Visualization and Prioritization of Privacy Risks in Software Systems

George O. M. Yee
Computer Research Lab, Aptusinnova Inc., Ottawa, Canada
Dept. of Systems and Computer Engineering, Carleton University, Ottawa, Canada
email: george@aptusinnova.com, gmyee@sce.carleton.ca

Abstract—**Software systems are ubiquitous in almost every aspect of our lives, as can be seen in social media, online banking and shopping, as well as electronic health monitoring. This widespread involvement of software in our lives has led to the need to protect privacy, as the use of the software often requires us to input our personal information. However, before privacy can be protected, it is necessary to understand the risks to privacy that can be found in the software system. In addition, it is important to understand how the risks can be prioritized since budgetary constraints usually mean that not all risks will be mitigated. Indeed, understanding the risks and prioritizing them is key to protecting privacy throughout the system's range of application. This paper presents straightforward methods for effectively visualizing, identifying, and prioritizing privacy risks in software systems, and illustrates the methods with examples.**

Keywords-software; system; privacy; risks; visualization; prioritization.

I. INTRODUCTION

The rapid growth of the Internet has been accompanied by numerous software systems targeting consumers. Software systems are available for banking, shopping, learning, healthcare, and Government Online. However, most of these systems require a consumer's personal information in one form or another, leading to concerns over privacy. For these systems to be successful, privacy must be protected.

This work extends Yee [1] by expanding the sections on privacy and risk visualization. Further, a new section on risk prioritization has been added.

Various approaches have been used to protect personal information, including data anonymization [2] and pseudonym technology [3]. Other approaches for privacy protection include treating privacy protection as an access problem and then bringing the tools of access control to bear for privacy control [4]. However, these approaches presume to know where and what protection is needed. They presume that some sort of analysis has been done that answers the question of "where" and "what" with respect to privacy risks. Without such answers, the effectiveness of the protection comes into question. The total risks to data depends both on the number of vulnerable locations of the data (where) and on the severity of each vulnerability (what). For example, protection against house break-ins is ineffective if the owner only secures the front door without securing other vulnerable spots such as windows. An effective break-in risk analysis would have identified the windows as additional locations having break-in risks (where and what) and would have led to the windows also being secured. The result is a house that is better protected against break-ins. In the same way, privacy risk identification considering "where" and "what" is essential to effective privacy protection - this work proposes a visual method for such identification.

The objectives of this paper are to a) propose an effective method for visualizing privacy risks in software systems to identify where and what risks are present, b) propose a straightforward method for prioritizing the risks for mitigation, since not all risks can be mitigated due to financial constraints, and c) illustrate the method using examples.

In the literature, there are significant works on security threat analysis but very little work on privacy risk identification using visualization. In fact, the only works that are directly related to privacy risk identification appear to be those on "privacy impact assessment (PIA)", originating from government policy [5]. PIA is meant to evaluate the impact to privacy of new government programs, services, and initiatives. PIA can also be applied to existing government services undergoing transformation or re-design. However, PIA is a long manual process consisting mainly of self-administered questionnaires. It is not focused on software systems nor does it employ visual techniques as proposed in this work.

This paper is organized as follows. Section II defines privacy, privacy preferences, privacy risks, and what they mean for software systems. Section III presents the proposed method for privacy risk visualization, together with examples. Section IV presents the method for prioritizing privacy risks. Section V examines the strengths and weaknesses of the approach, including potential improvements. Section VI discusses related work. Section VII presents conclusions and future work.

II. PRIVACY

As defined by Goldberg et al. in 1997 [6], privacy refers to the ability of individuals to *control* the collection, retention, and distribution of information about themselves. This leads to the following definition of privacy for this work.

DEFINITION 1: *Privacy* refers to the ability of individuals to *control* the collection, purpose, retention, and distribution of information about themselves.

Definition 1 is the same as given by Goldberg et al. except that it also includes "purpose". To see that "purpose" is needed, consider, for example, that one may agree to give out one's email address for the purpose of friends to send email but not for the purpose of spammers to send spam. This definition also suggests that "personal information", "private information" or "private data" is any information that can be linked to a person; otherwise, the information would not be "about" the person. Thus, another term for private information is "personally identifiable information (PII)". These terms are used interchangeably in this paper. In addition, controlling the "collection" of information implies controlling *who* collects *what* information. Controlling the "retention" of information is really about controlling the *retention time* of information, i.e. how long the information can be retained before being destroyed. Controlling the "distribution" of information is controlling to which other parties the information can be *disclosed-to*. These considerations motivate the following definitions.

DEFINITION 2: A user's *privacy preference* expresses the user's desired control over a) *PII* - what the item of personal information is, b) *collector* - who can collect it, c) *purpose* - the purpose for collecting it, d) *retention time* - the amount of time the information is kept, and e) *disclosed-to* - which other parties the information can be disclosed-to.

DEFINITION 3: A *privacy risk* is the potential occurrence of any action or circumstance that will result in a violation of any of the components PII, collector, purpose, retention time, and disclosed-to in a user's privacy preference.

For example, Alice uses an online pharmacy and has the following privacy preference:

PII: name, address, telephone number
Collector: A-Z Drugs
Purpose: identification
Retention Time: 2 years
Disclosed-To: none

This preference states that Alice allows A-Z Drugs to collect her name, address, and telephone number, and that A-Z Drugs must: use the information only to identify her, not keep the information for more than 2 years, and not disclose the information to any other party.

This work considers only privacy risks as defined in Definition 3. The privacy preference components PII, collector, purpose, retention time, and disclosed-to have, in fact, been standardized by the Canadian Standards Association in its Model Code for the Protection of Personal Information [7]. The Model Code is based on ten privacy principles as given in Table I. As can be seen in Table I, PII is reflected in principle 3 (which PII requires consent), collector is seen in principle 1 (collector's accountability) and principle 5 (disclosure to other collectors), purpose is contained in principles 2 and 4, and finally, retention time and disclosed-to are seen in principle 5. Further, these privacy preference components have been enacted by privacy legislation as fully describing the privacy rights of individuals in many countries, including Canada, the United States, the European Union, and Australia [8]. Thus, this work is consistent with privacy legislation, and treating only privacy risks defined by Definition 3 does not overly reduce the generality of this work.

TABLE I. Ten Privacy Principles Forming Basis of Model Code

Principle	Description
1. Accountability	An organization is responsible for personal information under its control and shall designate an individual or individuals accountable for the organization's compliance with the privacy principles.
2. Identifying Purposes	The purposes for which personal information is collected shall be identified by the organization at or before the time the information is collected.
3. Consent	The knowledge and consent of the individual are required for the collection, use or disclosure of personal information, except when inappropriate.
4. Limiting Collection	The collection of personal information shall be limited to that which is necessary for the purposes identified by the organization. Information shall be collected by fair and lawful means.
5. Limiting Use, Disclosure, and Retention	Personal information shall not be used or disclosed for purposes other than those for which it was collected, except with the consent of the individual or as required by the law. In addition, personal information shall be retained only as long as necessary for fulfillment of those purposes.
6. Accuracy	Personal information shall be as accurate, complete, and up-to-date as is necessary for the purposes for which it is to be used.
7. Safeguards	Security safeguards appropriate to the sensitivity of the information shall be used to protect personal information.
8. Openness	An organization shall make readily available to individuals specific information about its policies and practices relating to the management of personal information.
9. Individual Access	Upon request, an individual shall be informed of the existence, use and disclosure of his or her personal information and shall be given access to that information. An individual shall be able to challenge the accuracy and completeness of the information and have it amended as appropriate.
10. Challenging Compliance	An individual shall be able to address a challenge concerning compliance with the above principles to the designated individual or individuals accountable for the organization's compliance.

