Analysis of Factors that Influence Customers’ Willingness to Leave Big Data Digital Footprints on Social Media: A Systematic Review of Literature

Abstract:
Big data has been discussed extensively in existing scholarly works but scant consideration is given to customers’ willingness to generate and leave big data digital footprints on social media, especially in the light of the profusely debated issue of privacy and security. The current paper endeavours to address this gap in the literature by developing a conceptual framework. In doing so, this paper conducts a systematic review of extant literature from 2002 to 2017 to identify and analyse the underlying factors that influence customers’ willingness to leave digital footprints on social media. The findings of this review reveal that personal behaviour (intrinsic psychological dispositions), technological factors (relative advantage and convenience), social influence (social interaction, social ties and social support) and privacy and security (risk, control and trust) are the key factors that influence customers’ willingness to generate and leave big data digital footprints on social media. The conceptual framework presented in this paper advances the scholarship of technology adoption and use and provides useful direction for future empirical research for both academics and practitioners.

Key Words: Big Data Digital Footprint; Social Media; Privacy and Security, Technology, Personal Behaviour, Social Influence

Unified Theory of Acceptance and Use of Technology model (UTAUT2), which takes into account various aspects of customers’ use of technology to offer deep insights into the dynamics and kinetics of customers’ willingness to deposit digital DNA on social media
Introduction

Big data digital footprints are digital DNA that customers generate and leave on digital platforms when they interact with and use various media channels, including social media. The ever-increasing use of and interaction with social media has intensified researchers’ and practitioners’ interest in customers’ social media led digital footprints (Rauniar et al., 2013; Tuton and Solomon, 2015). However, research on customers’ perceptions of and attitude towards generating big social data digital footprints is still in its infancy.

Digital footprints are described as social data created by customers when they interact with media channels. Such digital footprints are not just identities but also memories, moments and behaviour. Social media providers who collect these huge digital chronicles can determine how and why users behave and purchase on digital platforms (Fish, 2009). Social media use has grown exponentially and has become an integral part of consumer life. With the advent of Web 2.0, digital footprint generation has increased significantly. It is estimated that 44 times more data generation would take place, from 2009 to 44 zettabytes of data by 2020 (CSC 2017). In addition, the exponential growth of mobile telephony (Sharma, 2017), cloud computing and 4G networks have created many more social media touchpoints. As a result, customers are found to be connected to smart devices (smartphones, tablets, smartwatches, Cortana, Siri and Alexa etc.) 24/7, generating and leaving behind huge digital trails for service providers. Furthermore, user-generated content on social media will be the main channel for the enrichment of information base for public administrative bodies and commercial firms (Baur, 2017).

Tuton and Solomon (2015) divided the use of social media into four zones based on different social media channels and vehicles: i.e. social community interaction, social commerce, social publishing and social entertainment. Customers create their digital DNA on channels in each of these zones by sharing comments, photos, videos, blogs, bookmarks, reviews, ratings and social shopping, linking with government applications etc. (Malhotra et al., 2012; Rosenberger et al., 2017). These digital trails exhibit their interests, social and cultural identities, and occupational and geographical attachments, which are essentially required by firms (Charlesworth, 2014; Michael et al., 2014). Moreover, these digital traces help firms to analyse customers’ sentiments and contents by using advanced analytics to gain deeper insight into their behaviour and develop their profiles (Charlesworth, 2014; DWork and Mulligan, 2013). Customers use social media excessively, but they may or may not be aware of the digital footprints that they leave for companies such as Google, Yahoo, Amazon and Facebook. The services of social media platforms (Facebook, Twitter and Instagram etc.) have redefined the ways in which business value can be generated, as these providers and tech giants use personal data to track customers and their behaviour through invasive and ubiquitous crawling. They use algorithms to generate powerful insight through data connections, inferences and data interpretations (DWork and Mulligan, 2013). Furthermore, as the competition amongst firms is increasing to seek innovative capabilities to mine digital trails and gain a competitive edge over rivals. Hence, the managerial implications for big data digital footprints are immense, as they can create value and promote development (Pulse, 2012).

