The AAAI 2017 Spring Symposium on
Designing the User Experience of Machine Learning Systems
Technical Report SS-17-04

Privacy by Design in Machine Learning Data Collection: A User Experience Experimentation

Jonathan Vitale, Meg Tonkin, Suman Ojha, Mary-Anne Williams
Centre For Artificial Intelligence – Innovation and Enterprise Research Lab (The Magic Lab)
University of Technology Sydney, 15 Broadway, Ultimo NSW 2007 – Australia

Xun Wang, William Judge
Commonwealth Bank Innovation Lab, 11 Harbour St, Sydney NSW 2000 – Australia

Abstract

Designing successful user experiences that use machine learning systems is an area of increasing importance. In supervised machine learning for biometric systems, such as for face recognition, the user experience can be improved. In order to use biometric authentication systems, users are asked for their biometric information together with their personal information. In contexts where there is a frequent and large amount of users to be enrolled, the human expert assisting the data collection process is often replaced in favour of software with a step-by-step user interface. However, this may introduce limitations to the overall user experience of the system. User experience should be addressed from the very beginning, during the design process. Furthermore, data collection might also introduce privacy concerns in users and potentially lead them to not use the system. For these reasons, we propose a privacy by design approach in order to maximize the user experience of the system while reducing privacy concerns of users. To do so we suggest a novel experiment in a Human-Robot interaction setting. We investigate the effects of embodiment and transparency on privacy and user experience. We expect that embodiment would enhance the overall user experience of the system, independently from transparency, whereas we expect that transparency would reduce privacy concerns of the participants. In particular, we forecast that transparency, together with embodiment, would significantly reduce privacy considerations of participants, thus maximising the amount of personal information provided by a user.

Introduction

Data collection, like the name suggests, is the process of gathering and measuring information. This is a necessary task for every machine learning system (Chen, Mao, and Liu 2014), and collecting data in an *efficient* and *responsible* manner is extremely important, especially in biometric authentication systems (Wayman et al. 2005). We propose a novel experimental methodology to test the user experience of data collection interfaces for face recognition systems, while at the same time gathering insights for dealing with privacy issues involved in the process. In developing the proposed data collection system we make use of a *privacy by design* approach. We plan to situate the experiment in a bank. Since their customers usually know they are protected by laws when asked for sensitive data (Turner, Zavod, and Yurcik 2001), this is a good context to test the effects on privacy concerns. In this paper we focus our investigation to supervised machine learning systems for face recognition. These systems provide on-line enrolment processes on first contact of the user with the biometric system. These are used to gather biometric samples of the user's face and related personal details.

A supervised machine learning algorithm, such as a Deep Convolutional Neural Network (Krizhevsky, Sutskever, and Hinton 2012), uses labelled training data to teach itself how to classify new observations in the correct classes. For the huge amount of data required for learning, the process of labelling can be efficiently realized using crowd-sourcing. However, the benefits of this methodology are limited to data gathered for off-line training, thus becoming unsuitable for on-line data collection. For instance, a face recognition system needs to register and enrol a new face and the related details of the person every time a new user is to be recognised by the system. Since users generally do not possess any previous experience with the system they are being registered with, the first measurement is often guided by a professional who explains the use of the biometric reader (Matyáš and Říha 2002). This Human-Human interaction can indeed facilitate the data collection process, making it more effective and efficient. However, in order to minimise the associated costs, this solution is often replaced in favour of the provision of software having a guided graphic user interface (GUI).

More importantly, the data to be collected can contain sensitive information. Thus, the data collection process can become a potential threat for users' privacy (Wayman et al. 2005). No matter how appealing and usable is the system; if users do not trust a system, they are likely not to use it (Pearson 2009). This problem is particularly significant when the user is aware of the data collection process and it is necessary to gain their consent to collect their private information.

Privacy issues that arise in real-time data collection can be addressed explicitly, by providing specific guidelines aimed at reducing the amount of sensitive data stored in the system (Fung et al. 2010) or by using encryption algorithms to safely store data (Basharat, Azam, and Muzaffar 2012). In addition, privacy can be implicitly addressed by the ma-

Copyright © 2017, Association for the Advancement of Artificial Intelligence (www.aaai.org). All rights reserved.

chine learning algorithm itself. For instance, the algorithm can learn which features to extract from inputs after a training phase. This new representation is then extracted from new observations in order to classify them with competitive recognition rates. However, the extracted features might not contain enough information for reconstructing the original input back (Amos, Ludwiczuk, and Satyanarayanan 2016), thus relieving privacy concerns. Despite these important precautions, the majority of naïve users do not know the underlying details of the machine learning algorithm used, what data would be stored by the system, and how. Thus, this lack of transparency might impact users' trust and lead to, possibly unfounded, privacy concerns.

