The ethical dilemma of implicit vs explicit data collection: Examining the factors that influence the voluntary disclosure of information by consumers to commercial organizations

Christos Themistocleous
Andrew Smith
Marketing Division
Business school, University of Nottingham
Nottingham, United Kingdom
lixct3@nottingham.ac.uk,
Andrew.p.smith@nottingham.ac.uk

Christian Wagner
Senior Member, IEEE
Horizon Digital Economy Research
Intelligent Modelling and Analysis Group
School of Computer Science
University of Nottingham
Nottingham, United Kingdom
Christian.wagner@nottingham.ac.uk

Abstract— This paper focuses on the examination, analysis and ethical evaluation of the explicit data collection method employed by commercial organizations. It builds upon the work of previous academics and introduces a multi-dimensional framework which provides an anatomy of what influences the voluntary disclosure of information by consumers to commercial organisations. Each of the framework’s dimensions analyses constructs that derive from different academic disciplines like psychology, consumer behavior and marketing allowing the mixture and cross examination of multi-disciplinary factors that influence voluntary disclosure. The presented dimensions of the framework are a) Psychological processes b) Relational factors and c) Instrumental factors. Additionally, an experimental approach is proposed in order to examine and analyse the framework. This approach allows the identification of the most efficient and explicit combination of construct conditions leading to more effective data collection methods for organizations while taking into consideration consumer concerns regarding their privacy.

I. INTRODUCTION

The collection of customer information presents one of the main challenges that organizations have to face in order to acquire a competitive edge. The evolution of marketing in the last decades showed that consumer information is vital for organizations. Academics [1], [2], argue that the correct use of consumer information enables organizations to build customer loyalty. Additionally acquired customer information enables organisations to expand their customer base through customised advertising and more personalised products [3]. This generates what was identified [4] as personalisation-privacy trade-off where consumers sacrifice an amount of their private information in return for more personalised products that match almost perfectly their demands.

Even though transactional data are obtained by organizations more easily when compared to non-transactional data some argue [5],[6] that non-transactional data is extremely important for organizations due to the insightful information it provides in terms of consumer lifestyle, shopping habits as well as psychological characteristics. What was identified as a drawback by some organizations is the fact that non-transactional data rely (or should rely) on the voluntary disclosure by individuals [7]. Vital is also the method in which consumer data is acquired. Explicit data collection is made by explicitly asking consumers to disclose private data while the implicit data collection is characterised as a “behind the curtain” approach [7] where information is being gathered without the direct involvement or consent of the individual. The implicit data collection method is considered to be the easiest of the two options which leads to the main explicit vs implicit data collection dilemma with which organisations are faced. Ethical issues arise because of the potential non-voluntary disclosure of non-transactional data that disregards the fundamental right of privacy of the individuals.

Important is also the fact that in recent years several data protection initiatives like Mydata in the UK are slowly becoming more dominant in the data protection area. Data collection initiatives are poised to provide consumers with access to their personal data that are held by marketers while alleviating information exchange asymmetry [8]. An asymmetric exchange occurs when consumers receive limited value for providing information to firms [9]. This arguably embraces the importance of explicit data collection and the need for minimising the amounts of data that are being collected in an implicit fashion.
In the literature the factors that influence disclosure of information emerge from different disciplines like psychology, consumer behavior, and marketing. From the marketing literature, in recent experiments conducted by Acquisti and Loewenstein [26] it was found that the order in which privacy related questions are presented to individuals influences disclosure. Additionally, findings indicate that when individuals are led to believe that others have disclosed similar information they tend to follow a herding behavior. Furthermore, in another experiment by Zimmer et al [25] they examined how organizations voluntary influence disclosure of information by integrating privacy related questionnaires with statements regarding the use of the acquired data by the organization. From these experiments it can be argued that the way in which privacy related questionnaires are designed influences voluntary disclosure of information.

From consumer psychology, Barnett [10] introduced a framework which takes into consideration positive and negative stimuli that influence disclosure of information. One of these stimuli was coined as relational depth and examined how the relations between individuals and organizations influence discloses. Even though the framework embraces the importance of this construct it defines it as the sum of three widely examined concepts; trust, intimacy and relationships failing to provide a clear context for each. The main limitation of this is that the vagueness of the relational depth parameter does not allow the smooth employment of the framework in an organizational environment. This is due to the fact that not all industries allow the development of intimacy between organizations and customers mainly because there is lack of one to one interactions. Additionally Branett’s framework fails to provide practical guidelines as to what kind of trust organizations need to aim for in order to influence voluntary disclosure.

