Government Information Quarterly xxx (2016) xxx–xxx

Contents lists available at ScienceDirect

Government Information Quarterly

journal homepage: www.elsevier.com/locate/govinf

On design and deployment of two privacy-preserving procedures for judicial-data dissemination

Mortaza S. Bargh [a,*], Sunil Choenni [a,b], Ronald Meijer [a]

[a] Research and Documentation Centre, Ministry of Justice and Security, The Hague, The Netherlands
[b] Creating 010, Rotterdam University of Applied Sciences, Rotterdam, The Netherlands

ARTICLE INFO

Article history:
Received 28 March 2015
Received in revised form 8 May 2016
Accepted 5 June 2016
Available online xxxx

Key words:
Action design research
Feedback
Information sharing
Privacy
Transparency

ABSTRACT

Institutions such as governmental and scientific organizations share information to gain the public trust. Information sharing, on the other hand, may cause privacy breaches and undermine the stakeholders' trust in such information sharing institutions. Thus, information sharing may work against the purpose of gaining trust through transparency. Moreover, fear of potential privacy breaches compels information disseminators to share minimum or no information. In this contribution, we present two procedures – the so-called restricted access procedure and open access procedure – to disseminate information for the contending purposes of transparency and privacy preservation. These procedures enable sharing of data with data requesters directly or via a trusted third party, respectively, in the context of our public judiciary organization. We have developed and operationalized these design artifacts in an organizational context and the resulting procedures have emerged from operational interactions within our organization. As such, our inquiry of knowledge thereto can be considered as an action design research. This contribution describes our approach and reflects upon our practice inspired research, where we share the gained insights and present some design guidelines – like providing usage control through implicit and explicit feedback, sharing data with scientists and for scientific purposes, and adopting a pre-commitment strategy – for the information systems that aim at sharing information in a real setting and in a privacy preserving way.

1. Introduction

Governments and scientific community seek for more openness and transparency through sharing information/data[1] with their stakeholders such as citizens, scientists, organizations and enterprises. The main motivation of these institutions for creating transparency is to gain the trust of stakeholders in their institutions (Kulk & van Loenen, 2012; Rajamäki et al., 2012; ROB, 2012; Zuiderwijk et al., 2012). Information sharing, on the other hand, may increase the chance of compromising privacy-sensitive data such as individuals' names, email and postal addresses, dates of birth, geo-locations, bank account numbers, photos, and political/personal opinions. When data sharing by institutions results in privacy risks and possibly harms citizens and individuals, the trust in these institutions may diminish.

Information systems that process, e.g., collect, enhance, and store, privacy-sensitive information are fairly vulnerable to information leakage, as we have witnessed many times in recent years (Bargh et al., 2012). This information leakage, which can lead to privacy breaches, stems from, for example, cyber attacks, compromised information systems, or (un)intentional disclosure of privacy-sensitive information through fusing of various information sources. Although there are numerous technologies for protecting privacy-sensitive information – such as data anonymization, pseudonymization and perturbation (O'Hara, 2011) – it is possible to derive sensitive personal data from the protected data when one combines various available data (Kulk & van Loenen, 2012; Narayanan & Shmatikov, 2008; Sweeney, 2000; van den Braak, Choenni, Meijer, & Zuiderwijk, 2012). When privacy breaches occur, not only is the trust in information sharing institutions on stake, there are also enormous costs inflicted on individuals, businesses and the society at large. Individuals can face, for example, emotional embarrassments, loss of employment/business opportunities, increased health and life insurance fees, and identity theft. For organizations and businesses there are direct costs such as legislative fines, shareholder lawsuits, third party and customer compensations, profit loss, and legal defense costs; indirect costs such as those for upgrading and maintaining of protective systems and safeguards; and implicit

* Corresponding author at: Research and Documentation Centre, Ministry of Justice and Security, WODC, Fl. 14, Turfmarkt 147, 2511 DP The Hague, The Netherlands.
E-mail addresses: m.shoae.bargh@minvenj.nl (M. S. Bargh), r.choenni@minvenj.nl, r.choenni@hr.nl (S. Choenni), r.f.meijer@minvenj.nl (R. Meijer).

[1] Throughout this contribution the terms 'information' and 'data' are used interchangeably.

http://dx.doi.org/10.1016/j.giq.2016.06.002
0740-624X/

costs such as those associated with reputation and branding damages, loss of goodwill, reduced turnover, and damaged customer loyalty. Privacy breaches impact also the society at large due to diminishing the collective trust of people in online services, upon which the foundation of our current networked society rests.

It is challenging to design and realize robust privacy preserving systems (Choenni, van Dijk, & Leeuw, 2010; Tschantz & Wing, 2009). This challenge originates from the contending driving forces that system designers have to simultaneously consider. For example, designers should deal with the preferences and wishes of users; limitations of technologies, constraints of laws and regulations; ill intention and malevolence of adversaries; and unforeseen side-effects of information fusion and analytics. User privacy preferences, moreover, are subjective and dependent of the context, e.g., location, time and situation. In addition, there exists a gap between what users want and how a system is realized due to lack of communication and cooperation between technical-oriented people, business-oriented people and (end) users (Choenni, van Waart, & de Haan, 2011).

