Research Paper

JIS

Journal of Information Science
1–15

DOI: 10.1177/0165551521992756
journals.sagepub.com/home/jis
SAGE

Delphi study of risk to individuals who disclose personal information online

David Haynes
School of Computing at Edinburgh Napier University, Edinburgh, UK; City, University of London, London, UK

Lyn Robinson
City, University of London, London, UK

Abstract
A two-round Delphi study was conducted to explore priorities for addressing online risk to individuals. A corpus of literature was created based on 69 peer-reviewed articles about privacy risk and the privacy calculus published between 2014 and 2019. A cluster analysis of the resulting text-base using Pearson's correlation coefficient resulted in seven broad topics. After two rounds of the Delphi survey with experts in information security and information literacy, the following topics were identified as priorities for further investigation: personalisation versus privacy, responsibility for privacy on social networks, measuring privacy risk, and perceptions of powerlessness and the resulting apathy. The Delphi approach provided clear conclusions about research topics and has potential as a tool for prioritising future research areas.

Keywords
Delphi study; online risk; personal data; privacy

1. Introduction

1.1. Objectives

The objectives of this study were twofold: (1) to examine statements about online privacy risk from the research literature and (2) to identify areas where there is no consensus and which could be prioritised for further research.

The scope of the study is to identify the risks to individuals associated with disclosure of personal information during online activities. These disclosures may be made consciously or may be the result of data leakage, such as that which occurs when cookies and other tracking technology is used by service providers.

1.2. Background

The study was part of a 2-year investigation into the nature of risk in a privacy calculus and was prompted by the perception that there needs to be a way of categorising personal risks online [1]. These findings could be applied in several areas, such as public safety, informing policy making and feeding into a better understanding of the potential for personal insurance in cybersecurity. There is an emerging market in cyber insurance to protect corporate bodies against online risks. It is possible to imagine that individuals may one day seek protection against the exposure to harm that results from online activity. This research will help to define that market.

A privacy calculus is a way of conceptualising individuals' motivations for disclosing personal information online. There are various ways in which this can be estimated. Based on interviews with subjects, several hypotheses are tested to establish the relationship between perceived risks and benefits and willingness to disclose personal information online [2]. This has been extended by some researchers to test behaviour in controlled environments [3,4]. The 'privacy

Corresponding author:
David Haynes, School of Computing at Edinburgh Napier University, Merchiston Campus, Edinburgh EH10 5DT, UK.
Email: d.haynes@napier.ac.uk

paradox' has been described as 'the observed dissonance between consumer concerns and hypothetical behavior and actual decision making' or 'the dichotomy of information privacy attitude and actual information privacy behaviour' [5,6]. Acquisti, Brandimarte and Loewenstein take a slightly broader view of the 'dichotomy, between people's self-reported mental states (attitudes, concerns, desires, etc.) regarding privacy and their actual behaviors' [7].

Much of this research centres on definitions of risk that all have elements of probability and negative impact on outcomes or activities that matter to the individual [6,8]. However, the concept of risk is problematical because it is treated in different ways in the literature and it is not always clear whether the same concept is being discussed. For example, in project management terms, risks can represent threats or opportunities [9]. When it comes to industrial safety, however, risks are exclusively about negative consequences [10].

A risk event or incident may itself be a consequence of an earlier event. Each risk event may result in several different consequences. This makes the relationship between risk events and their consequences complicated and potentially ambiguous. An event tree analysis may be applied to map a cascade of consequences (and events) from an initial event. The distribution of outcomes may lend itself to a probability tree to which different analytical methods such as a Monte Carlo simulation or Bayesian analysis could be applied [11].

In the absence of quantitative data, a qualitative approach was adopted in the knowledge that a qualitative approach cannot provide a probability of occurrence nor quantify the impact of a risk event. This could eventually be addressed by obtaining transactional data for a statistically significant sample of users representing a target population and analysing the types of risks that individuals face. Where statistical data are available, it may be possible to track the consequences of those events. For example, it may be possible to extract quantitative data from:

1. Statistical authorities (e.g. Office for National Statistics);
2. Previously gathered datasets;
3. Transactional data from an online service provider or website;
4. Reports in the press or literature about incidents of harm [12];
5. Reports of harm (consequences) mentioned in online forums.

Another approach is to ask experts in the appropriate field to provide estimates of the probable levels of future risk and harm that might occur. With the appropriate choice of experts and maintenance of anonymity, it should be possible to detect whether there is a consensus or at least stable views about the probable levels of risk of different risk scenarios. For instance, van Duersen et al. gathered data from 117 health authorities' incident registers from 2005 to 2010 [13]. They were able to identify 2018 separate incidents and on analysis they found that there were approximately 150 scenario types. They then recruited a panel of 12 experts with more than 5 years' experience of cybersecurity, data protection and information management in the health service in the United Kingdom and asked them to identify the scenarios that were likely to occur in the future. After three rounds of this Delphi-type consultation, the researchers were able to report on six scenarios that would most probably affect health trusts.

Paintsil evaluated a taxonomy of privacy and security risks using the Delphi technique [14]. The Delphi study verified the key risk indicators (KRIs) proposed by the researcher as well as identifying some possible new ones. In contrast, Di Gangi et al. used the literature to identify the types of harm or risk that could occur during use of online social media by employees [12]. Their focus was on risks to organisations rather than individuals. They also used a Delphi-type approach with a panel of experts to evaluate the current state of social media policy development.

Chen used cluster analysis for literature reviews to detect and visualise emerging trends in the literature [15]. The idea of clustering is based on detected similarities between items in a collection using measures such as modality, betweenness centrality and silhouettes [16,17]. This research carries forward the idea of using recently published literature as the source material for a Delphi study rather than attempting to evaluate a taxonomy of privacy and security risks. This allows access to a wider range of perspectives and reduces the risk of being aligned to a single, idiosyncratic view of online privacy and security.

2. Methodology

The Delphi technique was developed as a forecasting technique by the Rand Corporation for the US Air Force in the 1950s [18]. The original purpose was to try and predict what the Soviet Union's strategy might be for an A-bomb attack to disrupt US munitions production. One of the concerns was the lack of data and the degree of uncertainty involved in this type of planning. A Delphi survey was designed to get a range of views independently from a panel of experts and to ascertain whether there was any consensus on possible answers to the question.