Given the significance of these issues, in this paper we propose a methodology to investigate the effect that *embodiment* and *transparency* have on the *user experience* and on the *privacy considerations* of users during real-time data collection.

We expect that embodying a face enrolment system in a robotic platform, and enhancing the transparency of the underlying machine learning system through an appropriate GUI would mitigate users' privacy concerns, while at the same time providing a better user experience.

Privacy by Design in Human-Machine Interactions

A privacy by design approach ensures privacy protections are built by default into a system from the beginning and this may also be considered good business practice (Williams 2009). Existing systems built without including privacy considerations as a core part of their development often result in poor privacy management. This may be experienced as a disparity between stated privacy policies and actual privacy controls (Anthonysamy, Greenwood, and Rashid 2013).

Many governments address privacy through guidelines and laws. For example, Australia provides a principal based law for the use, collection and management of personal data. The Australian Privacy Principles (APPs), included in Schedule 1 of the Privacy Act 1988[1], state that "entities manage personal information in an open and transparent way". Indeed, labelling and storage of sensitive biometric information, such as a biometric template for machine learning of facial information, is specifically included in the APP. Users must provide consent before an entity is allowed to store personal information about them and must be informed as to what the stored information will be used for. Hence, designing a system accounting for these requirements, while also targeting the best user experience, becomes crucial.

In the context of institutions like banks, academia, corporate or government departments, information collected from people (say, customers) can have significant privacy issues. For instance, it is not ethical for such organizations to deceptively collect personal information that might be exposed to the public or other technical employees of the organization. Sometimes it might be essential to provide (disclose) the information for analytical and reporting purposes but institutions "have to evaluate the potential ethical or unethical use of disclosed information" (Turilli and Floridi 2009). As such, it is important that the organization enables a mechanism of full transparency in the information collection procedure. "Information transparency is not an ethical principle per se" (Turilli and Floridi 2009, pag. 107) but having a less transparent method of information collection from people might impose some privacy concerns. This might in turn prevent them from willingly providing some information which they think is private for them and may be disclosed by the organization collecting the information. Hence, it is the requirement of an organization to have a more transparent mechanism of collecting information and explaining how the information will be stored and used so that the people providing the information can make suitable choices and are not concerned about threats to their privacy (Milne 2000).

Making use of a human assistant to collect the data while providing information about the system, the data collected and the rights of the person can enhance users' trust and relieve privacy concerns. However, this solution is not always feasible in machine learning systems, because of its high costs. Thus, often the alternative is to provide software to assist the user during the data collection process. It has been shown that robots can be an effective alternative to human-human interactions, because of their embodiment, which simple computer based applications cannot exhibit (Kidd and Breazeal 2004). In fact, robots can speak, gesture, direct their gaze, *etc.*, thus resembling a social human interaction and facilitate the data collection process, while providing a greater sense of trust and reliability than a simple disembodied software (Kidd and Breazeal 2004). Hence, similarly to transparency, investigating the impact of embodiment in privacy by design settings becomes crucial too.

In the following section we provide an experimental methodology designed to test the effect of embodiment and transparency on privacy and user experience in the context of a face enrolment system situated in a bank.

Proposed Methodology

Participants

We propose to design a preliminary survey to select our participants. We aim to include a population homogeneously ranging from 18 to 45 years old, evenly split into males and females, and balanced in terms of computer literacy. Preconditions for access to this experiment would be to be proficient in English reading and listening, and own a profile on Facebook, Instagram, LinkedIn and Twitter.

Used Material

The experiment is a comparative study using two settings. The first setting involves using a robot guiding users during a face enrolment process. In the second setting participants experience the same face enrolment process without the involvement of a robot.

Robotic Platform. We use a PAL REEM service robot[2], a wheel-based humanoid robot equipped with a pair of stereo

[1] https://www.legislation.gov.au/Details/C2016C00838

[2] http://pal-robotics.com/wp-content/uploads/2016/03/REEM-Datasheet.pdf

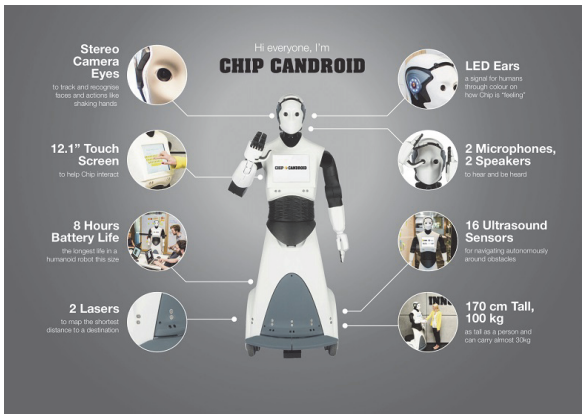

Figure 1: Technical details of the used robot platform for the embodied setting of the experiment.

cameras on its head. The robot also has a pair of 4-degrees of freedom (DOF) arms and 7-DOF hands that can perform human-like gestures. More importantly, the robot has a built-in touch screen, used during the experiment to display the graphic user interface of the face enrolment system (Fig. 1).

