RC22912 (W0309-132) September 22, 2003
Computer Science

IBM Research Report

Views of Privacy: Business Drivers, Strategy and Directions

Clare-Marie Karat, Carolyn Brodie, John Karat
IBM Research Division
Thomas J. Watson Research Center
P.O. Box 704
Yorktown Heights, NY 10598

Research Division
Almaden - Austin - Beijing - Haifa - India - T. J. Watson - Tokyo - Zurich

1.0 Executive Summary

Privacy legislation and the growing public awareness regarding privacy issues are contributing to the importance of implementing policies to deal with privacy issues for organizations that collect and store Personally Identifiable Information (PII) data about their customers and constituents and / or their employees.   In many geographies and industries, legislation is either in place or being considered which regulates the use of PII. Likewise, there have been an increasing number of news stories detailing cases in which PII data has been exposed either because of accidental lapses in privacy policy enforcement or because of malicious actions on the part of a few individuals. For this reason, many organizations are now scrambling to ensure that the PII data with which they are entrusted is protected and handled properly. The purpose of this work is to create an understanding of the business drivers, strategies and anticipated future privacy needs of organizations that collect and use PII data. The results of this work will inform the future direction of IBM products and services.

To accomplish this goal, the user-experience based privacy research team at IBM's T. J. Watson Research Center has completed a series of interviews with 13 representatives from business and government organizations in North America and Europe who are concerned with privacy issues. Each participant was sent a pre-session questionnaire that they were asked to complete before the interview. The respondents were then each asked to participate in a one-hour, structured interview.
Quantitative data were collected from the pre-session questionnaire. The data collected includes participants:

1. top privacy concerns regarding their businesses, and
2. the top three types of privacy functionality they would like to have available to address their privacy concerns regarding their businesses.

In regard to respondent concerns related to privacy, participants mentioned "The economic harm that would result to this company if a privacy breach regarding customer data became public." and "Keeping internal employees from violating the privacy of others' data." as their top two concerns. Regarding the ranking of functionality seen as important to addressing the respondents privacy concerns, participants selected "One integrated solution for all legacy and Web data." as their top choice and "The ability to associate privacy policy information with individual data elements in a customer's file." as their second choice. Two other choices (i.e., "Privacy protection for data stored on servers from IT staff with no need to view data content" and "Application-specific privacy policy authoring, implementation, auditing, and enforcement") also received strong support.

The analysis of the qualitative data gathered from the interviews provided a rich picture of business drivers related to privacy and privacy's relation to security, business process efficiency, personalization, customer relationship management, and trust. Participants identified legislative regulations and contractual obligations, improved customer service, risk mitigation of brand damage, and more efficient and effective IT infrastructure as issues driving them to create and implement privacy policies. A number of the

participants explained that their requirements for privacy solutions grow from their need to offer great services and return value to their share-holders. To summarize this flow, participants felt that they need to use PII data in order to provide high quality services to their customers. However, the use of this data forces them to comply with privacy regulations in the jurisdictions in which they operate. It also leaves them vulnerable to security breaches that could harm their reputations with their clients. For these reasons, they are adopting strategies such as the creation and implementation of privacy policies within their organizations. Participants are involved in data classification projects and business process data flows related to PII data. They describe the major risks as brand damage, misuse of PII data by internal staff, larger scale breaches related to PII data due to automation and consolidation of data, and the physical security of PII data. Finally, the participants reported that they do believe that there are advantages to implementing privacy solutions that will save them money in addition to mitigating the risks involved with the use of PII data. These advantages include removing redundancies in business processes, more effective use of IT infrastructure, employing privacy as a competitive differentiator, and facilitating the flow of information across boundaries. During the remainder of this year, the user-experience based privacy research team will investigate the issues raised during these interviews.

2.0 Goals and Approach for Interviews on Privacy

Privacy has become a very important issue for any organization that collects and stores Personally Identifiable Information (PII) data about their customers and constituents and / or their employees. Both legislation and public opinion have contributed to this new emphasis. In many geographies and industries, legislation is either in place or being considered which regulates the use of PII. Likewise, there have been an increasing number of news stories detailing cases in which PII data has been exposed either because of accidental lapses in privacy policy enforcement or because of malicious actions on the part of a few individuals. For this reason, many organizations are now scrambling to ensure that the PII data with which they are entrusted is protected and handled properly. The purpose of this work is to create an understanding of both the status of privacy efforts underway and the future privacy needs of organizations that collect and use PII data. The results of this work will inform the future direction of IBM products and services.

2.1 Interview Goals

The goals of the interviews were to understand the top-priority privacy concerns for both businesses from a wide industries and government, the privacy functionality needed to address those concerns, the value that this privacy technology would have to their organizations, scenarios of the flow of Personally Identifiable Information (PII) as it passes through the organization from the time it is collected until the organization disposes of it, the strengths and weaknesses of these current business processes with PII data, the additional privacy functionality they need, and identification of significant challenges for research in this area to create desired privacy technology and solutions for the future.