Since that time, Delphi studies have evolved to become an established technique for forecasting [19]. In addition to providing a structured approach to gathering expert opinions, it avoids 'group think' by consulting the experts independently. It depends on two or more rounds to narrow the field of investigation to those areas where there is no consensus, and arguably requiring the greatest attention. Ideally, there should be several rounds of opinion seeking until there is stability – that is, no significant change in the distribution of opinions from one round to the next. Lund suggests that library and information science researchers could use Delphi studies to develop frameworks, taxonomies and forecasts [20].

A literature search was conducted using the following terms in the titles of articles in the Serial Solutions consolidated bibliographic databases (including Emerald Insight, Library and Information Science and Technology Abstracts, SCOPUS and Web of Science). These databases were chosen for their broad coverage of literature of library and information science, computing and sociology:

'privacy calculus' OR 'privacy paradox' 62 articles
'privacy' AND 'risk' 136 articles

The results were limited to scholarly articles published in the 5 years prior to the search (from August 2014 onwards). A 5-year limit was used because this is a rapidly evolving area of research. Recent publications will reflect relevant findings from earlier research. These searches yielded 170 items after duplicates had been removed from the combined set. This was filtered down to 89 documents following a review of the titles and abstracts to eliminate items that did not address risk to individuals. Articles on medical genomics, privacy issues affecting children and privacy related to the Internet of things were also removed. Although these are important areas, they are beyond the scope of this study. A further 20 items were eliminated after reading the full text of all candidate articles, because they were not specifically focused on risk to individuals.

A corpus of 69 full-text items (Appendix 1) was created for content analysis using NVivo12, a qualitative data analysis software tool. Word frequency counts were used to identify potential search terms that could form the basis of topics for the study. This was followed by a cluster analysis based on word similarity in the full text of journal articles, using the Pearson correlation coefficient. The articles in each cluster were examined to identify their common themes and this formed the basis for labelling each cluster. The resulting categories were then checked for overlap and in some cases were merged into a new category, which was labelled accordingly. The text extracts were coded according to the topics identified from the word frequencies and the cluster analysis. This allowed text extracts about a particular topic to be listed together for comparison. The individual text extracts were consolidated and normalised into a set of statements, for which respondents could indicate their level of agreement or disagreement on a 5-point Likert-type scale. The resulting statements (Table 2) were tested in a pilot run with a small number of researchers before the final instrument was created for the study. The following categories were identified from the corpus:

- Mobile technology;
- Images;
- Tracking and advertising;
- Social media and social networks;
- Privacy harms;
- Health apps;
- E-government;
- E-commerce and mobile payments;
- Personalisation;
- Location-based services;
- Risk awareness and perception;
- Privacy paradox.

A panel of nine experts was recruited from the cybersecurity sector, the social informatics research community and the library and information profession (Table 1). The experts were identified in consultation with committee members of the British Computer Society's (BCS) Information Security Special Group and the Information Literacy Group of CILIP, the UK library and information association. Of the nine experts who completed the first round, seven completed the second round of the study.

In the first round, the panel was asked to rate the level of agreement or disagreement with statements about information privacy and online risk. They were also asked to comment on the choice of topics and to add any that they felt were

Table 1. Expertise of the Delphi panel.

Risk assessment and management	2
Privacy and data protection	6
Cyber security	2
Social media and online social networks	4
Other	1 each
Digital inclusion	
Information and digital literacies	
Information literacy	
Copyright and online learning	
Resilience and crisis management	

missing. A 5-point Likert-type scale was used to rate the level of agreement with each statement. The results were analysed and areas of consensus were identified.

In the second round, the modified list of non-consensus topics was listed and the seven remaining experts were asked to review their scores considering the previous results and to give reasons if their scores differed significantly from the median score. They were also informed about the topics where consensus was reached but were not allowed to vote on them a second time. The topics were again analysed to identify any further consensus topics and to extract opinions where there was no consensus.

Von der Gracht's review of Delphi methodologies points out the consensus measures such as mean and standard deviation, which are commonly used for continuous scales or ratios, are not appropriate for Likert-type scale surveys [19]. Mode and interquartile ranges (IQRs) are more relevant here, with an IQR < 1 indicating consensus. An advantage of this approach is that outliers do not unduly affect the average score or the dispersal of scores.

Throughout this process, the experts were not made known to one another, so that the main interaction was through the Delphi study.

There were nine respondents for the first round and, of these, seven respondents completed the second round of the survey. Although the recent trend has been towards larger panels, early studies such as the initial one conducted by the Rand Corporation had a panel of seven experts and there is no indication that smaller panel size undermines the validity of results of the Delphi studies [19,20]. Table 1 shows the expertise of the panel members (Note that some individuals belonged to more than one category).

All the respondents were considered experts in their respective specialisms, as evidenced by their membership of relevant professional bodies. Three respondents declared that they had 5–9 years' experience, three had 10–14 years' experience and three had 15 or more years' experience in their specialism.

Not surprisingly, the membership distribution reflected the recruitment from CILIP and BCS members. The following professional memberships were declared by the nine respondents (bearing in mind that some individuals belonged to multiple professional groups):

UK chartered bodies
CILIP (Library and Information Association) (3)
BCS (British Computer Society) (2)
CIISec (Chartered Institute of Information Security) (1)
Other professional groups (1 each)
ISACA (Information Systems Audit and Control Association)
CiSP (Cybersecurity Information Sharing Partnership)
Higher Education Academy
Society for Research into Higher Education
Security Institute
IAPP (International Association of Privacy Professionals)
Data and Marketing Association
CILIP Information Literacy Group (non-CILIP member)

3. Results

The responses to rounds 1 and 2 of the Delphi study were analysed separately to identify areas where there was some degree of consensus between the respondents. The respondents were encouraged to give their immediate views on a range

of statements (listed in Table 2). The 5-point Likert-type scale was used, with values ranging from 1 – strongly agree to 5 – strongly disagree.

Answers to these scales were analysed arithmetically to determine the degree of divergence between items. The IQR was defined as the interval of values within which 50% of the responses were recorded. If this interval was 1 or less than 1 unit either side of the mode value, this was taken as an indication of consensus. In other words, if 50% of the scores were within 1 unit either side of the modal score, this was classed as consensus [19].

