AI in Public: The Effects of Technology Bias, Fears of Public Surveillance, and Moral Tradeoffs on Privacy Concerns

Abstract

Applications of AI in public surveillance contexts fuel polemics among consumers and public policy makers alike. In two experimental studies, we explore the mechanisms that affect citizens' attitudes towards government surveillance technologies. In Study 1, we show that the privacy and surveillance concerns are reduced when government (vs. firm) owns the data. Moreover, the fear of technology biases moderates the relationship between privacy concerns and willingness to adopt. In Study 2, we analyze the potential of anonymization of data collection to remedy the perceived privacy concerns. We find that the effect of anonymization of data collection on the willingness to support government surveillance technology goes through two parallel antecedents of privacy concerns: a reduction in perceived government intrusiveness and an increase in the perceived fairness and justice. Reduced privacy concerns ultimately increase the perceived usefulness of technological solution and increase the willingness to adopt.

Keywords: AI surveillance technology, privacy fears, anonymization

Track: Social Responsibility and Ethics

1. Introduction

"While there is much to be appreciated and valued in a technology-integrated society, there is much that is unknown or not well understood and concerning" (Walker et al., 2019 p. 403). The increased proliferation of smart technology solutions has transcended business domains. New artificial intelligence (AI) deployments are reshaping numerous areas of public administration, health system management, and local municipality services. While benefits and costs of AI have been discussed in all areas of consumer experiences (Puntoni et al., 2021), there has been a lack of research on idiosyncrasies of public AI implementations and their potential impact on society. AI solutions in public context have often been related to government surveillance technologies that threaten the rise of surveillance society, a digitally-enabled panopticon that allows government institutions to monitor and control citizens in increasingly more intrusive ways (van Zoonen 2016; Nam 2019). The fears of a surveillance society push regulators and policy makers to define yet unregulated risks and protect both citizens and core social values. In the new proposal to lay down harmonized rules on AI (Artificial Intelligence Act), European Commission explicitly forbids the introduction of any type of AI that may be used for social scoring and asks for strong compliance and AI conformity assessment for many other types of public AI use (EUR-Lex, 2021).

This study evaluates drivers and mechanisms that play a role in citizens' evaluations of public implemented AI. Existing marketing literature mostly focuses on privacy concerns' impact on disclosure of information in consumer-firm transactions (Okazaki et al, 2020). However, there is a lack of empirical evidence in studies exploring the mechanisms that drive cost-benefit tradeoffs in the public context, particularly those that go beyond the individual privacy concerns and addresses public safety and governmental concerns (Mulligan, Regan & King, 2020; Walker et al., 2019). In two studies, we explore the drivers of citizens' acceptance of public AI solutions, perceptions about the interactions between government and third parties in controlling the data, and propose recommendations on how policymakers can improve the adoption of government technologies.

2. Theoretical Rationales

Government implemented technologies are largely viewed as intrusive because of their adverse impact on individuals' privacy (Dinev et al.,2008; Fox et al., 2021). Privacy concerns have a central role in the technology adoption. Higher privacy concerns are shown to

significantly decrease positive evaluation, trust, willingness to disclose information, purchase, while they significantly increase perceived risk and protection behaviors (Okazaki et al., 2020). However, privacy has a strong contextual nature (people accept being monitored at airport but do not accept being monitored in their street) and is strongly affected by the perceived purpose of data collection (e.g., to improve shopping experiences online versus to maintain social order). For this reasons, standard findings in privacy literature in the business context cannot be easily transferred to public setting.

2.1.1 Privacy as an Individual versus Social Value

"Privacy is becoming a flashpoint in the surveillance economy, yet the concerns causing the fire go well beyond privacy. The harms expand beyond those construed by privacy (i.e., beyond information control)" *(*Mulligan et al., 2020, p. 771). The main problem in understanding the social impact of privacy is that privacy is primarily investigated as a preexisting individually-oriented right or preference, and not as a negotiated social construct defined in a dialog with others and socio-technical environment in which privacy desires are constructed (Nissenbaum, 2009; Mulligan et al., 2020). Consequently, in marketing as well, privacy has been primarily seen as an individual's right to share or withhold an information (typically with a firm and, in extension, with a government service as well) (Westin, 1967).