On the other hand, the pervasive use of digital footprints has raised privacy and security concerns amongst social media users. This area is of huge interest for individuals, public and
governments as to where to draw the line for privacy and security of unauthorised access to individual digital footprints. This has even led many countries to initiate measures to protect individual privacy and security as social media platforms and web technologies have become more pervasive and vulnerable to abuse and exploitation. Recently in the USA, privacy advocates allege that large Internet service providers can potentially encroach to consumers’ privacy as they have access to large volume of personal data (Waters and Bond, 2017). Even everyday objects, connected to the Internet, are collecting personal digital footprints (Kuchler, 2017).

Although consumers’ engagement with social media has received significant research attention (Al-Jabri et al., 2015; Charlesworth, 2014; Hajli, 2014; Hsu and Wu, 2011; Akar and Topçu 2011; Hau and Kim, 2010), there is paucity of research that identifies and analyses the factors that influence customers’ intention to generate and leave big data digital footprints on social media. Hence, further investigation will not only contribute to big data and social media literature, but will also advance privacy and security scholarship. Our study addresses this research gap by identifying and analysing the factors that determine customers’ willingness to generate and leave digital DNAs on social media and how privacy and security deter or facilitate digital footprint generation on social media.

Furthermore, the current literature provides the theoretical framework on customers use and adoption of social media (Hsu and Wu, 2011; Lin and Anol, 2008; Lu et al., 2009; Venkatesh et al., 2012; Wei et al., 2009), for instance, Technology Acceptance Model (TAM), Theory of Reasoned Action (TRA), Theory of Planned Behaviour (TPB), Uses and Gratifications (U&G) and the Unified Theory of Acceptance and Use of Technology (UTAUT1 and UTAUT2) etc. are widely cited theories and frameworks for assessing customers’ use of social media. These theories take into account factors that influence users’ acceptance and use of social media such as usefulness, relative advantage, ease of use, innovation, social and compatibility issues (Al-Gahtani et al., 2007; Davis, 1989; Davis et al., 1989; Hsu and Wu, 2011; Lin and Anol, 2008; Lu et al., 2009; Venkatesh et al., 2003; Venkatesh et al., 2012; Wei et al., 2009). However they do not fully capture the dynamics and kinetics of customers’ willingness to generate and leave big data digital footprints on social media and provide limited scope for generalisation. Furthermore, they do not address how privacy and security affect customers’ willingness to generate digital footprint on social media. The current paper aims to develop appropriate conceptual scaffodings for customers’ willingness to generate and leave big data digital footprints on social media by critically examining and synthesising the above mentioned theories and concepts that have been extensively cited and applied in information systems, management and marketing literature.

Therefore, we feel there is a need for a conceptual paper to develop a theoretical framework through a rigorous and systematic review of existing academic literature (Cropanzano, 2009). A conceptual paper will offer theoretical impetus for future scholarly works and a theoretical framework for empirical investigation. Hence, the current paper aims to present a conceptual framework that defines the nascent and potential inter-relationships amongst various constructs.

The paper is structured as follows: The first section describes the methodology of the study. The next section delineates the findings and analysis. The final section discusses the results and their theoretical and practical implications, along with limitations and future research direction.
Methodology of the Systematic Review

In order to study customers’ willingness to generate and leave digital footprints on social media, a desk-based study was carried out based on the approach of Ngai et al. (2009) and Ngai and Wat (2002) as follows.

Classification Process

The following steps were carried out in the systematic review.

1. Database search
2. First classification
3. Verification of first classification
4. Final verification

The literature around the use of social media and big data digital footprint was found in a number of different journals belonging to a wide range of disciplinary origins. The following online journal databases were used to identify published peer reviewed articles: EBSCO Host, Emerald, IEEE/IET, Electronic Library, ProQuest, Sage, Science Direct and Scopus.