Table 2 lists the statements and results of the two rounds of the Delphi consultation, the implications of which are explored in the 'Discussion' section. The statements were extracted from the articles in each cluster and consolidated where the statements were similar in meaning.

There was no consensus for the following statements after two rounds:

PERSONALISATION
To operate effectively, intelligent user interfaces need to acquire rich information about the user.
The intrusiveness of personalised ads outweighs the benefits of personalisation.
SOCIAL MEDIA AND ONLINE SOCIAL NETWORKS
It is possible to learn a great deal about someone from what his or her connections have said about them on social media.
The social need for sharing information online outweighs potential privacy risks.
Keeping their information private is primarily the personal responsibility of social media users.
RISK ASSESSMENT
Social media connectedness is a good indication of a user's privacy risk.
USER BEHAVIOUR AND MOTIVATION
The experience of being online highlights feelings of powerlessness.

4. Discussion

After two rounds of the Delphi survey, a strong level of consensus was achieved about most of the statements presented to the panel. This may reflect the small number of participants who took part, although this was considered in setting the parameters for consensus. Consensus was defined as occurring where the IQR of responses was within one unit of the modal value. This meant, for instance, that all the statements about risk perception, privacy risks and protection and mitigation were consensual after two rounds (Table 2).

4.1. Personalisation

Personalisation was the first area considered by the expert panel, where there were two statements where no consensus was achieved.

The first is the idea about how much personal information is required by user interfaces to operate effectively. Hazard and Singh acknowledge that 'The most valuable of such information is potentially sensitive and revealing; it can pose a threat to the user's safety, finances or dignity … ' [21]. They go on to discuss the ways in which personalised services can undermine identity and sense of autonomy.

There was also disagreement about the balance of benefits and costs of personalisation. Karwatzki et al. showed that individuals' disposition to value privacy (DTVP) and situational factors affected their intention to disclose information to personalise services [22]. Their survey revealed that personalisation motivates information disclosure. However, those people who value privacy tend not to want to be profiled online. Their third finding was 'that transparency features do not directly impact the information disclosure of individuals'. In contrast, Brinson and Eastin found that 'advertising personalization did not have a significant effect on attitude toward the ad, but inclusion of the AdChoices icon did' [23].

This comes back to a discussion about risk: 'How much risk is perceived is further moderately predicted by the user's trust in the recipient's ability to protect his/her data, the degree of personalization that is gained by data disclosure and the perceived relevance of the collected information' [24]. Gerber et al. only saw a weak relationship between degree of personalisation and prediction of perceived risk.

Gutierrez et al. in reviewing previous research suggest that personalisation is seen as a benefit in mobile, location-based advertising [25]. Their survey of 252 respondents showed that 'The overall indications are that mobile users' MLBA acceptance is guided primarily by judgments on the perceived balance between intrusiveness and monetary rewards, and secondarily by their perceptions of personalisation or IPC' [25, p 300]. 'The effect of monetary rewards was found to be a more significant driver for MLBA acceptance than personalisation. This implies that, while users

Table 2. Results of consensus analysis of Delphi statements.

Statement	Mean of responses (R1)*	Mean of responses (R2)*	% Mode (R1)	% Mode (R2)	IQR (R1)	IQR (R2)	Relevant IQR	Consensus?
PERSONALISATION								
Higher ad relevance reduces anxieties about privacy	4.3		56		1.0		1.0	Y
To operate effectively, intelligent user interfaces need to acquire rich information about the user	3.3	3.3	33	33	2.0	2.5	2.5	N
The intrusiveness of personalised ads outweighs the benefits of personalisation	1.9	2.1	44	44	2.0	2.5	2.5	N
Personalised services increase the benefits of sharing information with mobile app service providers	2.9	3.3	44	22	1.0	2.0	1.0	Y
The quality of personalisation offerings outweighs privacy concerns	4.2		56		1.0		1.0	Y
A personalised service had no direct impact on whether or not users would like to use a specific application	3.2		44		1.0		1.0	Y
Individuals should determine whether location information is communicated to third parties	1.0		100	0	0.0		0.0	Y
Are there any other issues you would like to add under this heading? (These may be included in the second round of the Delphi survey)								
SOCIAL MEDIA AND ONLINE SOCIAL NETWORKS								
Privacy is a collective effort that requires the cooperation of those with whom we connect on social media	1.4		56		1.0		1.0	Y
In the social networking context, personal pleasure plays a much stronger role than utilitarian benefits	1.9	1.9	44	44	1.0	0.5	1.0	Y
It is possible to learn a great deal about someone from what his or her connections have said about them on social media	2.2	3.0	33	44	2.0	2.0	2.0	N
The main danger for social media users is a lack of balance between risk and reward	2.3	2.4	33	33	2.0	1.0	1.0	Y
The social need for sharing information online outweighs potential privacy risks	3.4	2.7	44	33	3.0	3.5	3.5	N
Keeping their information private is primarily the personal responsibility of social media users	3.3	3.0	33	22	2.0	2.0	2.0	N
Participants on social media understand privacy risks	4.3		56		1.0		1.0	Y
Are there any other issues you would like to add under this heading? (These may be included in the second round of the Delphi survey)								
RISK PERCEPTION								
Online risk perception does not have a strong influence on actual risk-avoiding behaviour	2.3		33		1.0		1.0	Y
Good company reputation reduces the perceived risk in disclosing personal information online	2.4	2.7	56	22	1.0	1.5	1.0	Y
A positive online brand image increases consumer-perceived benefits of information disclosure	1.8		78		0.0		0.0	Y
A good privacy policy reduces the perceived privacy risk	1.9		56		1.0		1.0	Y
Perceived enjoyment has replaced ease-of-use in accepting digital technologies	2.6		44		1.0		1.0	Y
Individuals' perception of control is a significant predictor of risk perception when consumers are providing their personal information	2.1	2.1	33	56	2.0	0.0	0.0	Y
Are there any other issues you would like to add under this heading? (These may be included in the second round of the Delphi survey)								
RISK ASSESSMENT								
Data protection impact assessments (DPIA) do not provide an effective solution to assess and prioritise privacy risks	3.1		44		1.0		1.0	Y

(continued)

Table 2. (continued)