On the other hand, surveillance fears are primarily a social phenomenon that require different framing and measuring of privacy (Nissenbaum, 2009). People are uncertain about the nature of their privacy trade-offs as well as their own preferences over them (Acquisti et al., 2015). The context of security/privacy trade-offs in public is especially a complex one because the privacy harms are often intangible and uncertain, while security risks are seen as low and constant over time. People are in general uncertain about what data has been collected and how it might be used, or what exact harms may befall them. In public contexts, people are likely to diminish the perceived risk that some security issue would occur that would warrant surveillance intrusion (Acquisti et al., 2015). Because of these biases, citizens undermine the uncertain and delayed benefits of government technology implementations (Acquisti et al., 2020) and weight them against highly accessible image of threats of surveillance state formed by science fiction movies and sensationalistic nature of social media and some public media stories (Kaplan & Haelein, 2020).

2.2. Privacy Intrusion Versus Fears of Surveillance, Algorithmic Bias or Distrust in Government

Due to the underlying weaknesses of analyzing privacy concerns in public settings through the lenses of personal control over information, previous studies tend to either ignore or bundle together with privacy various other fears like: fears of surveillance, fears of algorithm fallibility and capacity for discrimination (Jago & Laurin, 2021), perceived inaptness for algorithms to make morally-relevant decisions (Dietvorst & Bartels, 2021; Nagtegaal, 2021), perceived trust in government to use the technology to surveil citizens (Nam, 2019) and the perceived procedural (un)fairness and justice (Yalcin et al., 2020). Indeed, the few existing studies show mixed results. In the general context of using Internet, the fear of government intrusion and the perceived need for government surveillance significantly impact the willingness to provide personal information required to complete transactions on the Internet (Dinev et al.,2008). Conversely, in the context of adopting privacy tracing app among Irish citizens, privacy concerns do not influence acceptance of app prior or post launch (Fox et al.,2021) At the same time, more than 53% of US population is unwilling to install a COVID tracing app that would provide information to public health authorities (Acquisti et al., 2020). Potential reason for these disparate findings could be the self-selection of citizens and the lack of understanding of the mechanisms affecting privacy concerns in public contexts.

Many of the fears in public contexts transcend the typical questions used to explore privacy concerns or willingness to disclose information (cf. Okazaki et al., 2020). The set of issues related to fears of *technology bias* are often grouped under the terms of *algorithmic aversion* (fears and objections to automated decision-making by algorithms) and *algorithmic bias* (the fears that algorithms may mirror and augment the racial and other discriminatory human biases) (Puntoni et al, 2020; Walker et al., 2019).

Very few existing studies disentangle some of these fears and trade-offs. Previous studies differentiate between privacy concern and intrusion or surveillance concerns (Nam, 2019; Dinev et al., 2008). Nam (2019) found that surveillance concerns and perceived acceptability of government retaining records over time affects the acceptability of government surveillance among US citizens much more than general information privacy intrusion measures. Studies on algorithmic aversion and algorithmic bias often use business context, but when public versus business applications are considered, the studies suggest that fears of technological biases are

stronger in public context, particularly when the algorithmic decision-making concerns morally or ethically complex decisions, like justice, crime recidivism, hiring or promotion at work (Nagtegall, 2021; Yacin et al., 2020, Araujo et al., 2020).

2.3 Firms versus Government Use of Data

Privacy and government surveillance studies emphasize the fear of being observed/surveilled (Nam, 2019, Van Zoonen, 2019), which may heighten the negative potential consequences of public surveillance. Moreover, consumers are largely unaware of the diverse ways in which firm collect and use their data themselves and with third-parties, under-weigh potential costs to privacy, show higher illusion of control over data choices, and follow the common herding bias believing that it is safe to share and disclose information because many other people do the same (Acquisti et al.,2020). Based on the evidence from privacy heuristic and surveillance economy literature, we would predict that people would be more likely to accept firm rather than government implementation of AI technologies. On the other hand, few studies (outside of marketing) and surveys of citizens suggest the opposite. Dutch citizens have shown higher willingness to share the data collected using smartphone sensors (GPS, camera, wearables) with government sponsor than a market research firm (Singer & Couper 2011; Struminskaya et al., 2020). There are two possible explanations for this phenomenon. On one side, people ignore the risks and they lack knowledge over situation, which would foster their feelings of dependance on the government which in turn increase the system justification and government trust (Shepherd and Kay, 2011). On the other side, people might appreciate convenience and ease of use. For example, people (74%) accepted potential corporate and government surveillance of crypto-currency wallets in exchange for the ease of use (Catalini & Tucker, 2016). In this context, "citizens may still accept government surveillance as necessary evil, despite being aware that surveillance can threaten civil liberties" (Nam, 2019). The explanations of the directions of the effects and the impact of data ownership on citizen perceptions are the essential questions in our first study.

