ ]" [4, pg.89].

$$Y_{it} = \beta_0 + \beta_1 X_{it1} + \beta_2 X_{it2} + \dots + \beta_p X_{itp} + \gamma_1 Z_{t1} + \gamma_2 Z_{t2} + \dots + \gamma_q Z_{tq} + \mu_t + \varepsilon_{it} \quad (1)$$

We fitted the equation to our model in different specialized software. The results presented here are obtained from `clmm` function from `ordinal` package from *R*. The `clmm` runs a cumulative linked mixed model fitted with the Laplace approximation. Different models were tested and compared based on their AIC, log-likelihood and others parameter.<sup>4</sup>

### 4.2 Analysis

The results show that the different levels of \$DP and \$AP are statistically significant in the willingness of downloading, but not \$OS. The analysis also shows that there are some differences between subjects who did and didn't download the app.

<sup>4</sup> Codes and database are available at [https://github.com/csnegri/CaiseForum\\_22\\_Negri](https://github.com/csnegri/CaiseForum_22_Negri).

Term	RI with L2		RI for those who	RI for those who
	\$Dwn Coef. (s.e.)		downloaded app	didn't download
	Coef. (s.e.)		Coef. (s.e.)	app Coef. (s.e.)
	[IC 95%]		[IC 95%]	[IC 95%]
Data protection				
High (ref.)	-		-	-
Medium	-1.273 (0.049)*** [-1.369 ; -1.178]		-1.287 (0.064)*** [-1.416 ; -1.161]	-1.282 (0.002)*** [-1.286 ; -1.278])
Low	-2.906 (0.057)*** [-3.018 ; -2.793]		-3.060 (0.076)*** [-3.209 ; -2.910]	-2.714 (0.003)*** [-2.721 ; -2.708])
App provider				
Gov.	Uni.	Priv.		
✓	(ref.)	-		-
	✓		0.662 (0.075)*** [0.516 ; 0.808]	0.336 (0.099)*** [0.143 ; 0.529]
		✓	-1.947 (0.079)*** [-2.101 ; -1.794]	-2.328 (0.103)*** [-2.530 ; -2.126]
	✓	✓	-1.483 (0.075)*** [-1.631 ; -1.338]	-1.681 (0.099)*** [-1.877 ; -1.487]
✓	✓		0.390 (0.073)*** [0.246 ; 0.534]	0.210 (0.098)* [-0.019 ; 0.402]
	✓	✓	-1.055 (0.074)*** [-1.200 ; -0.910]	0.610 (0.003)*** [0.605 ; 0.615]
✓	✓	✓	-0.877 (0.074)*** [-1.022 ; -0.723]	-0.647 (0.002)*** [-0.651 ; -0.642])
			-1.061 (0.098)*** [-0.704 ; -0.328]	-0.658 (0.002)*** [-0.663 ; -0.653])
Type of code				
Open source (ref.)	-		-	-
Proprietary code	-0.028 (0.156) [-0.333 ; 0.277]		-0.018 (0.217) [-0.442 ; 0.408]	-0.139 (0.191) [-0.515 ; 0.237]
Download				
No (ref.)	-		-	-
Yes	1.673 (0.231)*** [1.220 ; 2.125]		- -	- -
Intercept variance, ID (st.dev)	5.235 (2.288)		4.89 (2.211)	5.711 (2.39)
Respondents	414		231	183
Log likelihood	-15902		-9265	-6594

Significant p-values : ° p <0.1 \* p <0.05, \*\* p <0.01, \*\*\* p <0.001

**Table 2.** Multilevel models, compared by L2 groups (\$dwn). Bold elements have different p-values for each L2 group

*Impact of different app provider in willingness to download.* The analysis from the model supports the hypothesis that \$AP is significant in the willingness to download. Table 2 shows that app providers that include private companies have consistently negative beta coefficients on average, whereas app providers that include universities - without private companies - are consistently rated with a positive beta coefficient. *University* and *university and government* levels were evaluated at 0.662 and 0.390

points higher on average. In comparison, the levels that included private companies were all rated negatively on average. All these results are statistically significant.