Statement	Mean of responses (R1)*	Mean of responses (R2)*	% Mode (R1)	% Mode (R2)	IQR (R1)	IQR (R2)	Relevant IQR	Consensus?
PERSONALISATION								
Privacy risk assessments should consider threats to user privacy that come from user-profiling	1.3		67		1.0		1.0	Y
Social media connectedness is a good indication of a user's privacy risk	3.1	3.3	44	33	2.0	2.5	2.5	N
Privacy assessments influence individuals' actual behaviour	3.6	3.4	33	33	1.0	1.5	1.0	Y
Are there any other issues you would like to add under this heading? (These may be included in the second round of the Delphi survey)								
PRIVACY RISKS								
Self-disclosing of sensitive personal information poses a threat to safety	1.7	1.9	56	44	1.0	0.5	1.0	Y
Governments show little self-restraint in sacrificing citizens' rights to privacy in the face of real or perceived security threats	2.0	1.4	44	44	1.0	1.0	1.0	Y
Digital technology developments are threatening to overrun citizens' constitutional rights	2.0		78		0.0		0.0	Y
The privacy agenda is undermined by state surveillance	2.2	1.4	33	44	2.0	1.0	1.0	Y
Personal information leakage from apps on mobile devices is a critical concern	1.4		67		1.0		1.0	Y
Individuals trust businesses with access to their communications more than they trust government	2.2		78		0.0		0.0	Y
Are there any other issues you would like to add under this heading? (These may be included in the second round of the Delphi survey)								
USER BEHAVIOUR AND MOTIVATION								
Users often fail to enforce their own privacy preferences when judging image content	1.7		44		1.0		1.0	Y
The experience of being online highlights feelings of powerlessness	3.0	2.3	44	33	2.0	1.5	1.5	N
The increasingly complicated control of privacy settings leads users to believe any protection effort to be substantially useless	1.9		56		1.0		1.0	Y
Younger individuals are more aware about potential negative consequences of personal information disclosure in online environments	2.9	3.3	33	33	1.0	1.0	1.0	Y
Consumers are willing to pay a premium to purchase from privacy protective websites	2.9	3.3	44	22	1.0	1.5	1.0	Y
Are there any other issues you would like to add under this heading? (These may be included in the second round of the Delphi survey)								
PROTECTION AND MITIGATION								
Fake information as a privacy protection method should be avoided	2.3	2.6	33	44	2.0	1.0	1.0	Y
Without privacy measures, such as encryption, consumers won't feel safe enough to participate in e-commerce	2.1	2.0	44	44	1.0	0.5	1.0	Y
Fine-grained permissions (such as access for a specific file type) should be built into the design of apps	1.9	1.6	44	44	1.0	1.0	1.0	Y
Personalised warning notices significantly increase compliance compared to impersonal signs	2.1		67		0.0		0.0	Y
Online advertising based on personal online behaviour should be more tightly regulated	1.3		67		1.0		1.0	Y
Greater effort in strengthening consumer competency in technology use is required	1.3		67		1.0		1.0	Y

IQR: interquartile range; Y: yes; N: no.*

1: strongly agree; 5: strongly disagree.

welcome personalisation, monetary rewards have a higher compensating value for the perceived risk of accepting MLBA' (P 302).

Interestingly, Karwatzki et al, reported on a survey of 286 participants, which noted that 'we found that personalization itself motivates information disclosure' [22]. It would be interesting to investigate whether there is a difference in perceived benefits when individuals consciously (and deliberately) disclose personal data rather than having it automatically gathered by service providers.

Lee and Rha found that although there are benefits of personalisation, increased privacy risk inhibits take-up of these services [26]. They advocate more openness about the ways in which personal data are collected and used, allowing individuals to make their own assessments of risks involved.

In response to the two statements about the balance between privacy and the benefits of personalisation, the expert panel felt that privacy outweighed personalisation.

4.2. Social media and social networks

The panel did not agree about the following statements:

- It is possible to learn a great deal about someone from what his or her connections have said about them on social media;
- The social need for sharing information online outweighs potential privacy risks;
- Keeping their information private is primarily the personal responsibility of social media users.

Although it seems self-evident that it is possible to learn about someone from others' contributions to social media, the panel members did not have a consensus on the veracity of this statement. Hargittai and Marwick's report of interviews with young adults highlighted the case of the 21-year-old male who came from 'a family of oversharers' and that although he revealed little about himself online, it was still possible to learn a great deal about him from his family's posts on social media [27]. However, this situation may not apply in all cases, for instance, if the 'information' revealed is deliberately wrong or misleading.

This same article goes on to investigate the following research questions:

> 'RQ2: To what extent do participants feel that social needs for sharing information online outweigh potential privacy risks?
> RQ3: To what extent do participants believe that keeping their information private is their personal responsibility?' [27]

The authors explain the balance between risk and benefit in terms of networked privacy where 'individuals exist in social contexts where others can and do violate their privacy' [27]. This theme has been explored extensively in the literature on the privacy calculus from Dinev and Hart onwards [2].

RQ3 is interesting because it reflects the debate between individual responsibility and regulation of providers. Haynes, Bawden and Robinson suggest that a balance has to be struck between these modes of regulating access to personal data on social networks [28]. One element of this is user education, which requires a regulatory framework to provide a means of redress for privacy breaches. Responsibility for protecting privacy is the responsibility of all parties.

Chen's study in Hong Kong and the United States suggests a number of measures that individuals can take to protect their privacy by limiting: self-disclosure, friending and profile visibility [29]. The article suggests that with greater control over their profile visibility, individuals would more probably expand their social networks.

4.3. Risk assessment

There was no consensus about the following statement:

- Social media connectedness is a good indication of a user's privacy risk.

The analysis of social media networks by Alemany et al. resulted in the development of a Privacy Risk Score (PRS) [30]. They tested the performance of different topologies of synthetic networks before testing it on real networks of rumours. Global metrics such as closeness correlates with PRS, if there is information about the network structure. Social centrality metrics based on degree provides a good estimate of privacy risk [30]. Pensa et al. propose a privacy

score, partly inspired by Pagerank (also used by Alemany et al.), which provides an estimate of users' attitudes towards privacy [31].

4.4. User behaviour and motivation

The final area where consensus was not achieved was:

- The experience of being online highlights feelings of powerlessness.