The following works show the importance of privacy in the online world: Tene [9], Kambourakis [10], Ruiz-Martinez [11], and Ren and Wu [12]. In addition, Pfitzmann and Hansen [13] present some terminology for talking about privacy, e.g., "anonymity", "unlinkability".

III. METHOD FOR PRIVACY RISK VISUALIZATION

The proposed method for privacy risk visualization assumes the following common characteristics of a software system:

a) The software system requires the user's personal information in order to carry out its function. For example, an online bookseller requires the user's address for shipping purposes.

b) The software system may transmit the information (e.g., move it from one group to another within the software system's organization), store the information (e.g., store the information in a data base), and make use of the information to carry out its function (e.g., print out shipping labels with the user's address).

The method is based on the notion that the *location* of personal information gives rise to privacy risks. The importance of location is reflected in physical security, where sensitive paper documents are kept in a locked safe (a location) to protect privacy, rather than being left on a desk (a location). For a software system, storing the user's personal information in an encrypted database with secure access controls is the equivalent of storing it in a safe, with corresponding reduced privacy risks. The method employs notation, as given in Table II.

TABLE II. Notation for Visualizing Privacy Risks

Element		Description
Use Circle	○	Identifies where PII is used. Labeled with a letter together with a description of the use in a legend.
Data Store	□	Identifies where PII is stored. Labeled with a letter together with a description of the data store in a legend.
Same Physical Platform	⬚	Identifies use circles and data stores that execute on the same computing platform.
PII Data Flo w	⟶	Identifies the movement of PII from one location to another. Labeled with a number together with a description of the data in a legend.
Non-PII Data Flow	⇢	Identifies the movement of non-PII from one location to another. Labeled with a number together with a description of the data in a legend.
Legend		Descriptions corresponding to the letters or numbers with which the above notational elements were labeled.

The method, then, consists of i) determining all the possible locations in the software system where the user's personal information could reside, and ii) visualizing at each of these locations the possible ways in which the user's privacy preferences could be violated. The complete method is as follows:

A. *Method for Privacy Risk Visualization*

1. Draw the paths of all personal information flows within the software system, based on characteristic b) above, namely, that personal information can be transmitted, stored, and used. Use a solid arrow to represent the transmission of personal information items that are described by privacy preferences. Label the arrow with numbers, where each arrow number corresponds to a description of a personal data item in a legend. Use a square to represent the storage of personal information. Use a circle to denote the use of the information. Use a dashed rectangle to enclose circles or squares into physically distinct units. For example, two circles representing two uses would be enclosed by a dashed square if both uses run on the same computing platform. Physically separate units allow the identification of risks for any data flow between them. Circles or squares not enclosed by a dashed rectangle are understood to be already physically separate units. Label the squares and circles with letters. Each such label corresponds to a description of the type of storage or the type of use as indicated in the legend.
2. Use dashed arrows, numbered in the same way as the solid arrows in Step 1, to add to the drawing all non-personal information flows, if any, that are involved with the transmission, storage and use of the personal information. Non-personal information is information that is not personal or not private, i.e., information that cannot identify any particular individual, e.g., the price of something. The resulting drawing is called a Personal Information Map (PIM). Figure 1 illustrates steps 1 and 2 for the software system of an online seller of merchandise, e.g., Amazon.com, that requires the user's name, address, merchandise selection, and credit card number. These are considered as three personal information items where name and address together are considered as one item. Figure 1 also shows three non-personal information flows (4, 5, 6). The dashed rectangle enclosing A, B, and C indicates that A, B, and C all run on the same physical computing platform.
3. Inspect the PIM resulting from step 2, and for each location (flow arrow, storage square, and use circle) and each personal information item, visualize the possible ways in which a privacy preference may be violated in terms of violations of any of *PII*, *collector*, *purpose*, *retention time*, and *disclose-to* (see Section II). This may be achieved by asking risk questions for each component, as proposed in Table III, and drawing conclusions based on security and systems knowledge and experience. The risk questions are "how" questions, based on the idea that a risk arises where there is some way (i.e. how) for a violation to occur. This step actually calls for visualization since one is tasked with exploring the possible risks in conjunction with a visual notation, the PIM. Record the results in a Privacy Risks Table containing two columns: the left column for records of the form "(PII$_1$, PII$_2$, …/ locations)" and the right column containing the corresponding privacy risks. The Privacy Risks Table is the goal of the method. Table IV illustrates this step for the online seller of Fig. 1.

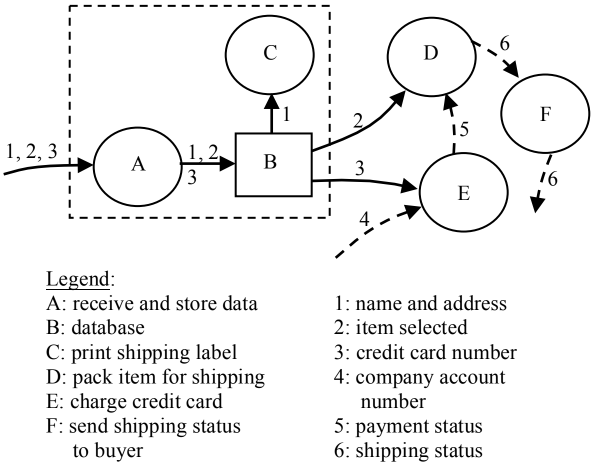


Figure 1. PIM for an online seller of merchandise.

TABLE III. Risk Questions

Component	Risk Questions
PII	How can the user be asked for other PII, either intentionally or inadvertently?
collector	How can the PII be received by an unintended collector, either in addition to or in place of the intended collector?
purpose	How can the PII be used for other purposes?
retention time	How can the PII retention time be violated?
disclose-to	How can the PII be disclosed either intentionally or inadvertently to an unintended recipient?