The \$OS factor has no statistical significance at the different level (*proprietary code*). The subjects rated the level *proprietary code* on average -0.028 points. It may suggest that for the subjects the difference between open source and proprietary is not be critical.

*The impact of different data protection levels in the willingness to download* The analysis reveals that \$DP is an important factor for the subjects. For instance, the beta coefficient of medium and low \$DP are rated at -1.273 points and -2.906 points lower accordingly. The level *low* of \$DP is the highest beta coefficient of all the levels in the study. Previous literature indicates that data protection was a topic that subjects would consistently mention when asked about downloading or not downloading CT apps [14,2].

*Difference between users who did and didn't download a French CT COVID-19 app*

Both groups gave importance to \$DP. Indeed, the factor level of *low data protection* has the biggest beta coefficient of all the levels in the study, rated on average -3.060 and -2.714 points lower in each group. The data suggest that subjects who haven't downloaded such apps place a bit more importance on this requirement.

Moving into \$AP, there are difference between both groups. There is a difference in the effect of the involvement of *universities and government* and *universities* in the providence of app. For subjects who didn't download the app, when the app is provided by *university and government*, the willingness to download is higher compared to those that did download the app. Indeed, the subjects who didn't download show beta coefficients 3 times higher than those who did. Therefore, between subjects who didn't download, universities seem to play a bigger role, compared to subjects who did download

On the other hand, subjects who did download seem to have a better perception of the government as app providers - as the *universities* beta coefficient isn't as big as compared to those who didn't download - but a lower perception of private companies in the same role, as seen by the beta coefficients. The beta coefficient of *private company* for those who downloaded the app is -2.328 (the worst beta coefficient in the whole factor level) compared to a -1.441 for those who didn't download. The beta coefficient of solely *private company* changes when a university is added, with an increase to -1.373 for those who downloaded the app. Yet, this coefficient is still worse when compared with subjects who didn't download the app, whose beta coefficient at this level is -0.647. This could also be that those who didn't download have a higher perception of universities.

If just the levels of \$AP at *private company, government and private company university and private company* are taken, it would be safe to conclude that those who did download the app trust private companies less. The beta coefficients when private companies are involved are always negative, and for the subjects who downloaded are bigger when compared to those who didn't download. bigger between subjects .



Moving on to \$OS, at both levels (L1 and L2) the gamma coefficients of both groups are not statistically significant. Yet, it is important to note that the beta coefficients between the two groups have different signs. For the group of people who didn't download the app, the beta coefficient of proprietary code is -0.139 whereas for those that did download is 0.018.

*Governance* Data protection is a critical requirement for subjects and should be specified to the highest standard when designing an COVID-19 CT app.

The results also let us infer that subjects who have downloaded this app aren't indifferent about their personal data, as both groups have similar beta coefficients. A variable that wasn't measured due to the scope and time of the project was the privacy trade-off phenomenon.

From an app provider perspective, that the app provenance impacts on the willingness to download, particularly if universities are involved. What's interesting is that the involvement of universities seems to have a positive impact in the subjects who didn't download the app, as long as there was no involvement of private companies in the development. Given this results, what specific roles could universities play as app providers? What is it exactly that makes universities a source of trust? These are questions that need further thinking.

Different governance designs could help increase the willingness to download for between those who didn't download a contact tracing app. Based on the data, the p-values and significance of *government* and *university* changes between subjects who did and didn't download. Between subjects who didn't download, its p-value seems to be smaller and have greater significance.

## 5 Threats to validity and future work

Previous research suggest that subjects trust more institutions/systems with which they are familiar [9]. Thus, this can help explain why subjects might trust more apps coming from universities, but doesn't explain the difference at the L2 level.