This is a statement that needs qualification as it may not be universally applicable and this might be the reason for the lack of consensus. Hoffman et al. conducted a series of focus groups to explore feelings of vulnerability and powerlessness associated with being online [32]. They studied the attitudes of both low- and high-skilled individuals and found that there was privacy cynicism resulting from the belief in both groups that any efforts to protect personal privacy are futile.

4.5. Privacy harms

Another element of the survey was to obtain views about the areas of greatest risk to online users. Financial loss and cyberbullying were the two most highly ranked issues. There was then a cluster of mid-ranking issues: online stalking, intrusion by advertising and filtering of content. Then the final group was: disclosure of medical data, loss of control of personal data and algorithmic decision-making. This suggests that financial loss, cyberbullying and cyberstalking could be prioritised in further research into privacy harms.

4.6. Limitations

The application of the Delphi technique in this study could be improved in two ways: extending the range of experts and following up responses with interviews to understand the reasons for the experts' choices.

The survey could be run in parallel groups to see whether there is any consistency about the topics identified. The setting up of the expert panel may have affected the degree of consensus. Although all the experts were professionally involved in some aspect of online delivery of information, whether as information professionals or as cybersecurity experts, they had diverse backgrounds, as demonstrated by the range of professional bodies to which they belonged. The surveys could also run for more than two rounds to see whether the areas of disagreement stabilise. The question arises: 'To what extent is the distribution of responses down to chance?' Is there a normal distribution of responses? Are there differences if the cohort of experts is broken down by different criteria such as subject discipline or location?

Future Delphi studies in this area could also provide a fuller briefing with contextual information about the statements to be evaluated by the panel. Behind each of these statements was a published paper or papers based on empirical research. The expert panel was not provided with these papers and indeed was asked to give an immediate 'gut feel' answer rather than a considered view about each statement. Lack of consensus about the statements may be due to genuine differences of opinion. However, as has been pointed out by several respondents, it may be due to lack of understanding of the statement and ambiguities in the terminology used. Many of the statements were biased, deliberately so, to evoke a reaction and to reflect conclusions of researchers who contributed to the corpus of literature used in this research. Further work could be done to refine the statements, so that the intent is more obvious to panel members.

5. Conclusion

The objectives of this study were twofold. The first was to identify themes from the research literature about online privacy risk. A cluster analysis of recent research literature about online privacy risk (using Pearson's correlation coefficient) revealed several emerging themes:

- Personalisation – the benefits of personalisation of services outweighs privacy concerns;
- Social media and online social networks – the risks associated with the use of social media and how they are balanced with the benefits and who is responsible for maintaining the privacy of users;
- Risk perception – the effects that perceptions of risk have on online behaviour;
- Risk assessment – identification of risk assessment methods that can be applied to privacy risk;
- Privacy risks – identification of the main threats to personal privacy;

- User behaviour and motivation – methods individuals can take to preserve their privacy online;
- Protection and mitigation – ways in which privacy can be protected.

The second objective of the study was to examine the degree of consensus on specific issues. By creating clear statements, members of the expert panel were forced to make a choice about the degree to which they agreed with each statement. The delivery of the survey allowed for segregation so that panel members were not able to influence one another. This research was not a consensus-building exercise; rather it was intended to detect areas of consensus and disagreement using the IQR to identify where consensus had occurred.

Where there is no consensus, there is scope for further investigation, because it suggests that the theme is emerging and has not had time to become established and widely accepted. Alternatively, it might mean that there are distinct schools of thought where there is genuine disagreement. In either case, these are probably the productive areas for further research. Working on this principle, the following specific potential research questions have been identified from the literature and the Delphi study:

- Do intelligent user interfaces need to acquire rich information about users in order to be effective?
- How intrusive are personal ads and does that intrusiveness outweigh the benefits?
- How much information about individuals is revealed by the online activity of their associates (such as connections on social media)?
- Is there a way of detecting whether information about an individual revealed by a social media connection is misleading or incorrect?
- Who is primarily responsible for maintaining the privacy of social media users, the individual or the service provider?
- What is the relationship between connectedness and privacy risk? This could be based on the number of primary connections, the number of secondary and primary connections, the betweenness centrality of an individual within a network. It could be measured as the number of adverse incidents that an individual has been subjected to, for instance. A social media provider could also do an analysis by identifying individuals affected by specific incidents such as trolling, or spamming or fraud and then look at the characteristics of those individuals compared to a control group. Other factors such as public visibility of the individual would need to be taken into account.
- There has been some interesting research into feelings of powerlessness of individuals when it comes to online privacy [32–34]. It would be worth investigating social media users' attitudes to different forms of privacy regulation following Lessig's model, which has been adapted for social media by Haynes, Bawden and Robinson [28,35]. This is the idea that Internet regulation is not just about the law but is also brought into effect by the market, through self-regulation and by the way in which systems are designed.

5.1. Further work

The research has identified a number of areas where there was no consensus among the panel of experts. This study could be extended to a group of experts representing a wider range of backgrounds to see whether a similar pattern of disagreement persists (as described above). Having identified areas of disagreement, they could be explored in more detail by asking respondents why they held their views and what would persuade them to change their minds. A follow-on study with a modified set of statements taking into account comments from this study would be presented to a new panel of experts. Additional rounds would be followed up with interviews with panel members to understand the reasoning behind their responses.

The other follow-on for this research would be to further investigate the research questions that have arisen from the analysis of non-consensus areas.

Acknowledgements

The Delphi survey was conducted at City, University of London. The authors gratefully acknowledge the input of the members of the expert panel who freely gave their time and expertise for this study. The write-up was completed at Edinburgh Napier University where the corresponding author is a member of the Centre for Social Informatics. Thanks to colleagues at both institutions and to Brian Detlor at McMaster University for their helpful feedback during the drafting of this paper.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This research was supported by the Royal Academy of Engineering and the Office of the Chief Science Adviser for National Security under the UK Intelligence Community Postdoctoral Fellowship Programme (Grant No. ICRF1718\1\54).