3. Overview of the Studies

To explore the mechanisms that drive citizens' willingness to accept public surveillance, we conducted two experimental studies using Prolific survey panel platform during spring-summer 2021. In Study 1, we test how the perceptions of who owns the data (government or firm) together with the fear of algorithmic biases affects privacy concerns, willingness to adopt

government surveillance technology and perceived wellbeing. In Study 2, we explore potential remedies that regulators can introduce to encourage citizens' willingness to support government surveillance technology. Namely, we examine the impact that anonymization of government data collection has on government intrusion perceptions, privacy concerns, perceptions of fairness and justice (moral equity), which affects perceived usefulness and willingness to support. In both Studies, we use the context of government surveillance technology introductions to monitor COVID-19 contract tracing. We use examples of introducing contact tracing government app (Study 1) and introduction of smart camera surveillance of public transport facilities to monitor citizen's compliance with the face-wearing regulations (Study 2). The context allows us to have participants with more concrete experiences on surveillance technologies. Both studies aim to provide better understanding of the antecedents of privacy concerns and mechanisms that affect citizens' attitudes towards government surveillance technologies.

3.1. *Study 1*

In Study 1, we aim to test the effect of government versus firm data ownership and its interaction with fear of technology biases on privacy concerns and willingness to adopt a government contact tracing app.

We recruited 400 participants (41% female, 1% non-binary; M_{age} = 25.6, SD_{age} = .09) on Prolific to participate in a 2(bias awareness: high, low) × 2(data owner: company, government) between-subjects design. At the beginning of the study, participants read a newspaper-like article discussing the decision of the Health Ministry of their country to introduce an AI system to collect data on users' infection, location, and movement, accessible through a mobile app. The conditions of data collection are such that: either the government or a high-tech company would be responsible for storing and analyzing the data. The data will be automatically deleted every 30 days. The adoption of the new mobile contact-tracing app would be voluntary. After reading about the purpose for the new AI system and the way in which the system would work, at the end of the article, participants read an expert opinion about AI being quite neutral versus biased tool policy makers could use to make decision. was given. After the article, participants reported their willingness to adopt the app. In addition, participants answered questions about their privacy concerns (Dinev et al., 2008), need for control over personal data (Lo, 2010), government intrusion concerns (Dinev et al., 2008), perceived need for government surveillance (Dinev et al., 2008), trust in government (Poznyak et al., 2013), AI literacy (Long, & Magerko, 2020), trust in

AI (Lee, 2018), fear of Covid-19. We also asked participants to report their perceived fear of AI bias (Mason et al., 2014). At the end of the study, participants reported some demographic information (e.g., gender, income, country of origin), and they answered the manipulation and attention checks.

3.1. *Results Study 1*

We excluded 79 participants who failed the attention checks in the study, leaving a final sample of 321 participants. First, we conducted a 2(bias awareness: high, low) × 2(data owner: company, government) ANOVA on the likelihood of adopting the app. The analysis revealed a significant main effect of data ownership ($F(1, 321) = 5.403$, $p = .021$). In particular, participants were more likely to adopt the app when the government was the owner of the data ($M = 5.06$, $SD = 1.30$) rather than the high-tech company ($M = 4.69$, $SD = 1.51$). The main effect of awareness of bias and the interaction effects between data ownership and bias awareness were not significant.

To show how fear of technology bias affect citizens' attitudes, we performed a moderation analysis with perceived fear of technology bias as the main moderator in the relationship between data ownership and willingness to adopt, using Process macro in SPSS (model 1; Hayes, 2017). Results revealed a significant main effect of data ownership ($b = 0.3182$, $SE = 0.1100$, 95% CI [0.0300, 0.6064], a main effect of fear of technology bias ($b = - 0.5946$, $SE = 0.0926$, 95% CI [- 0.7767, - 0.4125], and the critical interaction between fear of technology bias ($b = 0.2900$, $SE = 0.1293$, 95% CI [0.0355, 0.5444]. We probed the interaction between data ownership and fear of technology bias on adoption using the Johnson–Neyman technique (Hayes, 2017). Results of this analysis revealed that people with high fear in technology bias are more likely to adopt the app when the government is the owner of the data. For people with low fear in technology bias, there is no significant differences in adoption based on who owns the data.