TABLE IV. Partial Privacy Risks Table Corresponding to Fig. 1

(PIIs / locations)	Privacy Risks
(1, 2, 3 / path into A); (2 / path into D); (3 / path into E)	Man-in-the-middle attack violates *collector*, *purpose*, *retention time* and *disclose-to*.
(1, 2, 3 / A)	User could be asked for personal information that violates *PII,* i.e. asked for personal information other than as specified in the user's privacy preferences.
(1, 2, 3 / A); (1 / C); (2 / D); (3 / E)	Trojan horse, hacker attack use circles violating *collector*, *purpose*, *retention time*, and *disclose-to.*
(1, 2, 3 / B)	SQL attack on B violates *collector*, *purpose*, *retention time*, and *disclose-to.*
(1, 2, 3 / B)	*PII* in B could be kept past its *retention time.*

It is important to note that the PIM resulting from Step 2 is not a program logic flow diagram and one should not try to interpret it as such. It shows *what* PII is required, *where* PII goes, *where* PII is stored, and *where* PII is used, corresponding to the notion that the location of personal information is key to understanding privacy risks, as mentioned above.

Privacy risks and security risks are conceptually different. However, a privacy risk may be due to a security risk, and vice versa. For example, the privacy risk associated with a man-in-the-middle attack in Table IV is really due to the security risk of a man-in-the-middle attack. Again in Table IV, a higher security risk of theft can be attributed to the privacy risk of PII being kept past its retention time, since the longer the PII is retained, the greater the security risk of it being stolen.

Adding non-personal information flows in Step 2 is important to help identify potential unintended leakages of PII. For example, consider a "produce report" use circle that "anonymizes" (any obvious links to the information owner removed) PII and combines the result with non-personal information to produce a report for public distribution. The fact that both PII and non-PII flow into "produce report" could lead to identifying a personal information leakage risk.

It is recommended that this method be applied by a privacy risks identification team, consisting of no more than three or four people, selected for their technical knowledge of the software system and the work procedures and processes of the software system's organization. Good candidates for the team include the software system's design manager, test manager, and other line managers with the required knowledge. The team should be led by a privacy and security analyst, who must also be knowledgeable about the software system, and who must have the support of upper management to carry out the privacy risks identification. A definite advantage of the team approach would accrue to step 3, where the visualization would be more thorough by virtue of more people being involved.

B. First Application Example

Consider PatientBilling, a patient billing system running in a doctor's office. PatientBilling makes use of two business software systems: an accounting system PatientAccounting and an online payment system PatientPay.

Table V shows the user's personal information required by each system. The user provides her private information to PatientBilling which then discloses this information to PatientAccounting and PatientPay.

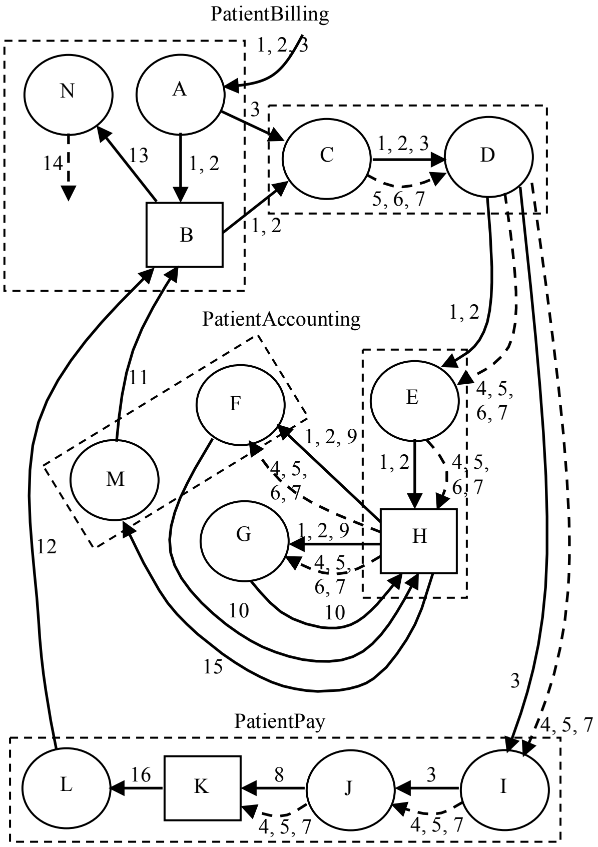
TABLE V. Personal Information Required

Software System	Patient Personal Information Required
PatientBilling	name and address, health complaint (patient name, health problem, health problem resolution), method of payment details (name, credit card number, credit card expiry date, health insurance number, health insurance expiry date)
PatientAccounting	name and address, health complaint (as above)
PatientPay	method of payment details (as above)

The proposed method for privacy risks visualization is carried out as follows:

Steps 1 and 2: Draw the PIM for each software system (see Fig. 2**).** As shown in Figure 2, the following uses of personal information are extra to the core function of each system. First, both PatientAccounting (M) and PatientPay

(L) send activity reports back to PatientBilling that contain personal information. These reports contain selections and re-arrangements of personal data (15, 16). Second, PatientBilling produces a publically accessible report for the medical association, giving statistics on the patients seen. To produce this report, PatientBilling (N) selects, re-arranges, and anonymizes personal data (13). Third, PatientAccounting allows its employees to partially work from home (G). Finally, the patient's method of payment details are used without being stored in databases.



Legend:
A: receive and store data
B: database
C: process billing
D: disclose data
1: name and address
2: health complaint
3: method of payment details
4: doctor id
5: billing id
6: time spent with patient
7: billing amount
8: doctor account update
9: current ledger record
10: updated ledger record
11: accounting report
12: payment report
13: patients seen data

E: receive and store data
F: update ledgers at work
G: update ledgers at home
H: database
I: receive and forward data
J: charge credit card or insurance; update doctor's account
K: database
L: compose payment report
M: compose accounting report
N: compose report for medical association
14: anonymized report for medical association
15: accounting data
16: payment data

Figure 2. PIM for PatientBilling, PatientAccounting, and PatientPay.

Step 3: Visualize privacy risks at private information locations. Table VI gives a partial Privacy Risk Table for locations in Fig. 2 that have interesting or serious privacy risks. The theft of personal information means that the information is under the control of an unintended party. Clearly, this can violate the corresponding privacy preference or preferences in terms of violating *collector*, *purpose*, *retention time*, and *disclose-to*. The risk of personal information theft arises so often that it is convenient to call it *CPRD-risk*, from the first letters of collector, purpose, retention time, and disclose-to.

TABLE VI. Partial Privacy Risks Table Corresponding to Fig. 2

(PIIs / locations)	Privacy Risks
(1, 2, 3 / path into A); (1, 2 / path between B and C, path between D and E); (3 / path between A and C, path between D and I); (12 / path between L and B); (11 / path between M and B)	Man-in-the-middle attacks lead to CPRD-risk.
(1, 2, 3 / A)	The patient could be asked for personal information that violates PII (i.e. asked for PII other than 1, 2, 3).
(1, 2, 3 / A, C, D); (13 / N); (1, 2 / E); (1, 2, 9 / F, G); (15 / M); (3 / J); (16 / L)	Trojan horse, or hacker attacks on the personal information use circles lead to CPRD-risk.
(1, 2, 11, 12 / B); (1, 2, 10 / H); (8 / K)	Potential SQL attacks on B, H, and K lead to CPRD-risk.
(13 / N)	A bad anonymization algorithm can expose personal information, leading to CPRD-risk.
(1, 2, 9 / G)	An insecure home environment, e.g., people looking over the shoulder or printed personal information lying on a desk in the clear, can also lead to CPRD-risk.
(1, 2, 9 / G)	If an employee works from home on a laptop and carries the laptop between home and work, possible theft or loss of the laptop can also lead to CPRD-risk for any of 1, 2, or 9 that might be temporarily stored in the laptop.
(1, 2, 9 / G)	If an employee works from home on a home PC and stores 1, 2, 9 on a flash memory stick, carrying the memory stick between home and work, possible theft or loss of the memory stick can also lead to CPRD-risk.