ORCID iD

David Haynes https://orcid.org/0000-0001-9191-9247

References

[1] Haynes D. Creating an ontology of risk: a human-mediated process. In: Haynes D and Vernau J (eds) *The human position in an artificial world: creativity, ethics and AI in knowledge organization (ISKO UK Sixth Biennial Conference*, London, 15–16 July 2019). Baden-Baden: Ergon Verlag Gmbh, 2019, pp. 167–180.
[2] Dinev T and Hart P. An extended privacy calculus model for E-commerce transactions. *Inf Syst Res* 2006; 17: 61–80.
[3] Keith MJ, Thompson SC, Hale J et al. Information disclosure on mobile devices: re-examining privacy calculus with actual user behavior. *Int J Hum Comput Stud* 2013; 71: 1163–1173.
[4] Martin K. Diminished or just different? A factorial vignette study of privacy as a social contract. *J Bus Ethics* 2012; 111: 519–539.
[5] Gerber N, Gerber P and Volkamer M. Explaining the privacy paradox: a systematic review of literature investigating privacy attitude and behavior. *Comput Secur* 2018; 77: 226–261.
[6] Aven T and Renn O. On risk defined as an event where the outcome is uncertain. *J Risk Res* 2009; 12: 1–11.
[7] Acquisti A, Brandimarte L and Loewenstein G. Secrets and likes: the drive for privacy and the difficulty of achieving it in the digital age. *J Consum Psychol* 2020; 30: 736–758.
[8] Fischhoff B, Watson SR and Hope C. Defining risk. *Policy Sci* 1984; 17: 123–139.
[9] Murray-Webster R. *Management of Risk: guidance for practitioners*. 3rd ed. London: TSO, 2010.
[10] ISO/IEC 27005:2011. Information technology – Security techniques – Information security risk management.
[11] Garlick AR. *Estimating Risk?: a management approach*. Aldershot: Gower Publishing, 2007.
[12] Di Gangi PM, Johnston AC, Worrell JL et al. What could possibly go wrong? A multi-panel Delphi study of organizational social media risk. *Inf Syst Front* 2018; 20: 1097–1116.
[13] van Deursen N, Buchanan WJ and Duff A. Monitoring information security risks within health care. *Comput Secur* 2013; 37: 31–45.
[14] Paintsil E. Evaluation of privacy and security risks analysis construct for identity management systems. *IEEE Syst J* 2013; 7: 189–198.
[15] Chen C. CiteSpace II: detecting and visualizing emerging trends and transient patterns in scientific literature. *J Am Soc Inf Sci Technol* 2006; 57: 359–377.
[16] Hofstetter H, Dusseldorp E, Van Empelen P et al. A primer on the use of cluster analysis or factor analysis to assess co-occurrence of risk behaviors. *Prev Med* 2014; 67: 141.
[17] Aldenderfer MS. *Cluster analysis*. Beverly Hills, CA: SAGE, 1984.
[18] Linstone HA and Turoff M (eds). *The Delphi Method: techniques and applications*. Reading, MA: Addison-Wesley Publishing Co, 1975.
[19] von der Gracht HA. Consensus measurement in Delphi studies: review and implications for future quality assurance. *Technol Forecast Soc Change* 2012; 79: 1525–1536.
[20] Lund BD. Review of the Delphi method in library and information science research. *J Document* 2020; 76: 0418.
[21] Hazard CJ and Singh MP. Privacy risks in intelligent user interfaces. *IEEE Internet Comput* 2016; 20: 57–61.
[22] Karwatzki S, Dytynko O, Trenz M et al. Beyond the personalization-privacy paradox: privacy valuation, transparency features, and service personalization. *J Manag Inf Syst* 2017; 34: 369–400.
[23] Brinson NH and Eastin MS. Juxtaposing the persuasion knowledge model and privacy paradox: an experimental look at advertising personalization, public policy and public understanding. *Cyberpsychology* 2016; 10(1): 7.
[24] Gerber N, Gerber P and Volkamer M. Explaining the privacy paradox: a systematic review of literature investigating privacy attitude and behavior. *Comput Secur* 2018; 77: 226–261.
[25] Gutierrez A, O'Leary S, Rana NP et al. Using privacy calculus theory to explore entrepreneurial directions in mobile location-based advertising: identifying intrusiveness as the critical risk factor. *Comput Human Behav* 2019; 95: 295–306.

[26] Lee J-M and Rha J-Y. Personalization–privacy paradox and consumer conflict with the use of location-based mobile commerce. *Comput Human Behav* 2016; 63: 453–462.
[27] Hargittai E and Marwick A. 'What Can I Really Do?' Explaining the privacy paradox with online apathy. *Int J Commun* 2016; 10: 3737–3757.
[28] Haynes D, Bawden D and Robinson L. A regulatory model for personal data on social networking services in the UK. *Int J Inf Manage* 2016; 36: 872–882.
[29] Chen HT. Revisiting the privacy paradox on social media with an extended privacy calculus model: the effect of privacy concerns, privacy self-efficacy, and social capital on privacy management. *Am Behav Sci* 2018; 62: 1392–1412.
[30] Alemany J, del Val E, Alberola J et al. Estimation of privacy risk through centrality metrics. *Futur Gener Comput Syst* 2018; 82: 63–76.
[31] Pensa RG, Di Blasi G and Bioglio L. Network-aware privacy risk estimation in online social networks. *Soc Netw Anal Min* 2019; 9: 1–15.
[32] Hoffman CP, Lutz C and Ranzini G. Privacy cynicism: a new approach to the privacy paradox. *Cyberpsychology* 2016; 10(4): 7.
[33] Xie W, Fowler-Dawson A and Tvauri A. Revealing the relationship between rational fatalism and the online privacy paradox. *Behav Inf Technol* 2019; 38: 742–759.
[34] Solove DJ. A taxonomy of privacy. *Univ PA Law Rev* 2006; 154: 477–564.
[35] Lessig L. *Code*. 2nd ed. New York; London: Basic Books, http://codev2.cc/download + remix/Lessig-Codev2.pdf (2006, accessed 2 June 2015).

Appendix 1

Articles reviewed for the Delphi Study

Selection criteria (in Title): (privacy AND risk) OR 'privacy calculus' OR 'privacy paradox'

Date range: published 2014–2019

Aditya P, Bhattacharjee B, Druschel P, et al. Brave new world: privacy risks for mobile users. *ACM SIGMOBILE Mob Comput Commun Rev* 2015; 18: 49–54.

Adjerid I, Peer E, Acquisti A. Beyond the privacy paradox: objective versus relative risk in privacy decision making. *MIS Q* 2018; 42: 465.

Alemany J, del Val E, Alberola J, et al. Estimation of privacy risk through centrality metrics. *Futur Gener Comput Syst* 2018; 82: 63–76.