The following criteria were used to screen and classify the articles. Only peer reviewed articles were chosen to represent advanced research output as suggested by Ngai et al (2009). Articles that have main themes around the use of social media, big data digital footprint, customers’ digital footprint, factors influencing the use of social media were chosen. In order to carry out a thorough desk based systematic review, the following appropriate search terms were identified after detailed discussions amongst the authors based on the literature review, following Nill and Schibrowsky (2007). Furthermore, they were searched individually through the search algorithm on the chosen databases. They were rephrased and cross checked to ensure articles fall into the area of the research as the extent literature suggests. “Customers’ digital footprint on social media”, “Why customers leave digital footprint on social media”, “Big data digital footprint of customers on social media”, “Customers’ use of social media”, “Why customers generate digital footprint on social media”. Based on the chosen approach, only peer-reviewed journal articles were selected. The period from January 2002 to May 2017 was selected considering the topicality of the issue.

The literature review resulted initially in the identification of 506 articles. The full text of each article was reviewed by the first and second authors, confirmed through discussion and verified by the third author. Feedback from two other independent academics was collected to ensure the validity and reliability of the process. A classification framework, shown in Figure 1, was developed based on the key emerging factors from the systematic review along with their sub-factors. These four key emerging themes were found to be the leading factors determining customers’ use of social media and their digital footprint generation. As a result, based on these four key factors, 58 articles were selected from the initial sample of 501 articles, although this selection was not exhaustive.

Figure 1: Classification Framework
Data Reporting

A six-column table (Table 1) was designed to report data captured from the final 58 articles. It was used to examine the constructs used by the previous studies to determine customers’ use and willingness to generate and leave big data digital footprints on social media. The following table delineates the details of constructs identified in each study, along with the choice of theories, research approach, design and findings respectively.

TABLE-1

Findings and Analysis

The systematic review of scholarly works led to the development of the above classification framework (Fig. 1) and identification of the key factors. These factors were found to be the leading determinants of the use of social media and customers’ digital footprint generation. Further analyses of each of the factor along with some sub-factors are provided below.

Personal behaviour

Personal behaviour consists of personal intrinsic and extrinsic motivation (sub-factors of personal behaviour are listed in Figure 2). It has been found that the use of social media is mostly influenced by intrinsic motivation such as intrinsic perceived functional and emotional benefits; self-enhancement, self-esteem and ego (Diffley et al., 2011; Hau and Kim, 2011). Users are driven by their perceived intrinsic experiential benefits, sensory pleasure (hedonic and emotional: Park and Kim, 2014) and self-enhancement, which satisfy their hedonic needs and extrinsic benefits. Similar findings are noted by Whiting and Williams (2013), who report that users tend to engage with social media to fulfil their intrinsic psychological needs of entertainment, relaxation and expression of opinions. The extant literature suggests that psychological intrinsic emotional factors of enjoyment, pleasure, and self-enhancement affect social media user behaviour (Al-Jabri et al., 2015; Campbell et al., 2014; Chen et al., 2014; Grace et al., 2015; Kim et al., 2011; Pereira et al., 2014).

Enjoyment comprises pleasure and flow (optimal psychological experience) and pleasure in the form of playfulness, fun and an intrinsic acceptance of social media (Hsu and Wu, 2011; Kim et al., 2011; Wu and Chang, 2005; Zolkepli and Kamarulzaman, 2015). Self-enhancement consists of self-esteem (self-status and image), which enables users to gratify their outcome expectations of personal influences on social media. High self-enhancers have high self-esteem, due to which they overwhelmingly update and present their self-image/status on social media and attract attention (Chen et al., 2014; Hennig-Thurau et al., 2004; Hepper et al., 2011; Hsu et al., 2007; Presi et al., 2014).

As such, it is found that users’ engagement with social media is based primarily on personal intrinsic motivational factors that affect their behaviour in the use of social media.
Technology

Technological factors are found to comprise performance expectancy (usefulness, utility), ease of use and relative advantage of the use of social media. Technological sub-factors are given in Figure 3. Performance expectancy is found to comprise perceived usefulness and utility of technology, which influences users’ behavioural intentions. Similarly, ease of use includes convenience of the use of technology and relative advantage involves the innovative compatibility of social media over other technologies (Al-Gahtani et al., 2007; Lin and Anol, 2008; Venkatesh et al., 2003) and the prevailing adoption of technology (Carter and Weerakkody; 2008). It can be summarised that perceived usefulness, ease of use, relative advantage and innovation significantly influence social media users’ behavioural intentions (Chiang, 2013; Hajli, 2014; Milewicz and Saxby, 2013; Pillai and Mukherjee, 2011; Zhang et al., 2014). Hence, technological factors which constitute social media usefulness, ease of use and compatibility, influence the intention to use of social media (Idemudia et al. 2016; Gironda and Korgaonkar, 2014; Lu et al., 2010; Wang et al., 2012; Zolkepli and Kamarulzaman, 2015).