To illustrate this step, the risks in the first 3 rows of Table VI were obtained as follows. For the first row, it was noticed that the personal information flows through transmission paths connecting physically distinct units. The risk questions of Table III were then considered, leading to possible man-in-the-middle attacks that give rise to CPRD-risk. For the second row, violations of PII are always possible unless strict controls are in place against it. For the third row, it was observed that the associated personal data are input to information use processes (e.g., A, C, D). The risk questions of Table III were again considered, leading to possible Trojan horse or hacker attacks that again give rise to CPRD-risk. For the fourth row, it was noticed that personal data are stored in databases. Once again the risk questions were considered, leading to possible SQL attacks against the databases, giving rise to CPRD-risk. In each of these four cases, knowledge of the system (personal data locations) and knowledge of information security (possible attacks) were needed to identify the risks. The remaining risks in Table VI were derived in a similar fashion.

B. Second Application Example

Consider an airline reservation system called AccuReserve offered by a Canadian airline with headquarters in Toronto, Canada. AccuReserve is a globally distributed system with modules in Canada, the United States, and Germany (serving the European Union).

Table VII shows the user's personal information required by the country specific modules of AccuReserve. The user provides her private information to each of these modules when she makes a travel reservation.

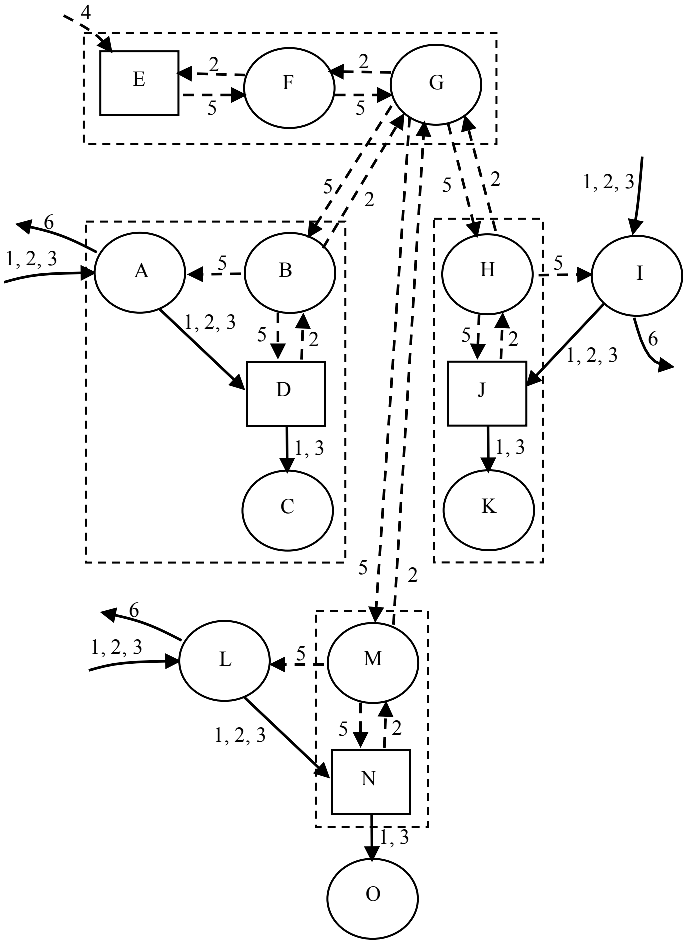
TABLE VII. Personal Information Required

Software Module	Patient Personal Information Required
Canada	Identification details (name, address, telephone number, nationality, passport number); payment details (credit card name, credit card number, credit card expiry date, credit card verification code)
United States	Same as above
Germany	Same as above

The proposed method for privacy risks visualization is carried out as follows:

Steps 1 and 2: The PIM for AccuReserve is shown in Fig. 3, and was obtained by drawing the PIM for each module (Main, Mod-US, and Mod-EU) and then linking the modules together with communication links. Main runs in Canada, Mod-US in the United States, and Mod-EU in Germany.

Step 3: Table VIII gives a partial Data Risk Table for locations in Fig. 3 that have PII risks. The privacy risks in Table VIII were obtained as follows. For the first and second rows, it was noticed that the personal information flows through transmission paths connecting physically distinct units. The risk questions of Table III were then considered, leading to possible man-in-the-middle attacks that give rise to CPRD-risk. Notice that "(1, 2, 3 / path between A and D)" is excluded because A and D both run on the same platform (so the path is not very accessible to attack). For the third row, violations of PII are always possible unless strict controls are in place against it. For the fourth row, it was observed that private data are input to information use processes A, I, L, C, K, O. The risk questions of Table III were again considered, leading to possible Trojan horse or hacker attacks that again give rise to CPRD-risk. For the fifth row, it was noticed that private data are stored in databases. Once again the risk questions were considered, leading to possible SQL attacks against the databases, giving rise to CPRD-risk. For the sixth row, it was noticed that private information stored in databases could be subject to insider attacks. For the seventh row, it was observed that the private data stored in the databases could be kept past their retention times. It should be noted that the links between G and B, G and M, and G and H are also vulnerable to man-in-the-middle attacks, but these attacks would not be privacy attacks, since these links are not used for private information.



Legend:

A, I, L: receive and store data
B, H, M: communicate with G
C, K, O: charge credit card
E: flights database (SD)
D, J, N: customer databases
F: flight availability manager
G: communicate with countries

1: identification details (PII)
2: flight details requested (non-PII, non-SD)
3: payment details (PII)
4: flight availability updates (SD)
5: flight details assigned (non-PII, non-SD)
6: travel itinerary (PII)

Figure 3. PIM for AccuReserve; Main consists of E, F, G, A, B, D, and C; Mod-US consists of L, M, N, and O; Mod-EU consists of H, I, J, and K.

IV. METHOD FOR PRIVACY RISK PRIORITIZATION

In this work, the concept behind privacy risk prioritization is that once a set of n privacy risks have been identified, we want to prioritize or select a subset k, $k < n$, of those risks for mitigation, given that we do not have sufficient financial resources to mitigate all n of the risks.

NOTATION: Let R be the set of identified privacy risks. Let P, $P \subset R$, be a subset of risks to be mitigated. Let ρ be the prioritization mapping such that $\rho\colon R \longrightarrow P$.