Bal G, Rannenberg K, Hong JI. Styx: Privacy risk communication for the Android smartphone platform based on apps' data-access behavior patterns. *Comput Secur* 2015; 53: 187–202.

Bhatia J, Breaux T. Empirical measurement of perceived privacy risk. *ACM Trans Comput Interact* 2018; 25: 1–47.

Brinson NH, Eastin MS. Juxtaposing the persuasion knowledge model and privacy paradox: an experimental look at advertising personalization, public policy and public understanding. *Cyberpsychology*; 10. Epub ahead of print 2016. DOI: 10.5817/CP2016-1-7.

Buschel I, Mehdi R, Cammilleri A, et al. Protecting human health and security in digital Europe: how to deal with the 'Privacy Paradox'? *Sci Eng Ethics* 2014; 20: 639–658.

Chen H-T. Revisiting the privacy paradox on social media with an extended privacy calculus model: the effect of privacy concerns, privacy self-efficacy, and social capital on privacy management. *Am Behav Sci* 2018; 62: 1392–1412.

Chen J, He J, Cai L, et al. Disclose more and risk less: privacy preserving online social network data sharing. *IEEE Trans Dependable Secur Comput* 2018; 1.

Chen J, Wang C, He K, et al. Semantics-aware privacy risk assessment using self-learning weight assignment for mobile apps. *IEEE Trans Dependable Secur Comput* 2018; 1.

Chen Y-NK, Wen C-HR. Taiwanese university students' smartphone use and the privacy paradox. *Comunicar* 2019; 27: 61–70.

Cheung M, She J. Evaluating the privacy risk of user-shared images. *ACM Trans Multimed Comput Commun Appl* 2016; 12: 1–21.

Choi B, Wu Y, Yu J, et al. Love at first sight: the interplay between privacy dispositions and privacy calculus in online social connectivity management. *J Assoc Inf Syst* 2018; 19: 124–151.

Choon MJK. Revisiting the privacy paradox on social media: an analysis of privacy practices associated with Facebook and Twitter. *Can J Commun* 2018; 43: 339–358.

Dienlin T, Trepte S. Is the privacy paradox a relic of the past? An in-depth analysis of privacy attitudes and privacy behaviors: the relation between privacy attitudes and privacy behaviors. *Eur J Soc Psychol* 2015; 45: 285–297.

Du S, Li X, Zhong J, et al. Modeling privacy leakage risks in large-scale social networks. *IEEE Access* 2018; 6: 17653–17665.

Elmisery AM, Rho S, Botvich D. Collaborative privacy framework for minimizing privacy risks in an IPTV social recommender service. *Multimed Tools Appl* 2016; 75: 14927–14957.

Feri F, Giannetti C, Jentzsch N. Disclosure of personal information under risk of privacy shocks. *J Econ Behav Organ* 2016; 123: 138–148.

Gerber N, Gerber P, Volkamer M. Explaining the Privacy Paradox–A systematic review of literature investigating privacy attitude and behavior. *Comput Secur*.

Gerber N, Reinheimer B, Volkamer M. Investigating people's privacy risk perception. *Proc Priv Enhancing Technol* 2019; 2019: 267–288.

Goldfeder S, Kalodner H, Reisman D, et al. When the cookie meets the blockchain: Privacy risks of web payments via cryptocurrencies. *Proc Priv Enhancing Technol* 2018; 2018: 179–199.

Gomez-Barroso JL, Feijoo C, Martinez-Martinez LJ. Privacy calculus: factors that influence the perception of benefit. *El Prof La Inf* 2018; 27: 341–348.

Guo J, Li N, Wu Y, et al. Examining help requests on social networking sites: integrating privacy perception and privacy calculus perspectives. *Electron Commer Res Appl*; preproof. Epub ahead of print 2019. DOI: 10.1016/j.elerap.2019.100828.

Gutierrez A, O'Leary S, Rana NP, et al. Using privacy calculus theory to explore entrepreneurial directions in mobile location-based advertising: Identifying intrusiveness as the critical risk factor. *Comput Human Behav* 2019; 95: 295–306.

Haber B. The digital ephemeral turn: queer theory, privacy, and the temporality of risk. *Media, Cult Soc* 2019; 1–19.

Hargittai E, Marwick A. 'What Can I Really Do?' Explaining the privacy paradox with online apathy. *Int J Commun* 2016; 10: 3737–3757.

Harkous H, Rahman R, Karlas B, et al. The curious case of the PDF converter that likes Mozart: dissecting and mitigating the privacy risk of personal cloud apps. *Proc Priv Enhancing Technol* 2016; 2016: 123–143.

Hart S, Ferrara AL, Paci F. Fuzzy-based approach to assess and prioritize privacy risks. *Soft Comput* 2019; 1–11.

Hazard CJ, Singh MP. Privacy risks in intelligent user interfaces. *IEEE Internet Comput* 2016; 20: 57–61.

Hoffman CP, Lutz C, Ranzini G. Privacy cynicism: a new approach to the privacy paradox. *Cyberpsychology*; 10. Epub ahead of print 2016. DOI: http://dx.doi.org/10.5817/CP2016-4-7.

Huckvale K, Prieto JT, Tilney M, et al. Unaddressed privacy risks in accredited health and wellness apps: a cross-sectional systematic assessment. *BMC Med* 2015; 13: 1–13.

James TL, Wallace L, Warkentin M, et al. Exposing others' information on online social networks (OSNs): perceived shared risk, its determinants, and its influence on OSN privacy control use. *Inf Manag* 2017; 54: 851–865.

Karwatzki S, Dytynko O, Trenz M, et al. Beyond the personalization-privacy paradox: privacy valuation, transparency features, and service personalization. *J Manag Inf Syst* 2017; 34: 369–400.

Kehr F, Kowatsch T, Wentzel D, et al. Blissfully ignorant: the effects of general privacy concerns, general institutional trust, and affect in the privacy calculus. *Inf Syst J* 2015; 25: 607–635.

Koohang A, Paliszkiewicz J, Goluchowski J. Social media privacy concerns: trusting beliefs and risk beliefs. *Ind Manag Data Syst* 2018; 118: 1209–1228.

Kordzadeh N, Warren J, Seifi A. Antecedents of privacy calculus components in virtual health communities. *Int J Inf Manage* 2016; 36: 724–734.