The REEM touch screen is used as the main interface to gather user inputs during the first setting of the experiment (Fig. 2). To entice a user to provide face enrolment information, the robot will perform various arm motions and audio instructions/cues in addition to the visual information displayed on its screen. The user interactions with the touch screen are recorded.

Disembodied Platform. As a comparison, the second settings of the experiment involves a face enrolment setup without the use of the robot. Instead of a robot, we use a tablet (an iPad mobile device) and a camera fixed on a blank wall (Fig. 2) mimicking the same robot physical configuration. The GUI on the touch screen is exactly the same as the one on the REEM robot. The same audio instructions are provided through the fixed mobile device, but no gestures would be available.

Face Recognition System. The experiment uses a face recognition system that is built upon the state-of-art OpenFace implementation (Amos, Ludwiczuk, and Satyanarayanan 2016). A face detector is employed to extract faces from an image first. Passing through a pretrained Deep Neural Network model, the face image patches are transformed into 128 dimension feature vectors for (face) classification. That is, raw images are no longer required for training a face classifier, since the biometric information of faces are fully encoded into vectors. Furthermore, we can experiment with a variety of machine learning techniques, *e.g.* support vector machines, to improve the performance of the face recognition system. The face enrolment process in our system is designed to be user self-driven. That is, the enrolment system shows to the user the captured face image through REEM's or iPad's camera, and the user is asked to self-annotate their captured facial image (*i.e.* providing the name and other personal information).

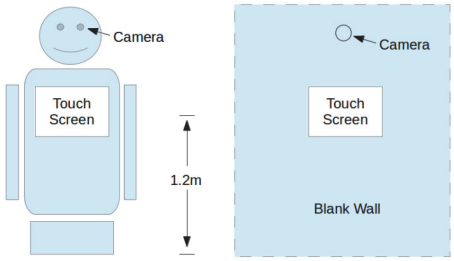

Figure 2: Face enrolment setup on the REEM robot (left) and the alternative disembodied setup on a blank wall (right).

Procedure

In the experiment we will consider two independent variables, namely *embodiment* and *transparency*. These variables are binary (*i.e.* either embodied or disembodied, and transparent or not transparent), thus leading to a 2×2 between-subjects design. The participants would be evenly divided for each of the four settings (*i.e.* embodiment with transparency, embodiment without transparency, disembodiment with transparency and disembodiment without transparency).

We measure the user experience with the data collection system through the User Experience Questionnaire (UEQ), (Laugwitz, Held, and Schrepp 2008) we measure the privacy consideration of users by counting the number of users allowing the system to collect images of their face, and, among them, the number of non compulsory information provided during the data collection process.

This additional information would include access to their social networks accounts. The reason behind this option is that a bank is usually perceived as a reliable institution by users in safely storing their personal information (Turner, Zavod, and Yurcik 2001). Hence asking users to provide simple additional information (*e.g.* address, telephone number, *etc.*) may not elicit privacy concerns by users. In a preliminary survey we conducted of 282 people (bank staff, customers, students), we discovered that a robot collecting information such as name, address and phone number was considered to be 'okay' (36.82%) or 'definitely okay' (20.94%) in a bank. On the contrary, we expect that by asking the users to provide access to their social networks accounts would be perceived as an intrusion outside the normal boundaries perceived appropriate for a bank institution. Indeed in our same survey, only 8.96% of users expressed to be 'okay' and 4.48% 'definitely okay' with the robot in a bank asking to connect to social networks. As such, by asking to connect on social networks we expect to create an impact on privacy concerns in the majority of users, allowing to efficiently measure the effects of embodiment and transparency on privacy.

The participants would stand in front of the enrolment system and interact with it, which would be either embodied (robot) or disembodied (tablet), depending on the experimental setting they belong to, and decide if providing their consent in saving their details and additional information. In

this first version of the experiment we do not investigate the effects on personification of the robot, thus not focusing on its personality or emotional capabilities.