Finally, we performed a moderated mediation analysis showing that the effect of data ownership on adoption is mediated by privacy concern and moderated by the fear of technology bias using Process (model 14; Hayes, 2017). In particular, people feel less privacy concerns when the government owns the data ($b = - 0.5403$, $SE = 0.1565$, 95% CI [- 0.8483, - 0.2323]), which in turn leads to higher willingness to adopt the technology ($b = - 0.4451$, $SE = 0.0598$, 95% CI [- 0.5628, - 0.3274]). The effect of privacy concerns on adoption is significantly moderated by perceived bias of the technology (Index of Moderated Mediation = 0.0564, BootSE = 0.0294,

95% CI [0.0093, 0.1225]). We probed the interaction between privacy concerns and perceived bias of the technology on adoption using the Johnson-Neyman technique. Results of the analysis revealed that if fear in technology bias is high, an increase in privacy concerns will lead to a decrease of adoption intentions. However, if fear in technology bias is low then a decrease in privacy concerns would lead to higher willingness to adopt. Overall, the results indicate that data ownership can reduce privacy concerns and lead people to adopt more surveillance technologies. However, for people who have high fear in technology bias, the ownership of data might not be enough. For this reason, in study 2, we propose an alternative instrument that governments can use to support adoption of surveillance technologies.

3.2. *Study 2*

In Study 2, we recruited 600 people (66% female, 2% non-binary) on Prolific, in June 2021. We employed a 2 (intrusiveness: low, high) between-subjects design. At the beginning of the study, participants read about the current spread of Covid-19 and the proposition of their government to introduce a solution to reduce the spread. In particular, we explained to participants that their government was planning to use an artificial intelligence system, embedded in surveillance cameras, to monitor whether people were wearing the mask properly in public transport systems. In the high (*low*) intrusiveness condition, the system was (*not*) able to identify the person (i.e., the anonymization of facial, gender and race characteristics of collected data was explained in the low intrusiveness setting). Participants also saw graphical illustrations of how the systems would work. Our main dependent variable was participants' willingness to support the introduction of the technology in their country. Participants also answered questions related to privacy concerns (Dinev et al., 2008), perceived usefulness of solution (Anton et al., 2021), moral equity in term of perceived fairness and justice (Anton et al., 2021), government intrusion concerns (Dinev et al., 2008), perceived need for government surveillance (Dinev et al., 2008), trust in government (Poznyak et al., 2013), and fear of Covid-19. At the end of the study, participants reported some demographic information (e.g., gender, income, country of origin), and they answered the manipulation and attention checks.

3.4. Results Study 2

We excluded 189 participants who failed the attention checks in the study, leaving a final sample of 411 participants. A one-way ANOVA showed that participants in the low intrusiveness condition (M = 3.93, SD = 1.92) were more willing to support the introduction of the surveillance

system than participants in the high intrusiveness condition (M = 3.45, SD = 2.04; F(1,410) = 6.100, p = .014). The results corroborate the idea that anonymization could be a valuable instrument for policy makers to increase acceptance of surveillance technologies. Interestingly, we show that the effect of anonymization on willingness to support goes through the impact that anonymization has on two parallel antecedents of privacy concerns, rather than through its direct impact on privacy concerns and usefulness of technology. A parallel mediation analysis in Process macro for SPSS (model 4; Hayes, 2017) shows that anonymization of data collection reduces government intrusions concerns (b = - 0.3386, SE = 0.1294, 95% CI [- 0.5928, - 0.0843]) which, in turn, decreases privacy concerns (b = 0.3576, SE = 0.0412, 95% CI [0.2767, 0.4386]). In parallel, anonymization increases perceived moral equity (fairness and justice) of the AI solution (b = 0.3695, SE = 0.1879, 95% CI [0.0001, 0.7390]), that again decreases privacy concerns (b = - 0.2913, SE = 0.0283, 95% CI [- 0.3470, - 0.2356]). Ultimately, the decrease in privacy concerns due to anonymization leads to the increase in perceived usefulness of government intervention (b = - 0.5650, SE = 0.0614, 95% CI [- 0.6856, -0.4443]) and increase in the willingness to support its introduction (b = 0.7279, SE = 0.0372, 95% CI [0.6547, 0.8011]). Importantly, both indirect effects of anonymization through government intrusion concerns (- 0.1211, BootSE = 0.0499, 95% CI [- 0.2236, - 0.0294]) and moral equity (- 0.1076, BootSE = 0.0565, 95% CI [- 0.2273, - 0.0018]) on privacy concerns were significant, while the direct path from anonymization on privacy concerns was not significant (- 0.1899, SE = 0.1006, 95% CI [- 0.3876, 0.0078]), indicating a full mediation.