Lee J, Kim S, Kim W. The effects of consumers' perceived privacy control on perceived privacy risk in location-based services. *Int J Contents* 2017; 13: 22–30.

Lee J-M, Lee B, Rha J-Y. Determinants of mobile payment usage and the moderating effect of gender: extending the Utaut model with privacy risk. *Int J Electron Commer Stud* 2019; 10: 43–64.

Lee J-M, Rha J-Y. Personalization–privacy paradox and consumer conflict with the use of location-based mobile commerce. *Comput Human Behav* 2016; 63: 453–462.

Lee N, Kwon O. A privacy-aware feature selection method for solving the personalization–privacy paradox in mobile wellness healthcare services. *Expert Syst Appl* 2015; 42: 2764–2771.

Li P, Cho H, Goh ZH. Unpacking the process of privacy management and self-disclosure from the perspectives of regulatory focus and privacy calculus. *Telemat Informatics* 2019; 41: 114–125.

Liu Z, Shan J, Pigneur Y. The role of personalized services and control: an empirical evaluation of privacy calculus and technology acceptance model in the mobile context. *J Inf Priv Secur* 2016; 12: 123–144.

Lo N-W, Yeh K-H, Fan C-Y. Leakage detection and risk assessment on privacy for android applications: LRPdroid. *IEEE Syst J* 2016; 10: 1361–1369.

Mense A, Steger S, Sulek M, et al. Analyzing privacy risks of mHealth applications. *Stud Health Technol Inform* 2016; 221: 41–45.

Metzger MJ, Suh JJ. Comparative optimism about privacy risks on Facebook. *J Commun* 2017; 67: 203–232.

Miltgen CL, Smith HJ. Exploring information privacy regulation, risks, trust, and behavior. *Inf Manag* 2015; 52: 741–759.

Morosan C, DeFranco A. Disclosing personal information via hotel apps: a privacy calculus perspective. *Int J Hosp Manag* 2015; 47: 120–130.

Munyoka W, Maharaj MS. Privacy, security, trust, risk and optimism bias in e-government use: the case of two Southern African Development Community countries. *South African J Inf Manag* 2019; 21: 1–9.

Mvungi B, Iwaihara M. Associations between privacy, risk awareness, and interactive motivations of social networking service users, and motivation prediction from observable features. *Comput Human Behav* 2015; 44: 20–34.

Noguerón-Liu S. 'Everybody Knows Your Business'/ 'Todo Mundo Se Da Cuenta': Immigrant adults' construction of privacy, risk, and vulnerability in online platforms. *J Adolesc Adult Lit* 2017; 60: 505–513.

Orekondy T, Schiele B, Fritz M. Towards a visual privacy advisor: understanding and predicting privacy risks in images. In: 2017 IEEE International Conference on Computer Vision (ICCV). *IEEE*, 2017, pp. 3706–3715.

Ozturk AB, Nusair K, Okumus F, et al. Understanding mobile hotel booking loyalty: an integration of privacy calculus theory and trust-risk framework. *Inf Syst Front* 2017; 19: 753–767.

Pensa RG, Di Blasi G, Bioglio L. Network-aware privacy risk estimation in online social networks. *Soc Netw Anal Min* 2019; 9: 1–15.

Pentina I, Zhang L, Bata H, et al. Exploring privacy paradox in information-sensitive mobile app adoption: A cross-cultural comparison. *Comput Human Behav* 2016; 65: 409–419.

Rider K. The privacy paradox: how market privacy facilitates government surveillance. *Information, Commun Soc* 2018; 21: 1369–1385.

Sampat B, Prabhakar B. Privacy risks and security threats in mHealth apps. *J Int Technol Inf Manag* 2017; 26: 126–153.

Sarabia-Sánchez F-J, Aguado J-M, Martínez-Martínez IJ. Privacy paradox in the mobile environment: the influence of the emotions. *El Prof la Inf* 2019; 28: 1–11.

Sun Y, Wang N, Shen X-L, et al. Location information disclosure in location-based social network services: privacy calculus, benefit structure, and gender differences. *Comput Human Behav* 2015; 52: 278–292.

Trepte S, Reinecke L, Ellison NB, et al. A cross-cultural perspective on the privacy calculus. *Soc Media + Soc*; 3. Epub ahead of print 2017. DOI: 10.1177/2056305116688035.

Van Schaik P, Jansen J, Onibokun J, et al. Security and privacy in online social networking: risk perceptions and precautionary behaviour. *Comput Human Behav* 2018; 78: 283–297.

Wang ES-T. Effects of brand awareness and social norms on user-perceived cyber privacy risk. *Int J Electron Commer* 2019; 23: 272–293.

Wang EST. Role of privacy legislations and online business brand image in consumer perceptions of online privacy risk. *J Theor Appl Electron Commer Res* 2019; 14: 59–69.

Wang T, Duong TD, Chen CC. Intention to disclose personal information via mobile applications: a privacy calculus perspective. *Int J Inf Manage* 2016; 36: 531–542.

Williams M, Nurse JRC. Optional data disclosure and the online privacy paradox: a UK perspective. In: Fourth International Conference on Human Aspects of Information Security, Privacy and Trust at the 18th International Conference on Human. Kent Academic Repository, 2016. Epub ahead of print 2016. DOI: 10.1007/978-3-319-39381-0_17.

Wright D, Raab C. Privacy principles, risks and harms. *Int Rev Law, Comput Technol* 2014; 28: 277–298.

Wu PF. The privacy paradox in the context of online social networking: a self-identity perspective. *J Assoc Inf Sci Technol* 2019; 70: 207–217.

Xu H, Hao S, Sari A, et al. Privacy risk assessment on email tracking. In: IEEE INFOCOM 2018–IEEE Conference on Computer Communications. *IEEE*, 2018, pp. 2519–2527.

Yin L, Wang Q, Shaw SL, et al. Re-identification risk versus data utility for aggregated mobility research using mobile phone location data. *PLoS One*; 10. Epub ahead of print 2015. DOI: 10.1371/journal.pone.0140589.

Zhu H, Ou CXJ, van den Heuvel WJAM, et al. Privacy calculus and its utility for personalization services in e-commerce: an analysis of consumer decision-making. *Inf Manag* 2017; 54: 427–437.