2.2 Approach

The user-experience based privacy research team interviewed 13 representatives from business and government organizations in North America and Europe who are concerned with privacy issues. Participants were recruited through a variety of mechanisms including follow-ups on attendance at privacy break-out sessions at IBM-sponsored conferences, referrals from stakeholders, and out-reach to organizations identified as having a strong privacy interest. The participants were chosen based on stated interest in the topic of privacy and because they represented industries and geographies where privacy legislation has taken effect. The sample of participants was also drawn to provide representation from three major types of responsibilities including chief privacy officer, organizational management, and IT privacy and security roles. All of the interviews were completed by telephone. Each interview lasted approximately 45 to 60 minutes. All interviews were conducted with a lead interviewer and a colleague who served as a second listener. The interviews consisted of seven open-ended questions. Follow-up questions were asked during the discussion of each question. The interviewees took written notes that were transcribed within 24 hours and reviewed by the team for completeness and accuracy. All interviewees were promised that all data would be kept confidential and only de-identified and summary results would be reported. The 13 interviews were completed during February and March of 2003. Data analysis of the qualitative, unstructured information was completed using an adapted and extended version of the Contextual Design method for affinity diagramming (Beyer and Holtzblatt, 1998). This report summarizes the results of the interview research. De-identified participant quotes are provided in the text for illustrative purposes and are shown in italics.

2.3 The Roles and Industries of Participants Interviewed

The participants whom we interviewed came from the following industries:

Industry	**Number of Interviewees**
Banking	3
Communications	1
Entertainment	1
Finance/Investment	1
Government	2
Healthcare	2
Health Insurance/HR	1
Travel	2
Total	13

Table 1. Interviews by Industry.

The participants included a mixture of individuals responsible for developing privacy and security strategies and policies; managing employees and business partners involved in business-critical processes involving PII data; designing the architecture and engineering the solutions for privacy; implementing policies through a mixture of education and training and manual and automated processes; and operationally sanctioning privacy systems.

Most of the businesses represented were multinational companies. Three of the thirteen individuals are part of organizations that were headquartered in Canada or Europe. The remainder of the participants belong to organizations were either multinational businesses or government organizations headquartered in the United States.

In this report we first present the results of the pre-interview data. We then discuss privacy-related concepts from the point of view of participants we interviewed. Next, we present the framework for a business case for privacy. We then discuss the visions of privacy technology expressed by the individuals we interviewed. We present a generalized view of the critical elements of a privacy solution as expressed by participants, both from a procedural and technology point of view. The report concludes with the questionnaire instruments and information on legislation on privacy.

3.0 Pre-Session Questionnaire Data

We asked the participants to answer three questions about their views on privacy before the interviews were conducted. The three questions were (1) "What are your top three privacy concerns regarding your business?", (2) "What are the top three types of privacy functionality you would like to have available to address your privacy concerns regarding your business?", and (3) " What action is your organization preparing to take to address the top privacy concerns you listed above?" For each question we provided a list of options to select and also provided "Other" options that the respondent could fill in. The questionnaire form is in the Appendix. During the interview, we began by asking the respondent follow-up questions based on their answers to the questions, and thus gained additional data for the interview analysis. The results are summarized in the figures below.

In question 1 in which we asked about top privacy concerns, the top issues raised included: "The economic harm that would result to this company if a privacy breach regarding customer data became public." and "Keeping internal employees from violating the privacy of others' data.". A range of concerns were mentioned in the “Other” category. While "Other” was the most frequent response, these responses did not reflect a single high-frequency concern.

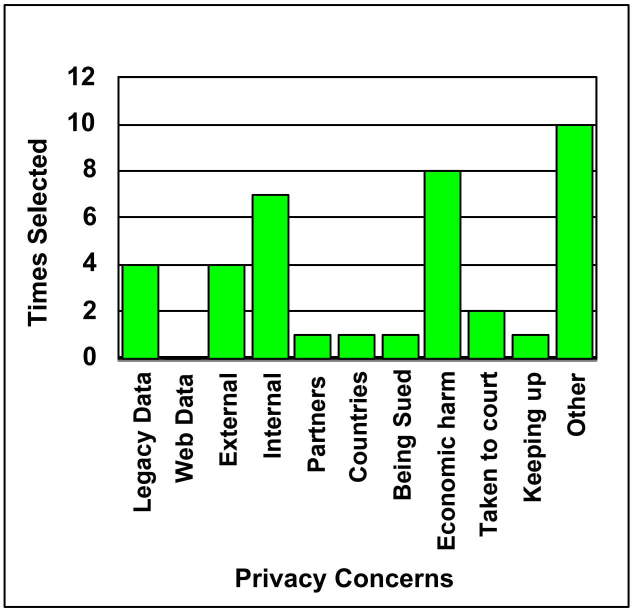

Figure 1. Top Privacy Concerns Expressed by Participants.