In this paper we focus on preserving privacy in design, realization, and intervention of those information systems that process privacy-sensitive information and share (part of) it with other parties. Our objective is to accommodate two contending properties of transparency and privacy preservation that relate directly and (possibly) indirectly to information sharing, respectively. The objective of information dissemination, in turn, is to enhance the trust of the public and individuals in governmental organizations. The key contributions of this paper lie in: (a) elucidating how we accommodated privacy and transparency in sharing judicial information, and (b) deriving the design principles contributed to the successful reconciliation of privacy and transparency in this judicial information sharing setting. The paper reports on the design and intervention of two organizational procedures to enable sharing of judicial information in a structured way. As such, our work represents a class of problems/solutions and delivers a set of design principles, which are derived from our research and practice in a time span of nine years. The work can be characterized as distinctive due to its reality in being applied to real judicial data of citizens, its longevity in being operational for nine years, its continuity in having overlapping phases, its flexibility in adapting to the changing environmental/contextual conditions, and its complexity in making hard decisions to accommodate privacy and transparency within the setting.

For this study we have carried out an Action Design Research (ADR) within the Research and Documentation Center (abbreviated as WODC in Dutch) of the Dutch Ministry of Security and Justice for sharing our judicial information. This contribution explains our approach for deriving privacy related requirements and design insights during the development/operation of the two procedures. The resulting design principles of our study – like providing feedback to data controllers about how their data is used, adopting a pre-commitment strategy to share data, learning about the context of data usage as well as data access, addressing the fear of data fusion with other sources – can be adopted by similar data dissemination initiatives in non-judicial settings.

The next section outlines the research setting and methodology. Section 3 provides some background information and the related work. Subsequently, Section 4 presents our design and intervention activities and results. Section 5 discusses the attained results in realizing and operationalizing the design artifacts. Finally, Section 6 presents our conclusions.

2. Research method

For the work presented in this contribution a design approach is adopted where the objective was to seek out for non-zero-sum solutions (O'Hara, 2011) in regard to the contending objectives of transparency and privacy preservation. To this end, the design should not only acknowledge the differences between these objectives, but also consider the differences among people and their values and relations in the context of use, and also fulfill the requirements of WODC's strategy namely: feasibility, sustainability, research-oriented and demand driven constraints, as explained in Section 3.1. The work, as a result, has become an enquiry of knowledge through design, realization and intervention of "IT [Information Technology] artifacts in an organizational context and learning from the intervention while addressing a problematic situation" (Sein et al., 2011). The resulting artifacts of restricted access and open access procedures emerged from the interactions between the design and the intervention of the artifacts within the organizational context.

The process of design, development, and intervention of these two procedures can be categorized as Action Design Research (ADR) (Sein et al., 2011) in the sense that: (a) the same individuals fulfilled the roles of researcher and practitioner because it was impossible to discern their mutual influence, and (b) the resulting artifacts emerged from interactions within the organizational context. Being shaped by the organizational context during their development and use stages, the outcomes have turned into real artifacts in a preferred state in comparison to the initial state in 2005. Note that the ADR activities for the open access procedure are described in details in (Zuiderwijk et al., 2014) with a focus on those generic – thus not privacy specific – activities in 2012 and 2013. This contribution, however, covers the ADR activities related to *privacy preservation aspects* of *both* artifacts since 2005.

3. Background and foundations

This section serves as the "problem formulation" (Sein et al., 2011) stage of the ADR method. We shall describe the problem background in Section 3.1, the theoretical foundation of the study in Section 3.2, which corresponds to the "practice inspired research" and "theory ingrained artifacts" principles of the ADR method (Sein et al., 2011), and the related work in Section 3.3.

3.1. Setting and motivations

The study has been carried out at WODC, which is the research center of the Dutch Ministry of Security and Justice. WODC systematically collects, stores and enhances the Dutch judicial information directly or indirectly through its external partner organizations. The resulting criminal-justice information is used to define the future research agenda, to answer the policy-related questions, and to assess the possible implications of standing policies of the ministry. In order to accomplish its mission, WODC has adopted the strategy of *being demand-driven* in fulfilling the demands of the stakeholders within the ministry, *being research-oriented* in addressing practical challenges and producing scientific outcomes, *having feasible objectives* in fulfilling those demands that are achievable given the resources and expertise available at WODC, and *doing sustainable activities* in reusing the knowledge of executed projects in future projects. The rationale behind our strategy is to serve our stakeholders and increase WODC's viability through identifying new challenges and acquiring new knowledge. The work presented in this contribution, therefore, was intended to generate the knowledge that can be applied to similar problem classes and not just to solve the problems at hand.

WODC works with two types of data: judicial registration data, which is the raw judicial data that can be extracted from a number of government databases, and judicial research data, which is the enriched data that WODC or its partners have produced in various research projects. In addition, WODC produces and possesses the so-called statistical information, which can be regarded as non-confidential aggregated data. This statistical information is produced based on the judicial registration data and the judicial research data. In summary, WODC's data can be categorized, with a decreasing order of privacy-sensitivity, as: judicial registration data, judicial research data, and statistical information.

Please cite this article as: S. Bargh, M., et al., On design and deployment of two privacy-preserving procedures for judicial-data dissemination, *Government Information Quarterly* (2016), http://dx.doi.org/10.1016/j.giq.2016.06.002