Our purpose in this section is to define the prioritization mapping ρ. In other words, we seek a method for selecting risks for mitigation (determining the set P). Intuitively, one would want to mitigate risks that are highly probable to be realized, and that once realized, would result in very costly damages. Due to financial budgetary constraints, we feel that we can ignore the risks that tend not to be realized and even if realized would cause very little damage. Determining which risks to mitigate may be assisted though weighting the risks according to certain criteria.

TABLE VIII. Partial Data Risks Table Corresponding to Fig. 3

(PIIs / locations)	Privacy Risks

(1, 2, 3 / path into A);
(1, 2, 3 / path into I);
(1, 2, 3 / path into L); (6 / path from A); (6 / path from I); (6 / path from L)
	Man-in-the-middle attacks lead to CPRD-risk.

(1, 2, 3 / path between I and J); (1, 2, 3 / path between L and N); (1, 3 / path between N and O)	Man-in-the-middle attacks lead to CPRD-risk.
(1, 2, 3 / path into A); (1, 2, 3 / path into I); (1, 2, 3 / path into L)	The user could be asked for personal information that violates PII (i.e. asked for PII other than 1, 2, 3).
(1, 2, 3 / A, I, L); (1, 3 / C, K, O)	Trojan horse, or hacker attacks on the personal information use circles lead to CPRD-risk.
(1, 2, 3 / D, J, N)	Potential SQL attacks on D, J, and N lead to CPRD-risk.
(1, 2, 3 / D, J, N)	Potential insider attack steals private information from D, J, and N resulting in CPRD-risk.
(1, 2, 3 / D, J, N)	Private information in D, J, and N could be kept past the retention time.

Salter et al. [14] proposed a method for applying weights to various forms of attacks in order to determine if a particular attack would be probable. They focused on three aspects of an attack, namely "risk", "access", and "cost", where "risk" is risk to the safety of the attacker, "access" is the ease with which the attacker can access the system under attack, and "cost" is the monetary cost to the attacker to mount the attack. To avoid confusion between "risk" to the safety of the attacker and "risk" to privacy, we use "safety" for "risk" to the safety of the attacker. The weight values are simply "L", "M", and "H" for Low, Medium, and High, respectively. These attack aspects can be represented using a 3-tuple, as [safety, access, cost] and so [H, M, L] would be an instance of the weights. For example, consider a physical attack such as a mugging incident in a park. In this case, the risk to the safety of the attacker would be high (the person being mugged could be an undercover police officer), the attacker's ease of access would be high (people stroll through the park all the time), and the attacker's cost would be low (not much needed to mount the attack). Thus, this attack has the weights [H, H, L].

In this work, we add a fourth aspect of an attack, namely the resulting damages from the attack. Thus, we use the 4-tuple [safety, access, cost, damages] with the same weight values L, M, and H. Hence, we would definitely want to defend against privacy risks leading to attacks with weights [L, H, L, H]. We feel that we can ignore privacy risks having attacks with weights [H, L, H, L]. In reality, there is a spectrum of weights between these two boundaries, where a decision to defend or ignore may not be clear, and ultimately a judgment, perhaps based on other factors, may be needed. For example, it is not clear whether or not a privacy risk with associated weights [L, L, H, H] should be ignored, and one would decide to defend if one believes that no matter how improbable the attack, the resulting damages must never be allowed to occur.

The uncertainty of deciding which risks to mitigate using the weights may be remedied through the use of a Prioritization Policy, which would be developed by the privacy and security analyst (see Section IIIA). This policy would identify the 4-tuples of weights whose associated risks are to be prioritized or mitigated. For example, the policy might state that risks with associated 4-tuples [L, *, *, H] and [L, *, *, M] are to be mitigated, where "*" indicates possibilities L, M, and H. We are now ready to define ρ.

DEFINITION 4: (Method for Privacy Risk Prioritization, ρ) Apply weights to the privacy risks in R using the procedure described in Section IV above. Select the risks for prioritization (or mitigation) based on the Prioritization Policy.

Prioritization Examples

Examples of the application of Definition 4 may be obtained by re-visiting and prioritizing the risks found in the privacy risk tables above (Tables IV, VI, and VIII). Two extra columns are added to each privacy risk table: one column for the weights, and one column identifying P, the set of risks that have been prioritized.

Re-visiting Table IV, adding the weights, and prioritizing using a Prioritization Policy that states "only prioritize (mitigate) risks with weights [*, *, L, H]", where * admits possibilities L, M, H gives Table IX.

The weights in Table IX were assigned by the privacy and security analyst as follows. For the man-in-the-middle attack, the risks to the attacker's safety is low since he or she is attacking at a distance; the access is high since it's the Internet; the cost is low as not much equipment is needed; the damages would be high since the attacker could post the private information leading to heavy damages to the company's reputation. Similar considerations apply to the weight assigned to the Trojan horse or hacker attack. For the

SQL attack on B, accessibility was assigned as low and cost as high because improvements to the database user interface were recently carried out to guard against SQL attacks. The risk of the user being asked for information violating PII and the risk of information kept past the retention time were considered as potential accidents caused by the company itself. Therefore, the risk to safety, the accessibility, and the costs were deemed to be low, high, and low respectively. The resulting damages were considered to be medium because the accidents would likely be quickly discovered through auditing and remedied.

TABLE IX. Partial Prioritized Privacy Risks Table Corresponding to Fig. 1

(PIIs / locations)	Privacy Risks	Weights	In *P*
(1, 2, 3 / path into A); (2 / path into D); (3 / path into E)	Man-in-the-middle attack violates *collector*, *purpose*, *retention time* and *disclose-to*.	[L, H, L, H]	Yes
(1, 2, 3 / A)	User could be asked for personal information that violates *PII,* i.e. asked for personal information other than as specified in the user's privacy preferences.	[L, H, L, M]	No
(1, 2, 3 / A); (1 / C); (2 / D); (3 / E)	Trojan horse, hacker attack use circles violating *collector*, *purpose*, *retention time*, and *disclose-to*.	[L, H, L, H]	Yes
(1, 2, 3 / B)	SQL attack on B violates *collector*, *purpose*, *retention time*, and *disclose-to*.	[L, L, H, H]	No
(1, 2, 3 / B)	*PII* in B could be kept past its *retention time.*	[L, H, L, M]	No

Table VI is prioritized next giving Table X. This time the Prioritization Policy used states "only prioritize (mitigate) risks with weights [*, H, L, H]" where * admits possibilities L, M, H. The analyst assigned the weights in Table X as follows. The weights for the man-in-the-middle attack, the violation of PII, and the Trojan horse or hacker attack are the same as in Table IX since they are the same attacks. The SQL attack was assigned the same weight as the Trojan horse or hacker attack since they have similar safety, access, and cost requirements, and the aftermath of which would also be highly damaging. The bad anonymization algorithm is considered as accidental and is assigned the same weight as the violation of PII, which is also considered accidental. The insecure home environment is assigned H for safety since the attacker could be easily caught, M for access since it's a private home, L for cost since it does not cost anything to look, and H for damages since lost of the information is highly damaging. The theft of the laptop (theft is considered here rather than accidental loss) is assigned H for safety since the thief could be observed and caught, M for access since the laptop can be a