4. Conclusions

This paper aims to explore the interplay between privacy concerns, surveillance fears and fears of technological biases on the citizens' willingness to support government surveillance technologies. Moreover, we analyze the potential of governments to improve the acceptance of public AI applications and reduce the perceived government intrusion concerns through the anonymization of data collection efforts and ownership of the collected data. We show that in surveillance contexts citizens appreciate more the government data collection and AI deployment than firms', despite the inherent fears of being surveilled. Our findings also emphasize the need for disentangling various AI-related concerns from the privacy concerns in public domains.

5. References

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2015). Privacy and human behavior in the age of information. Science, 347(6221), 509-514.

Acquisti, A., Brandimarte, L., & Loewenstein, G. (2020). Secrets and likes: the drive for privacy and the difficulty of achieving it in the digital age. Journal of Consumer Psychology, 30(4), 736-758.

Anton, E., Kus, K., & Teuteberg, F. (2021). Is Ethics Really Such a Big Deal? The Influence of Perceived Usefulness of AI-based Surveillance Technology on Ethical Decision-Making in Scenarios of Public Surveillance. In Proceedings of the 54th Hawaii International Conference on System Sciences (p. 2121).

Catalini, C., & Tucker, C. (2016). Seeding the S-curve? The role of early adopters in diffusion (No. w22596). National Bureau of Economic Research.

Dietvorst, B. J., & Bartels, D. M. (2021). Consumers object to algorithms making morally relevant tradeoffs because of algorithms' consequentialist decision strategies. Journal of Consumer Psychology.

Dinev, T., Hart, P., & Mullen, M. R. (2008). Internet privacy concerns and beliefs about government surveillance–An empirical investigation. The Journal of Strategic Information Systems, 17(3), 214-233.

Fox, G., Clohessy, T., van der Werff, L., Rosati, P., & Lynn, T. (2021). Exploring the competing influences of privacy concerns and positive beliefs on citizen acceptance of contact tracing mobile applications. Computers in Human Behavior, 121, 106806.

Jago, A. S., & Laurin, K. (2021). Assumptions About Algorithms' Capacity for Discrimination. Personality and Social Psychology Bulletin, 01461672211016187.

Kaplan, A., & Haenlein, M. (2020). Rulers of the world, unite! The challenges and opportunities of artificial intelligence. Business Horizons, 63(1), 37-50.

Lee, M. K. (2018). Understanding perception of algorithmic decisions: Fairness, trust, and emotion in response to algorithmic management. Big Data & Society, 5(1), 2053951718756684.

Lo, J. (2010). Privacy Concern, Locus of Control, and Salience in a Trust-Risk Model of Information Disclosure on Social Networking Sites. In AMCIS (p. 110).

Long, D., & Magerko, B. (2020, April). What is AI literacy? Competencies and design considerations. In Proceedings of the 2020 CHI Conference on Human Factors in Computing Systems (pp. 1-16).

Mason, O. J., Stevenson, C., & Freedman, F. (2014). Ever-present threats from information technology: the Cyber-Paranoia and Fear Scale. Frontiers in Psychology, 5, 1298.

Mulligan, D. K., Regan, P. M., & King, J. (2020). The Fertile Dark Matter of Privacy takes on the Dark Patterns of Surveillance. Journal of Consumer Psychology, 30(4), 767-773.

Nagtegaal, R. (2021). The impact of using algorithms for managerial decisions on public employees' procedural justice. Government Information Quarterly, 38(1), 101536.

Okazaki, S., Eisend, M., Plangger, K., de Ruyter, K., & Grewal, D. (2020). Understanding the strategic consequences of customer privacy concerns: A meta-analytic review. Journal of Retailing, 96(4), 458-473.

Poznyak, D., Meuleman, B., Abts, K., & Bishop, G. F. (2014). Trust in American government: Longitudinal measurement equivalence in the ANES, 1964–2008. Social Indicators Research, 118(2), 741-758.

Shepherd, S., Kay, A. C., Landau, M. J., & Keefer, L. A. (2011). Evidence for the specificity of control motivations in worldview defense: Distinguishing compensatory control from uncertainty management and terror management processes. Journal of Experimental Social Psychology, 47(5), 949-958.

Struminskaya, B., Lugtig, P., Keusch, F., & Höhne, J. K. (2020). Augmenting surveys with data from sensors and apps: Opportunities and challenges. Social Science Computer Review, 0894439320979951.

Walker, K. L., Milne, G. R., & Weinberg, B. D. (2019). Optimizing the future of innovative technologies and infinite data.

Westin, A. F. (1967). Special report: legal safeguards to insure privacy in a computer society. Communications of the ACM, 10(9), 533-537.