This paper proposes a new framework which incorporated Barnett’s disclosure management framework and amplifies the concept of impersonal trust from psychology in order to mitigate its main limitation. Additionally it integrates it with the design of data-capturing questionnaires in order to make the accumulation of information more practical, effective and efficient. Apart from the organisational perspective, this framework takes into consideration the consumer perspective by putting voluntary disclosure together with consumer knowledge and consent for use of information, as its foundations.

The structure of this paper is as follows: Section II introduces the proposed framework explaining also how it is integrated to the literature that led to its development. Section III provides a description of future research, more specifically the experiment that shall be used in order to examine each of the framework’s dimensions. Finally section IV provides an overview of this study together with potential academic and managerial contributions.

II. FRAMEWORK-BACKGROUND

The proposed framework serves as the integration of key parameters that influence disclosure of information with the organizational environment, while taking into consideration ethical issues in order to protect consumers. These parameters emerge from different disciplines like psychology, consumer behavior, and marketing each of which is included in one of the framework’s three dimensions which are entitled as:

- A. Psychological Processes
- B. Relational Factors
- C. Instrumental Factors

These are further described in the subsections below. Fig. 1 summarises all three dimensions of the framework together with their constructs.

A. Psychological processes

The psychological processes dimension includes the key concept of Disclosure Management by Barnett [10] which is the cognitive process that consumers go through before disclosing (or not) certain information. This concept is based on social exchange and prospect theory and it offers a simplified and comprehensive understanding of what influences the decision making of consumers when it comes to divulging information. Additionally both sides of the consumer’s mental scale are being examined, through the Perceived Disclosure Consequences (PDCs) and the benefits offered from organizations that seek to compensate individuals for their voluntary disclosure of information. These lead to the development of the concept’s four main constructs which are (i) Loss of face [11] and (ii) Loss of privacy [12,13] on the one side of the mental scale and on the other (iii) Compensation for disclosure of information [14,15] and (iv) Relational Depth [16]. All four constructs are incorporated in our framework with relational depth analysed as a dimension on its own for reasons explained later on.

![Fig.1 Visual Summatisation of the three dimensions that consist the framework together with the constructs of each.](image)

Note: The relational depth construct is analysed as a dimension on its own

(i) The loss of face has been identified as the potential embarrassment individuals encounter when disclosing sensitive information and it serves as a negative stimulus in the disclosure management process [10],[11]. As an addition the element of familiarity and relational depth that the
projector of sensitive information has with the recipient can result to greater embarrassment and thus to greater loss of face. [11]

(ii) Loss of privacy also serves as a negative stimulus to this cognitive process. A reasonable argument would be that the type of information being disclosed has a direct influence to the amount of privacy being lost [7],[12],[17]. From literature the distinction between sensitive and non-sensitive information can be traced back in more than a century ago [18] followed by more recent papers [1], [12], [13] that defined even more the two notions while providing examples of each. Sensitivity of information is defined by “the potential loss associated with the disclosure of that information” while sensitive information is linked to psychological, physical and material losses from disclosure [13:pp.77].

(iii) The offering of compensation that leads to disclosure of information is a widely explored concept [14],[15],[19]. Compensation serves as a positive stimulus in the disclosure management process since its presence in a potential exchange of information can alleviate concerns of potential losses that individuals might have from disclosure of information. Interestingly the type of compensation whether it is monetary or non-monetary does have an influence on disclosure of information while it is supported that monetary compensation serves as a better stimulus for disclosure of information [14].

(iv) From the literature, three main relational concepts that affect disclosure of information in the organisational level have been identified. These are relationships [16],[27] intimacy [28],[29] and trust [20,30,31]. The relational factors dimension of the proposed framework focuses on the last one and it is presented in the following section. It has been acknowledged that trust is an extremely wide and broad concept which in most cases “result(s) in a confusing potpourri of definitions applied to a host of units and levels of analysis” [20 pp625]. The examination of trust as a dimension on its own allows the better definition of the context in which it shall be examined in order to alleviate the broadness of the term.