Social factors

Social factors include social interaction, social ties and social support (Bharati et al., 2014; Chiasson and Lovato, 2011; Grace et al., 2015; Talukder and Quazi, 2011; Venkatesh et al., 2003). These factors drive social interaction, which is a desire to connect, collaborate and communicate with others on social media (Chang and Chuang, 2011; De Valck et al., 2009; Hussain, 2012; Trivedi et al., 2106). Social Sub-factors are given in Figure 4.

Social interaction can be described as the desire to communicate, interact with others and build relationships on social media (Al-Jabri et al., 2015; Ko et al., 2005). Similarly, social media are perceived to enhance social interaction, connect people almost anywhere, give control over interaction and maintain social relations with others (family, friends etc.). It is a platform to release anxiety and depression, and to increase companionship and interpersonal utility, as suggested by the extant literature (Ellison et al., 2007; Grieve et al. 2013; Oldmadow et al., 2013; Park et al., 2009; Whiting and Williams, 2013). Moreover, it is found that customers’ pleasurable experience and peer pressure enhance social interaction on social media (Grace et al., 2015; Junglas et al., 2013).

Social ties denote building and maintaining relationships with other social media users (Rishika et al., 2013; Wang et al 2012). They are psychological goals of users to develop and maintain social relations with others, release anxiety and enhance interpersonal utility (Ellison et al., 2007; Grieve et al. 2013; Whiting and Williams, 2013), as it is human nature to socialise and interact with others (Dyson, 1998).

Similarly, social support includes willingness to help others and share anything that would assist others on social media (Liang et al., 2011; Zhang et al., 2014). It is a social aspect of exchange to share information with others (Ali, 2011; Crocker and Canevello, 2008) and also a major social value for social media users from other community members that use and
interact with social media (Obst and Stafurik, 2010). As such, it can be summarised from the above findings that social factors (social interaction, social ties and social support) enhance the use of social media (Dalla Pozza, 2014; Foster et al., 2011; Gironda and Korgaonkar, 2014; Hsiao et al., 2010; Hsu and Wu, 2011).

**Figure 4: Social Factors**

**Privacy and Security**

Privacy and Security are found to comprise perceived risk, control and trust. Privacy involves users’ willingness to share information online and the ability to control and choose to divulge personal information, whereas security relates to protection against the threat from the unauthorised access to personal information on social media (Belanger et al., 2002; Eastlick et al., 2006). Moreover, privacy also includes individuals’ location, communication and information privacy (accumulation, treating and sharing information: Dinev et al., 2013). Sub-factors of privacy and security are given in Figure 5.

Privacy is described as users’ perceived sense of risk regarding information about oneself and security is perceived as protection against the threat from unauthorised access to information about someone (Boyd, 2008; Lee et al. 2013). By joining and interacting with social media, users create their profiles, connect and share interests and personal information with others, which may potentially lead to personal privacy and security risks (Cheung et al., 2015; Tan et al., 2012).

Likewise, in existing literature trust has also emerged as one of the leading factors for privacy and security, and comprises users’ confidence in the ability of a service provider to protect and monitor their personal information or reduce their uncertainty about the use of the service, as they also pay considerable heed to providers’ (vendor and social media) integrity and benevolence (Cao et al., 2015; Cheung et al, 2015; Cheung and Lee, 2006; Krasnova et al., 2010). Trust is also noted to have a key relevance to social media users’ perceived risk and it is developed through quality and source credibility (Burgess et al., 2011; Chen and Sharma, 2013; Gamboa and Gonçalves, 2014; Pentina et al., 2013; Shin, 2010; Zhou et al., 2016). From the above discussion, it is found that privacy and security can affect the use of social media and digital footprint generation.
Theoretical Foundation

A diverse theoretical constructs have been used and applied in the existing literature to conceptualise the above issue. TAM, TRA, TPB and U and G theories have been most frequently used in examining customers’ use and adoption of social media applications. More information has been provided in the Table-1.