TABLE X. Partial Prioritized Privacy Risks Table Corresponding to Fig. 2

(PIIs / locations)	Privacy Risks	Weights	In *P*
(1, 2, 3 / path into A); (1, 2 / path between B and C, path between D and E); (3 / path between A and C, path between D and I); (12 / path between L and B); (11 / path between M and B)	Man-in-the-middle attacks lead to CPRD-risk.	[L, H, L, H]	Yes
(1, 2, 3 / A)	The patient could be asked for personal information that violates PII (i.e. asked for PII other than 1, 2, 3).	[L, H, L, M]	No
(1, 2, 3 / A, C, D); (13 / N); (1, 2 / E); (1, 2, 9 / F, G); (15 / M); (3 / J); (16 / L)	Trojan horse, or hacker attacks on the personal information use circles lead to CPRD-risk.	[L, H, L, H]	Yes
(1, 2, 11, 12 / B); (1, 2, 10 / H); (8 / K)	Potential SQL attacks on B, H, and K lead to CPRD-risk.	[L, H, L, H]	Yes
(13 / N)	A bad anonymization algorithm can expose personal information, leading to CPRD-risk.	[L, H, L, M]	No
(1, 2, 9 / G)	An insecure home environment, e.g., people looking over the shoulder or printed personal information lying on a desk in the clear, can also lead to CPRD-risk.	[H, M, L, H]	No
(1, 2, 9 / G)	If an employee works from home on a laptop and carries the laptop between home and work, possible theft or loss of the laptop can also lead to CPRD-risk for any of 1, 2, or 9 that might be temporarily stored in the laptop.	[H, M, L, H]	No
(1, 2, 9 / G)	If an employee works from home on a home PC and stores 1, 2, 9 on a flash memory stick, carrying the memory stick between home and work, possible theft or loss of the memory stick can also lead to CPRD-risk.	[H, M, L, H]	No

little difficult to get to (e.g., inside a car), L for cost since it does not cost much to execute, and H for damages as again such a loss would be very damaging. The theft of the memory stick (theft is considered rather than loss) is assigned the same weights as the theft of the laptop since they have similar dangers and requirements for the attacker, and is also very damaging.

Table VIII is the last privacy risks table to be prioritized, giving Table XI. This time the Prioritization Policy used states "only prioritize (mitigate) risks with weights [L, *, *, H]" where * admits possibilities L, M, H. The analyst assigned the weights in Table XI as follows.

TABLE XI. Partial Prioritized Data Risks Table Corresponding to Fig. 3

(PIIs / locations)	Privacy Risks	Weights	In *P*
(1, 2, 3 / path into A); (1, 2, 3 / path into I); (1, 2, 3 / path into L); (6 / path from A); (6 / path from I); (6 / path from L)	Man-in-the-middle attacks lead to CPRD-risk.	[L, H, L, H]	Yes
(1, 2, 3 / path between I and J); (1, 2, 3 / path between L and N); (1, 3 / path between N and O)	Man-in-the-middle attacks lead to CPRD-risk.	[M, M, L, H]	No
(1, 2, 3 / path into A); (1, 2, 3 / path into I); (1, 2, 3 / path into L)	The user could be asked for personal information that violates PII (i.e. asked for PII other than 1, 2, 3).	[L, H, L, M]	No
(1, 2, 3 / A, I, L); (1, 3 / C, K, O)	Trojan horse, or hacker attacks on the personal information use circles lead to CPRD-risk.	[L, H, L, H]	Yes
(1, 2, 3 / D, J, N)	Potential SQL attacks on D, J, and N lead to CPRD-risk.	[L, H, L, H]	Yes
(1, 2, 3 / D, J, N)	Potential insider attack steals private information from D, J, and N resulting in CPRD-risk.	[L, H, L, H]	Yes
(1, 2, 3 / D, J, N)	Private information in D, J, and N could be kept past the retention time.	[L, H, L, M]	No

A weight of [L, H, L, H] was assigned to the first row after the same considerations as that described for man-in-the-middle attacks in Table IX. A weight of [M, M, L, H] was assigned to the second row since the paths in this row are relatively short (connecting components in the same module), leading to greater risk for the attacker (greater risk of being seen) and lower accessibility (fewer places to access the link). A weight of [L, H, L, M] was assigned to the third and last rows out of the same considerations as in Table IX, for the risk of the user being asked for information that violates PII and the risk of private information kept past the retention time. A weight of [L, H, L, H] was assigned to the Trojan horse or hacker attack in the fourth row and the SQL attacks in the fifth row since the attacker could operate from a distance with easy access through the Internet and with relatively low costs. A weight of [L, H, L, H] was assigned to the risk of an insider attack in the sixth row since an insider can hide in plain sight, has high access by virtue of being an insider, and carry out the attack at zero cost to herself.

V. DISCUSSION OF STRENGTHS, WEAKNESSES, AND IMPROVEMENTS

Some of the strengths of the approach include: a) provides a structured straightforward way to identify and prioritize privacy risks, b) user friendly common sense graphical notation, and c) based on the locations that involve PII, a concept that is easily understood.

Some weaknesses of the method are: a) drawing the PIM, filling out the Privacy Risks Table, and prioritizing the risks require expertise in how personal information is used as well as expertise in security and privacy, b) drawing the PIM is manual and is prone to error, c) the prioritization is partly subjective, and d) the method can never identify all the risks. Weakness a) is unavoidable as the expertise must be available somehow. This requirement for expertise is common to many technical endeavors, e.g., software engineering. Weakness b) can be addressed by building tools for automatically drawing the PIM. Similar tools already exist for rendering a software architecture diagram from the reverse engineering of code, e.g., Nanthaamornphong et al. [15]. Furthermore, automated analysis of the PIM should be feasible by using a rules engine to automate the visualization or enumeration of privacy risks, based on machine understanding of the graphical notation in this work. These automations should improve both the accuracy of the PIM and the identification of the privacy risks. Weakness c) may be attenuated by having a team of experts assign the weights through consensus. The accuracy of the prioritization may also be improved by considering other factors such as the nature and frequency of recent attacks, as well as the cost of mitigating a risk. Weakness d) may also be unavoidable, as it is mostly due to the nature of security, that no system can be completely secure. However, the above automated tools and rules engine should improve risk coverage.

VI. RELATED WORK

The literature on works by other authors, dealing *directly* with privacy risk visualization for software systems, appears to be non-existent. However, the following authors have written on topics that are related to privacy risk analysis. Hong et al. [16] propose the use of privacy risk models to help designers design ubiquitous computing applications that have a reasonable level of privacy protection. Their privacy risk model consists of two parts: a privacy risk analysis part and a privacy risk management part. The risk analysis identifies the privacy risks while the risk management part is a cost-benefit analysis to prioritize the risks and design artifacts to manage the risks. Visualization is not used.