The external validity of this study could be improved with other sampling techniques, taking into account age groups and cultural differences. This article should be understood within its scope, that it was carried out between students from French universities. Also, D-efficient design could be used to select the vignette sample. Finally, more factors and levels could have been added to the vignettes. Future studies could also research why subjects downloaded the apps and explore cultural differences, to further investigate the privacy trade-off phenomenon and research if conclusions travels between societies.

## 6 Conclusion

This research is an exploratory study of the relationship of DPT with the willingness to download COVID-19 CT apps. The data was collected using the FSE method and analyzed with a random intercept mixed model.

Analysis has shown that app providers and data protection are statistically significant. Furthermore, subjects who downloaded and didn't download give different importance to both data protection and app provider. For those who didn't download, having a *university* (with or without *government*) involved in the creation of the app is statistically significant. Indeed *university's* beta was 0.839, making it the highest positive beta of the whole study. From this, we can infer that including universities in the providence of apps can have a positive impact on the willingness to download between subjects who didn't download TousAntiCovid.

## References

1. Aljazzaf, Z.M., Perry, M., Capretz, M.A.: Online trust: Definition and principles. In: 2010 Fifth International Multi-conference on Computing in the Global Information Technology
2. Altmann, S., Milsom, L., Zillesen, H., Blasone, R., Gerdon, F., Bach, R., Kreuter, F., Nosenzo, D., Toussaert, S., Abeler, J.: Acceptability of app-based contact tracing for covid-19: Cross-country survey study. *JMIR mHealth and uHealth* **8**(8) (2020)
3. Atzmüller, C., Steiner, P.: Experimental vignette studies in survey research. *Methodology: European Journal of Research Methods for The Behavioral and Social Sciences* **6** (2010)
4. Auspurg, K., Hinz, T.: Factorial survey experiments, vol. 175. Sage Publications (2014)
5. Bano, M., Zowghi, D., Arora, C.: Requirements, politics, or individualism: What drives the success of covid-19 contact-tracing apps? *IEEE Softw.* **38**(1) (2021)
6. Bhatia, J., Breaux, T.D., Reidenberg, J.R., Norton, T.B.: A theory of vagueness and privacy risk perception. In: 2016 IEEE 24th International Requirements Engineering Conference (RE) (2016)
7. Corritore, C.L., Kracher, B., Wiedenbeck, S.: On-line trust: concepts, evolving themes, a model. *International journal of human-computer studies* **58**(6) (2003)
8. Hibshi, H., Breaux, T.D., Broomell, S.B.: Assessment of risk perception in security requirements composition. In: 2015 IEEE 23rd International Requirements Engineering Conference (RE) (2015)
9. Hoff, K.A., Bashir, M.: Trust in automation: Integrating empirical evidence on factors that influence trust. *Human factors* **57**(3) (2015)
10. Jonker, M., de Bekker-Grob, E., Veldwijk, J., Goossens, L., Bour, S., Rutten-Van Mölken, M.: Covid-19 contact tracing apps: Predicted uptake in the netherlands based on a discrete choice experiment. *JMIR mHealth and uHealth* **8**(10) (2020)
11. Mayer, R.C., Davis, J.H., Schoorman, F.D.: An integrative model of organizational trust. *Academy of management review* **20**(3) (1995)
12. McKnight, D.H., Choudhury, V., Kacmar, C.: Developing and validating trust measures for e-commerce: An integrative typology. *Information systems research* (2002)
13. Negri Ribalta, C.: A method to deal with social bias and desirability in ethical requirements. In: Fischbach, J., Condori-Fernández, N., et al. (eds.) Joint Proceedings of REFSQ-2022 Workshops, Doctoral Symposium, and Poster & Tools Track (2022)
14. Simko, L., Calo, R., Roesner, F., Kohno, T.: Covid-19 contact tracing and privacy: Studying opinion and preferences (2020)
15. Williams, S.N., Armitage, C.J., Tampe, T., Dienes, K.: Public attitudes towards covid-19 contact tracing apps: A uk-based focus group study. *Health Expectations* (2020)