Research Approaches

Positivist methods involving quantitative approaches and statistical testing appear to be more popular and common compared to interpretivist approach and/or qualitative methods. Quantitative research being a dominant method in this scholarship has used closed ended surveys. For model testing and data analysis, Structural Equation Modelling using AMOS and LISREL is a popular tool. For qualitative research, focus groups and semi-structured interviews were found to be the most widely used methods. Table 2 shows that 43 of the 58 studies chose quantitative approaches, whereas 10 studies used qualitative approaches and only 5 studies chose mixed methods respectively.

TABLE -2

Table 3 shows the distribution of journals, the number of articles from each journal and the percentage breakdown.

TABLE-3

Conceptual Framework

The current paper presents a conceptual framework (Figure 6) based on the relevant constructs for the antecedents of customers’ willingness to generate and leave big data digital footprints on social media. The model is based on the Unified Theory of Acceptance and Use of Technology (UTAUT 2). UTAUT 2 is used because it takes into account various aspects of consumers’ use of technology, such as motivation, innovation, technology utilisation and social aspects, as highlighted by the findings of this research. UTAUT 2 integrates elements on the use of technology from UTAUT 1 with similar themes (Venkatesh et al., 2012; Dwivedi et al., 2017). Hence, the proposed conceptual framework is developed by expanding on the UTAUT 2 model to examine the underlying factors that lead customers to generate and leave big data digital footprints on social media. The proposed framework has three antecedents of customers’ behavioural intention (technological factors, social influence and personal behaviour), with privacy and security as moderator. From the findings delineated above, privacy and security are hypothesised to have moderating influence (facilitate or deter) on the relationship between customers’ behavioural intention and its antecedents.
Figure 6 presents the proposed framework and the following section explicates each component of the research model.

**Figure 6: Conceptual Framework**

**Antecedent of customers’ behavioural intention**

UTAUT 2 adopted from UTAUT 1 which unified the constructs for technology adoption in order to develop a more holistic understanding. The model suggests performance expectancy by combining similar themes such as perceived usefulness, utility and relative advantage. Likewise, perceived ease of use and complexity are brought within effort expectancy; social norms, social factors and social image are combined for social influence; perceived behavioural control and compatibility are termed as facilitating conditions. UTAUT2 extends the model by integrating enjoyment within hedonic motivation, cost within price value and habit as determinants of customers’ behavioural intention (Venkatesh et al., 2012). Some of the earlier works, such as Dwivedi et al. (2008) also ascertained some of these factors in their study of technology adoption, acceptance and diffusion indicating the validity of the model. The current paper has taken a similar stance and identified the following three antecedents of customers’ behavioural intention in terms of their willingness to generate and leave digital footprints on social media. Details of each factor are given below. These factors, as delineated above, affect customers’ digital footprint generation.

**Technological factors:** The systematic review of literature reveals that the technological factors in the form of usefulness, ease of use (convenience) and relative advantage can influence customers’ interaction with social media and subsequent generation of digital footprints. Moreover, relative advantage comprises perceived technological innovation, which embodies convenience and compatibility, affecting customers’ intention to use social media. Likewise, perceived usefulness was found to enhance job performance and to be instrumental in achieving valued outcomes, as suggested by the extant literature, such as Chen et al. (2009), Chiang (2013), Ho and Wu (2011), Lee et al. (2011) and Lin (2011). Hence, this study makes contribution by explicating the key technological factors of convenience and relative advantage that have a major influence on customers’ big data digital footprint generation.

**Social influence:** This study has found that social influence is based on customers’ perceived psychological social pressure, comprising social interaction, social ties and social support. This study contributes by unearthing the key social factors of social interaction, social ties and social support that hypothesised and tested to have influence on customers’ psychological needs (Liang et al., 2011; Zhang et al., 2014), which result in big social data digital footprint generation on social media.