A second class of related work applies privacy risk analysis to specific application areas. Biega et al. [17] propose a new privacy model to help users manage privacy risks in their Internet search histories. They assume a powerful adversary who makes informed probabilistic inferences about sensitive data in search histories and aim for a tool that simulates the adversary, predicts privacy risks, and guides the user. Paintsil [18] presents an extended misuse case model and a tool that can be used to check the presence of known misuse cases and their effect on security and privacy risks in identity management systems. Das and Zhang [19] propose new design principles to lessen privacy risks in health databases due to aggregate disclosure. None of these works employ visualization.

A third class of related work is of course the work on privacy impact analysis (PIA) [5] (Section I). There are also works that support PIA. Meis and Heisel [20] present a method with tool support, based on a requirements model, that facilitates the PIA process. Tancock et al. [21] describe plans for a PIA tool that can be employed in a cloud environment to identify potential privacy risks and compliance. Joyee De and Le Métayer [22] present a Privacy Risk Analysis Methodology (PRIAM) for conducting privacy risk analysis in a systematic and traceable way, suitable for application in a PIA.

A fourth class of related work consists of security and privacy threat analysis, e.g., Nematzadeh and Camp [23]. Security and privacy threats are related risks. For example, a Trojan horse attack (security threat) can lead directly to the lost of private data (privacy threat). These works also do not use visualization as described here.

A fifth class of related work concerns earlier work on privacy visualization by this author. Yee [24] presents a notation for representing the software and hardware components of a computer system as well as the data flows between the components. It then checks each component for vulnerabilities that could violate a privacy policy. It differs from this work in terms of the notation (lower level than this work), the method of identifying vulnerabilities, and the use of privacy policies. Yee [25] featured the first use of the PIM but for web services only and involved privacy policies. In this work, we have extended the PIM to software systems in general and removed the need to work with privacy policies.

A sixth class of related work also involves visualization of risks but with different goals than in this work. They are works on the visualization of information intended to assist the decision making process under risk or improve the understanding of system security and risks. They differ from this work as follows: a) they concern the visualization of *security* risks rather than privacy risks, b) their goals are to assist in decision making or improve security understanding, whereas the goal of this work is to identify privacy vulnerabilities, and c) their visualizations are lower level in general and resemble more the objects being visualized, whereas this work uses a high level more abstract visualization. Three works representative of this class are Daradkeh [26], Takahashi et al. [27], and Kai et al. [28]. Daradkeh evaluates an information visualization tool for the support of decision making under uncertainty and risk. Takahashi et al. discuss the architecture of a tool for security risk visualization and alerting to increase security awareness. Kai et al. present a security visualization system for cloud computing that displays security levels computed over information gathered at monitoring points. Their visualization system is similar to visualizations provided by a security information and event management system (SIEM) [29].

A seventh class of related works deals with privacy by design. Guerriero et al. [30] provide a tool prototype to assist the process of continuous architecting of data intensive applications for the purpose of offering privacy by design guarantees. They also present a research roadmap for ensuring privacy by design for Big Data DevOps. Spiekermann [31] writes about the challenges of privacy by design. Le Métayer [32] presents a formal framework for use in the design phase of privacy by design, which checks if an architecture meets the requirements, including privacy requirements, of the parties involved with a system. Perera et al. [33] offer a conceptual framework with guidelines that employ privacy by design principles to direct software engineers in systematically assessing the privacy capabilities of Internet of Things applications and platforms.

Finally, no references were found that deals directly with the prioritization of privacy risks. However, abundant work exists on the assessment of security risks, which is closely related to prioritizing privacy risks. Alizadeh and Zannone [34] present a risk-based framework that facilitates the analysis of business process executions. The framework detects non-conforming process behaviors and ranks them according to criticality, which is determined by the execution's impact on organizational goals. The criticality ranking enables a security analyst to prioritize the most severe incidents. Jorgensen et al. [35] propose decomposing risk associated with a mobile application into several risk types that are more easily understood by the application's users and that a mid-level risk summary be presented that is made up of the dimensions of personal information privacy, monetary risk, device availability/stability risk, and data integrity risk. Their work suggests that privacy risk prioritization, as in this work, may be facilitated by decomposing the risks into more easily understandable categories or dimensions. Islam et al. [36] present a framework for threat analysis and risk assessment of automotive embedded systems to systematically tackle security risks and determine security impact levels. The latter serve to prioritize the severity of the risks. The framework aligns with several industrial standards.

VII. CONCLUSION AND FUTURE WORK

This work has proposed a straightforward method for visualizing and prioritizing privacy risks applicable to software systems, based on locations involving PII. Such locations are important for risk evaluation because they represent varying levels of vulnerabilities or risks, and they contribute to total risks. Although the approach has weaknesses, the weaknesses can be remedied, as described in Section V.

Future work includes the automations and improvements to the method for risk prioritization mentioned in Section V, along with a validation of the effectiveness of the approach. For this validation, it is envisioned that a software system with known privacy risks and prioritization (reference risks and prioritization) would be defined to act as the reference system. Different teams of privacy and security experts who do not have prior knowledge of the reference risks and prioritization would then be invited to apply the approach to the reference system. Their results would be compared to the reference risks and prioritization to gage the effectiveness of the approach. If the effectiveness was found to be inadequate, a follow-up analysis could point to the reasons for the discrepancy and could give insight into ways to improve the approach.

REFERENCES

[1] G. Yee, "Visualization of Privacy Risks in Software Systems," Proceedings of the Tenth International Conference on Emerging Security Information, Systems and Technologies (SECURWARE 2016), pp. 289-294, 2016.

[2] V. S. Iyengar, "Transforming Data to Satisfy Privacy Constraints," Proceedings of the 8th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'02), Edmonton, Alberta, pp. 279-288, 2002.

[3] R. Song, L. Korba, and G. Yee, "Pseudonym Technology for E-Services," chapter in Privacy Protection for E-Services, edited by G. Yee, Idea Group, Inc., 2006.

[4] C. Adams and K. Barbieri, "Privacy Enforcement in E-Services Environments," chapter in Privacy Protection for E-Services, edited by G. Yee, Idea Group, Inc., 2006.

[5] Treasury Board of Canada Secretariat, "Directive on Privacy Impact Assessment," available on March 27, 2016 at: http://www.tbs-sct.gc.ca/pol/doc-eng.aspx?id=18308

[6] I. Goldberg, D. Wagner, and E. Brewer, "Privacy-Enhancing Technologies for the Internet," IEEE COMPCON'97, pp. 103-109, 1997.

[7] CIPP Guide, "CSA Model Code," available on Feb. 22, 2017 at: https://www.cippguide.org/2010/06/29/csa-model-code/

[8] G. Yee, L. Korba, and R. Song, "Legislative Bases for Personal Privacy Policy Specification," chapter in Privacy Protection for E-Services, edited by G. Yee, Idea Group, Inc., 2006.