During the transparent condition only, the data collection system would explain how the machine learning system is realized through a simple and user friendly example. Right after taking a picture of the user's face, and before asking user's consent, we will present the picture of a celebrity which is the closest match for the retrieved picture of the user. The enrolment system would suggest that, given the present state of knowledge of the system, the user looks like such celebrity. We will explain through visual examples that the underlying system represents and save pictures of faces similarly to spatial coordinates (in order to allow everybody to understand the system we will use a 2-dimensional space as example). Hence, we will show a 2-dimensional plane with the picture of the user just taken together with a limited neighbourhood of other celebrities' faces in order to provide a visual example of the Euclidean distances between faces in such 2-dimensional space. The system would tell the user that the picture taken would not be stored 'as-if', but only as an abstract coordinate of such space, thus reducing risks for user's privacy.

Hypotheses

From the analysis of the results we would expect that:

- Embodiment of the system would positively impact on the user experience of the system, in both the transparent and not transparent conditions;
- Transparency of the system would significantly increase the number of users giving consensus for storing their face;
- Transparency in the embodied condition would increase the number of users releasing additional information about their social network accounts.

We believe that a robot platform having a humanoid form would encourage users to interact with natural interactions, as it happens with a human expert assisting the process. Thus, this would increase the ease of use of the system and enhance the overall user experience of the system. Furthermore, previous studies demonstrated that an embodied robot is perceived more trusting and reliable than virtual agents or computer applications (Kidd and Breazeal 2004). Hence, embodiment would not only impact on user experience, but also on trust and consequently privacy considerations. Finally, informing the users about the architecture underlying the face recognition system and give them the choice to go ahead with the process or not would provide them more sense of control, and consequently reduce their privacy concerns.

We think that this experiment would be really crucial for gathering information to make future data collection systems more efficient, user friendly, and most of all giving control to the users to manage their privacy from the very early stages of development.

Acknowledgments. This research is supported by an Australian Government Research Training Program Scholarship. We thank the University of Technology Sydney; ARC Discovery Project scheme; and CBA-UTS Social Robotics Partnership.

References

Amos, B.; Ludwiczuk, B.; and Satyanarayanan, M. 2016. Openface: A general-purpose face recognition library with mobile applications. Technical report, CMU-CS-16-118, CMU School of Computer Science.

Anthonysamy, P.; Greenwood, P.; and Rashid, A. 2013. Social networking privacy: Understanding the disconnect from policy to controls. *Computer* 46(6):60–67.

Basharat, I.; Azam, F.; and Muzaffar, A. W. 2012. Database security and encryption: A survey study. *International Journal of Computer Applications* 47(12).

Chen, M.; Mao, S.; and Liu, Y. 2014. Big data: a survey. *Mobile Networks and Applications* 19(2):171–209.

Fung, B.; Wang, K.; Chen, R.; and Yu, P. S. 2010. Privacy-preserving data publishing: A survey of recent developments. *ACM Computing Surveys (CSUR)* 42(4):14.

Kidd, C. D., and Breazeal, C. 2004. Effect of a robot on user perceptions. In *Intelligent Robots and Systems, 2004.(IROS 2004). Proceedings. 2004 IEEE/RSJ International Conference on*, volume 4, 3559–3564. IEEE.

Krizhevsky, A.; Sutskever, I.; and Hinton, G. E. 2012. Imagenet classification with deep convolutional neural networks. In *Advances in neural information processing systems*, 1097–1105.

Laugwitz, B.; Held, T.; and Schrepp, M. 2008. Construction and evaluation of a user experience questionnaire. In *Symposium of the Austrian HCI and Usability Engineering Group*, 63–76. Springer.

Matyáš, V., and Říha, Z. 2002. Biometric authentication-security and usability. In *Advanced Communications and Multimedia Security*. Springer. 227–239.

Milne, G. R. 2000. Privacy and ethical issues in database/interactive marketing and public policy: A research framework and overview of the special issue. *Journal of Public Policy & Marketing* 19(1):1–6.

Pearson, S. 2009. Taking account of privacy when designing cloud computing services. In *Proceedings of the 2009 ICSE Workshop on Software Engineering Challenges of Cloud Computing*, 44–52. IEEE Computer Society.

Turilli, M., and Floridi, L. 2009. The ethics of information transparency. *Ethics and Information Technology* 11(2):105–112.

Turner, C. W.; Zavod, M.; and Yurcik, W. 2001. Factors that affect the perception of security and privacy of ecommerce web sites. In *Fourth International Conference on Electronic Commerce Research, Dallas TX*, 628–636. Citeseer.

Wayman, J.; Jain, A.; Maltoni, D.; and Maio, D. 2005. *An introduction to biometric authentication systems*. Springer.

Williams, M.-A. 2009. Privacy management, the law & business strategies: A case for privacy driven design. In *International Conference on Computational Science and Engineering, 2009*, volume 3, 60–67. IEEE.