**Personal behaviour:** This study reveals two key personal behaviour factors of self-enhancement (self-efficacy and self-esteem) and perceived experiential and sensory pleasure benefits (hedonic and emotional; joy and enjoyment) that satisfy customers’ hedonic needs (Hau and Kim, 2010; Park and Kim, 2014). Hence the paper makes a key contribution to the current scholarship by highlighting the key personal behavioural factors that determine customers’ behavioural intention in generating big data digital footprints on social media.
Privacy and security: This study reveals that privacy and security comprise of perceived risks, control and trust. It has found that customers’ self-perceived sense of risks in leaving personal information on social media and their sense of trust in social media providers (privacy) have a huge impact on their use of social media and their digital footprint generation. Equally, their perceived control and trust in service providers’ ability and giving unauthorised access to others (security) affects their digital footprint generation on social media. This study, therefore, makes a contribution by revealing that customers’ sense of privacy is enhanced when their perceived sense of risk increase and their perceived information control and perceived trust in social media decrease, which accordingly facilitates or deters digital footprint generation on social media.

Theoretical contributions

First of all, this research advances the literature on the factors that lead to customers’ willingness to generate and leave big data digital footprints on social media. In doing so, it has extracted and analysed the key antecedents influencing customers’ behavioural intentions in the form of technological factors, social influence and personal behaviour. It has also developed a model which exhibits the apposite and significant association among factors, providing valuable insights in determining customers’ underlying behaviour in depositing big data digital footprints on social media, which can have practical implications for managers and practitioners. Furthermore, this study provides significant implications for the role of privacy and security for service providers (social media and vendors). Hence, the paper offers strong conceptual underpinning for assessing the dynamic and dichotomous nature of users’ social media engagement. While, the model is suggested for social media based interaction, the privacy and security issue can also be applied for other technological applications such as cloud computing and smartphones.

Conclusion

This desk-based study has applied a systematic literature review of 501 peer-reviewed articles. The articles were scattered across 38 different journals, which substantiates that the area of research is of interest in many disciplines. However, marketing, international business and information systems are more in the limelight. This study has found that personal behaviour, technology, social influence and privacy and security are the key factors that influence customers’ willingness to generate and leave big data digital footprints on social media. It has also developed a conceptual framework that offers deeper insights into the dynamics and kinetics of customers’ social media use and their willingness to deposit digital DNA on these media. In addition, this study highlights the theories used profusely in previous scholarly works and articulates the research gap. Addressing the research gap not only advances and enriches the widely used technology adoption theories but can also facilitate future empirical works. A more holistic and robust understanding of consumers’ ambivalence regarding the use of social media and their willingness (or a lack of it) of leaving digital footprints would enlighten future researchers and pave the way to more scholarly works in this field.

Nevertheless, the paper provides useful insights to the practitioners who can use the model to ascertain their target customers’ motivation for and perceptions of using social media and
other related interfaces. Business should realise that customers do not want to be chased; rather they would like to be wooed. Not all customers would have same motivation and/or similar level of desire to engage with social media. Hence, the social media providers such as Facebook and Instagram and the businesses who promote their products and services through these platforms and collect customer information should be aware of the sensitivities and intricacies pertaining to customers’ privacy and security.

**Limitations and Future Research Directions**

There are some notable limitations of this study. Firstly, it has focused mainly on peer-reviewed academic articles; future research could include monographs and industry reports. Secondly, the articles were clustered around the chosen four dimensions; future research could combine and choose the articles which have common factors. Thirdly, this study includes some articles around online social commerce; future research could focus just on social media customers. Last but not the very least, the conceptual framework offers a more generic model for assessing customers’ willingness to leave digital footprints. Future research could include cultural factors to posit the model in specific socio-cultural contexts.
References


Fish, T. (2009). My Digital Footprint A two-sided digital business model where your privacy will be someone else's business!: futuretext.


Hau, Y.S. and Kim, Y.G. (2010). “Why would online gamers share their innovation-conducive knowledge in the online game user community?


Kuchler, H. (2017, April 4). The Internet of things: Home is where the hackers are. *Financial Times*. Retrieved from https://www.ft.com/content/cb880bc2-057c-11e7-ace0-1ce02ef0def9