[9] O. Tene, "Privacy: The New Generations," International Data Privacy Law, Vol. 1, Issue 1, pp. 15-27, February 2011. Available on May 31, 2017 at: https://academic.oup.com/idpl/article-lookup/doi/10.1093/idpl/ipq003

[10] G. Kambourakis, "Anonymity and Closely Related Terms in the Cyberspace: An Analysis by Example," Journal of Information Security and Applications, Vol. 19, Issue 1, pp. 2-17, Elsevier, February 2014.

[11] A. Ruiz-Martinez, "A Survey on Solutions and Main Free Tools for Privacy Enhancing Web Communications," Journal of Network and Computer Applications, Vol. 35, Issue 5, pp. 1473-1492, Elsevier, September 2012.

[12] J. Ren and J. Wu, "Survey on Anonymous Communications in Computer Networks," Computer Communications, Vol. 33, Issue 4, pp. 420-431, Elsevier, March 2010.

[13] A. Pfitzmann and M. Hansen, "A Terminology for Talking About Privacy by Data Minimization: Anonymity, Unlinkability, Undetectability, Unobservability, Pseudonymity, and Identity Management," Version v0.34, 98 pages, Aug. 10, 2010. Available on May 31, 2017 at: https://dud.inf.tu-dresden.de/literatur/Anon_Terminology_v0.34.pdf

[14] C. Salter, O. S. Saydjari, B. Schneier, and J. Wallner, "Toward A Secure System Engineering Methodology," Proceedings of the New Security Paradigms Workshop, pp. 2-10, 1998.

[15] A. Nanthaamornphong, K. Morris, and S. Filippone, "Extracting UML Class Diagrams from Object-Oriented Fortran: ForUML," Proceedings of the 1st International Workshop on Software Engineering for High Performance Computing in Computational Science and Engineering (SE-HPCCSE'13), pp. 9-16, 2013.

[16] J. I. Hong, J. D. Ng, S. Lederer, and J. A. Landay, "Privacy Risk Models for Designing Privacy-Sensitive Ubiquitous Computing Systems," Proceedings, 2004 Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques, Cambridge, MA, USA, pp. 91-100, 2004.

[17] J. Biega, I. Mele, and G. Weikum, "Probabilistic Prediction of Privacy Risks in User Search Histories," Proceedings of the 1st International Workshop on Privacy and Security of Big Data, pp. 29-36, Nov. 2014.

[18] E. Paintsil, "A Model for Privacy and Security Risks Analysis," Proceedings of the 5th International Conference on New Technologies, Mobility and Security (NTMS), pp. 1-8, May 2012.

[19] 14-G. Das and N. Zhang, "Privacy Risks in Health Databases From Aggregate Disclosure," Proceedings of the 2nd ACM International Conference on Pervasive Technologies Related to Assistive Environments (PETRA'09), article no. 74, June 2009.

[20] R. Meis and M. Heisel, "Supporting Privacy Impact Assessments Using Problem-Based Privacy Analysis (Technical Report)." Available on May 31, 2017 at: https://www.uni-due.de/imperia/md/content/swe/pia-formal.pdf

[21] D. Tancock, S. Pearson, and A. Charlesworth, "A Privacy Impact Assessment Tool for Cloud Computing," Proceedings of the 2nd IEEE International Conference on Cloud Computing Technology and Science, pp. 667-676, 2010. Available on May 30, 2017 at: http://barbie.uta.edu/~hdfeng/CloudComputing/cc/cc47.pdf

[22] S. Joyee De and D. Le Métayer, "PRIAM: A Privacy Risk Analysis Methodology," Research Report RR-8876, Inria - Research Centre Genoble - Rhône-Alpes, 51 pages, 2016. Available on May 31, 2017 at: https://hal.inria.fr/hal-01302541/document

[23] A. Nematzadeh and L. J. Camp, "Threat Analysis of Online Health Information System," Proceedings of the 3rd International Conference on Pervasive Technologies Related to Assistive Environments (PETRA'10), article no. 31, June 2010.

[24] G. Yee, "Visualization for Privacy Compliance," Proceedings of the 3rd International Workshop on Visualization for Computer Security (VizSEC'06), pp. 117-122, Nov. 2006.

[25] G. Yee, "Visual Analysis of Privacy Risks in Web Services," Proceedings of the IEEE International Conference on Web Services (ICWS 2007), pp. 671-678, July 2007.

[26] M. Daradkeh, “Exploring the Use of an Information Visualization Tool for Decision Support under Uncertainty and Risk,” Proceedings of the International Conference on Engineering & MIS 2015 (ICEMIS’15), article no. 41, 2015.

[27] T. Takahashi, K. Emura, A. Kanaoka, S. Matsuo, and T. Minowa, “Risk Visualization and Alerting System: Architecture and Proof-of-Concept Implementation,” Proceedings of the First International Workshop on Security in Embedded Systems and Smartphones (SESP’13), pp. 3-10, 2013.

[28] S. Kai, T. Shigemoto, T. Kito, S. Takemoto, and T. Kaji, “Development of Qualification of Security Status Suitable for Cloud Computing System,” Proceedings of the 4th International Workshop on Security Measurements and Metrics (MetriSec’12), pp. 17-24, 2012.

[29] Wikipedia, “Security information and event management,” available on June 12, 2016 at: https://en.wikipedia.org/wiki/Security_information_and_event_management

[30] M. Guerriero, D. Tamburri, Y. Ridene, F. Marconi, M. Bersani, and M. Artac, “Towards DevOps for Privacy-by-Design in Data-Intensive Applications: A Research Roadmap,” Proceedings of the 8th ACM/SPEC on International Conference on Performance Engineering Companion (ICPE ’17), pp. 139-144, April 2017.

[31] S. Spiekermann, “The Challenges of Privacy by Design,” Communications of the ACM, Vol. 55, Issue 7, pp. 38-40, July 2012.

[32] D. Le Métayer, “Privacy by Design: A Formal Framework for the Analysis of Architectural Choices,” Proceedings of the 3rd ACM Conference on Data and Application Security and Privacy (CODASPY ’13), pp. 95-104, 2013.

[33] C. Perera, C. McCormick, A. Bandara, B. Price, and B. Nuseibeh, “Privacy-by-Design Framework for Assessing Internet of Things Applications and Platforms,” Proceedings of the 6th International Conference on the Internet of Things (IoT ’16), pp. 83-92, November 2016.

[34] M. Alizadeh and N. Zannone, “Risk-based Analysis of Business Process Executions,” Proceedings of the 6th ACM Conference on Data and Application Security and Privacy (CODASPY’16), pp. 130-132, 2016.

[35] Z. Jorgensen, J. Chen, C. Gates, N. Li, R. Proctor, and T. Yu, “Dimensions of Risk in Mobile Applications: A User Study,” Proceedings of the 5th ACM Conference on Data and Application Security and Privacy (CODASPY’15), pp. 49-60, 2015.

[36] M. Islam, A. Lautenbach, C. Sandberg, and T. Olovsson, “A Risk Assessment Framework for Automotive Embedded Systems,” Proceedings of the 2nd ACM International Workshop on Cyber-Physical System Security (CPSS’16), pp. 3-14, 2016.