B. Relational factors

Trust is identified as the “social relationship in which principals invest resources, authority, or responsibility in another to act on their behalf for some uncertain future return” [20pp.626]. On an interpersonal level, trust is defined as “having clear, mutually agreed upon expectations”[12:pp6]. Additionally the importance of trust is embraced by the fact that in most cases consumers are dependent and rely on strangers in order to act on their behalf. This is the case for disclosure of information by consumers to organizations where individuals trust organisations to use their information as intended and without manipulations while protecting it from misconduct [21]. When there are fewer opportunities for interpersonal trust to be developed impersonal trust can take place.

A relatively underdeveloped concept, impersonal trust is identified as the trust towards the impersonal organizational factors from vision and strategy to top management, fair processes and structures [22]. Impersonal trust is vital especially “(...) when faceless and readily interchangeable individual or organizational agents exercise considerable delegated power or privilege on behalf of customers who can neither specify, scrutinize, evaluate, nor constrain their behaviour” [22; pp113]. The importance of impersonal trust can be reflected in organisations where the addressed audience from which the organization seeks to extract information has limited scope for the development of interpersonal relationships. More precisely this can be attributed to the lack of strong Customer Relationship Management practises by the organisation as well as the lack of promotion for one to one interactions with consumers by the industry itself. Impersonal trust is mainly influenced by the following three parameters:

a) Fair Information Practices (FIPs): FIPs have a major impact on the development of impersonal trust [21]. FIPs are being identified as the statements and actions of organizations that seek to protect the acquired data from misconduct, as well as the empowerment of consumers by providing access to their information and the option of information withdrawal [4]. This reinforces the view that statements regarding privacy made by the organization affect positively the view that consumers have towards the organization’s credibility which is considered vital for the development of trust [23].

b) Perceived risks: Impersonal trust is also influenced by the perceived risks from disclosure of information. These perceived risks are linked with lack of customer knowledge regarding the reasons behind the collection of information as well as the processes that are to follow after the organisations acquires that information [1], [24]. Perceived risks can be reduced with the explicit explanation of the use and purpose of the acquired information [23] which is also expressed as the knowledge consumers have regarding the use of their information by the organisation who acquires it [12], [21].

c) Individuals who had good prior experiences with an organisation accompanied by the presence of consumer knowledge regarding that organisation’s FIPs enables the better development of impersonal trust [21].

C. Instrumental factors

This dimension explores how consumers perceive different methods of data capture in terms of privacy-intrusiveness which can be translated broadly into the following: How the structure of data capturing questionnaires influences disclosure. Interviews are purposely excluded due to the obvious impracticalities and inefficiencies of interviewing large numbers of individuals in order to acquire private information. Three key concepts are included in this dimension:

1) The concept of dyadic relationship which focuses on how organisations that disclose information to consumers prior to asking for information can positively influence disclosure of information.
2) Comparative nature: A concept which is interested on whether individuals compare themselves with others in terms of disclosure of information.

3) Question sequence: This concept examines whether the order in which questions related to privacy are presented in a questionnaire influence disclosure of information. All three concepts are analysed below.

1) Dyadic relationship: This concept follows a qui pro quo approach where the organisation offers certain information regarding the use of the acquired information from individuals in order to induce voluntary disclosure. The three types of dyadic relationship are a) Reasoned dyadic, b) Unreasoned dyadic and c) Non-dyadic relationship [25]. The information provided in the reasoned dyadic relationship reflects directly the potential uses of the requested information. The unreasoned condition provides information that is indirectly associated with the question that it is to follow. The non-dyadic relationship simply provides information of what it is to follow in the next question of the questionnaire.

2) Comparative nature: As mentioned comparative nature is based on the examination of a herding behaviour between individuals when it comes to voluntary disclosure of information. Acquity and Lowestine [26] indicate that a herding behaviour is in place when individuals are led to believe that other individuals have disclosed certain similar sensitive information. This is translated as an alleviation of the initial privacy concerns of the individual when is informed that other individuals did provide that information.

3) Question sequence: This concept derives from the examination of the order of the questions included in a data-capturing questionnaire. Question sequence is identified to have three conditions: ascending order of intrusiveness in the order of questions, descending order and random order[26]. In the descending sequence condition the most intrusive questions are presented first followed by less intrusive ones while in the ascending sequence condition intrusive questions are presented towards the end. In the random sequence questions were presented in no particular (random) order.

Apart from the proposed framework, this paper also introduces an experimental approach which will allow a suitable and comprehensive examination of the framework. The following section analyses that will be used for analysing the framework.

III. FUTURE WORK- AN EXPERIMENTAL APPROACH FOR EXAMINING THE FRAMEWORK

When examining the proposed framework the main objective is to capture how consumers perceive each of the constructs that compose the framework’s three dimensions. In doing so an experimental, questionnaire-based approach will be followed. The experiment will include a series of different questionnaires that are to be distributed to UK adult citizens.

Additionally individuals will be presented with a series of privacy-related questions that cover areas from medical conditions to sexual preferences and criminal activity. These thirty questions were pre-tested generating an ascending and descending order in terms of privacy invasiveness. In the experiment all subjects that took part in the pre-test will be excluded.

The objective of the experiment will be to present these pre-tested questions and induce individuals to engage and interact with them in order to understand how each of these is related to their privacy. This interaction process will include asking individuals to provide answers to these questions or choose to avoid disclosure. This experiment will be mainly interested on the asked questions that would follow the interaction with the pre-tested questions. The questions following this shall be based on the concepts of loss of privacy, loss of face and compensation all of which are part of the psychological processes dimension.

A. Conditions

In order to maximise the quality of the acquired responses, combinations of the different conditions of constructs will be made. In particular, the three conditions of dyadic relationship (reasoned, unreasoned, non) together with the three of the question sequence (ascending, descending, random) and the three of comparative nature (low, high, neutral) all of which consist the instrumental factors dimension, shall be included in a 3X3X3 matrix resulting to the generation of 27 different conditions. Each of these generated conditions will be reflected through the order, structure and information accompanying each of the thirty pre-tested questions with which respondents will interact with. This will provide a comprehensive understanding of which of the generated conditions maximise disclosure of information in a context in which the respondent feels comfortable and secure.

B. Questionnaire Structure

In the first section of the questionnaire the thirty pre-tested questions shall be presented in one of the 27 conditions accompanied by questions regarding the psychological processes dimension. Interesting is the fact that the compensation construct allows an assessment of the efficiency of each condition since in some cases where loss of face and loss of privacy is reduced, the compensation required in return for disclosure of information might be high. Responses are measured with the use of a seven-point Likert scale and the items used for each of the three constructs of psychological processes were in accordance to the literature [10], [11], [14].

The second section will cover the concept of impersonal trust which composes the third dimension- relational factors. As mentioned the relational factors dimension is an extended examination of the relational depth concept of the psychological processes dimension. The examination of impersonal trust will be based on a scenario followed by fourteen questions regarding its three constructs [23] (FIPs, Perceived risks from disclosure, prior experiences). In order to
enhance the impersonal element in the scenario, a limited amount of information regarding an anonymous data-capturing organisation shall be provided together with information regarding the current relationship that the respondent has with it. Measurement of these constructs will be based on ten items with the use of a seven-point Likert scale [23].

IV. CONCLUSION-CONTRIBUTION

This paper is concerned with the explicit data collection, more precisely the factors that influence voluntary disclosure of information by consumers to commercial organisations and seeks to pinpoint, examine, analyse and evaluate key parameters that influence it. In order to do so this paper introduces a multi-dimensional framework in which key parameters related to voluntary disclosure are introduced accompanied by an experimental approach which will allow its examination.

The experimental approach that will be employed for investigating the constructs of the framework allows the cross examination of different concepts from different disciplines regarding voluntary disclosure allowing also the identification of superior combinations of the different conditions of these concepts that lead to the maximisation of the efficiency of explicit data collection methods. Additionally from an academic standpoint the development of a framework like the one proposed here will allow the verification of correlations between variables as identified by the literature as well as the potential identification of new correlations regarding voluntary disclosure.

From a managerial perspective, this project will help organizations to understand and appreciate what factors influence voluntary disclosure of information by consumers. The research design not only allows the identification of effective data collection methods but also enables an assessment of the efficiency of these through the examination of the compensation construct. Furthermore by providing some insight on what influences disclosure of information will enable organizations to strategically formulate appropriate customer friendly strategies that will help them acquire much needed non-transactional data in an efficient and effective manner.

REFERENCES