Below are the “Other” responses to Question 1:

- Managing web site/page compliance against Internet policies
- Architecting the management of customer preferences against business rules and regulatory needs
- The biggest issue is culture - need to constantly remind business side about need for security.
- Separate privacy issues into "secure messaging" and "protecting data stored in a database"
- The economic, political and confidence harm that would result if a privacy breach became public.
- Difficulties in implementing different laws from different jurisdictions.
- Ensuring that PI is appropriately managed within the organization.
- Leveraging the company's information assets for economic gain in a privacy sensitive manner.
- Internal employees violating the privacy of internal data.
- External agencies violating the privacy of internal data in research and other projects.

For Question 2, which addressed the functionality seen as important to addressing the respondents privacy concerns, participants selected "One integrated solution for all legacy and Web data." as their top choice and "The ability to associate privacy policy information with individual data elements in a customer's file." as their second choice. Two other choices (i.e., "Privacy protection for data stored on servers from IT staff with no need to view data content" and "Application-specific privacy policy authoring, implementation, auditing, and enforcement") also received strong support.

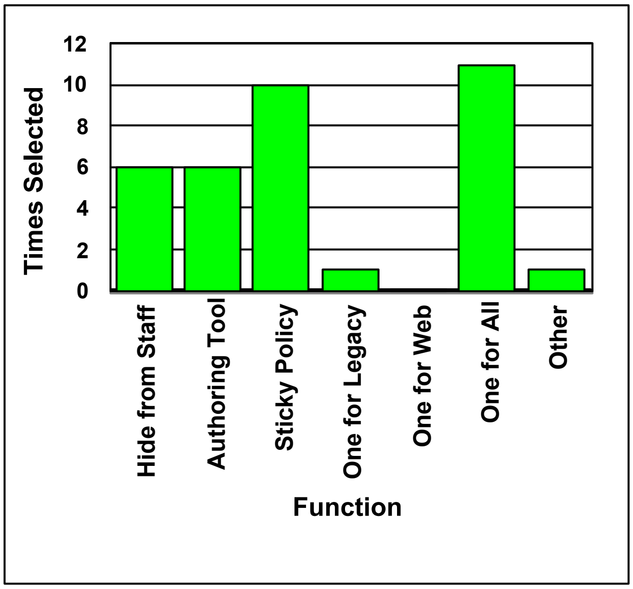

Figure 2. Participant Views of Top-Valued Privacy Functionality

The single “Other” response to Question 2 is listed below:

- Maintaining privacy controls over data in a data warehouse.

For Question 3, which asked about actions planned by the respondents organization regarding privacy, there were obviously different perspectives on the question (see Figure 3 below). In some cases all of our listed responses might apply - that is, parts of an organization might be done developing privacy functionality, while other parts had "no plans" or were working with "consultants". In general, almost all of the participant organizations we talked to had begun some level of technology development to support privacy. The specific actions that were being taken ranges from "developing architecture" to "building privacy protection in to individual applications". Participants were mixed in their plans to purchase or build their own solutions.

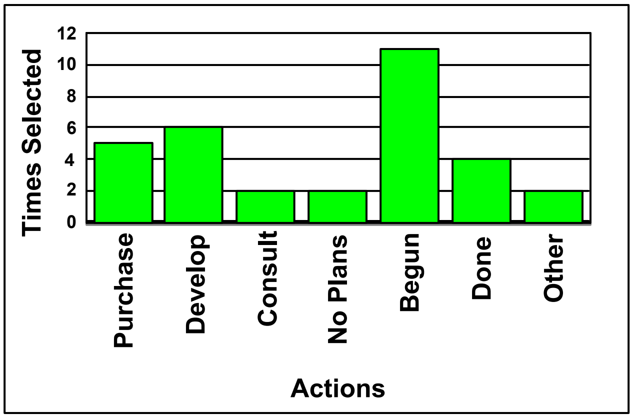

Figure 3. Participant Privacy Actions

Below are the “Other” responses to Question 3:

- Privacy architecture implementation. Expand use of privacy impact assessments.
- We are constantly monitoring where the security product market is going.

4.0 Privacy Concepts

Privacy in the context of computer infrastructure is a relatively new concept compared to areas like security. The rise of the Internet and growing awareness of the possibility of identity theft crimes have made the average person begin to worry about how personal data is stored and who has access to it. Because of this, many of the participants we talked with discussed how privacy related to other concepts such as security, personalization, trust, and education.

.

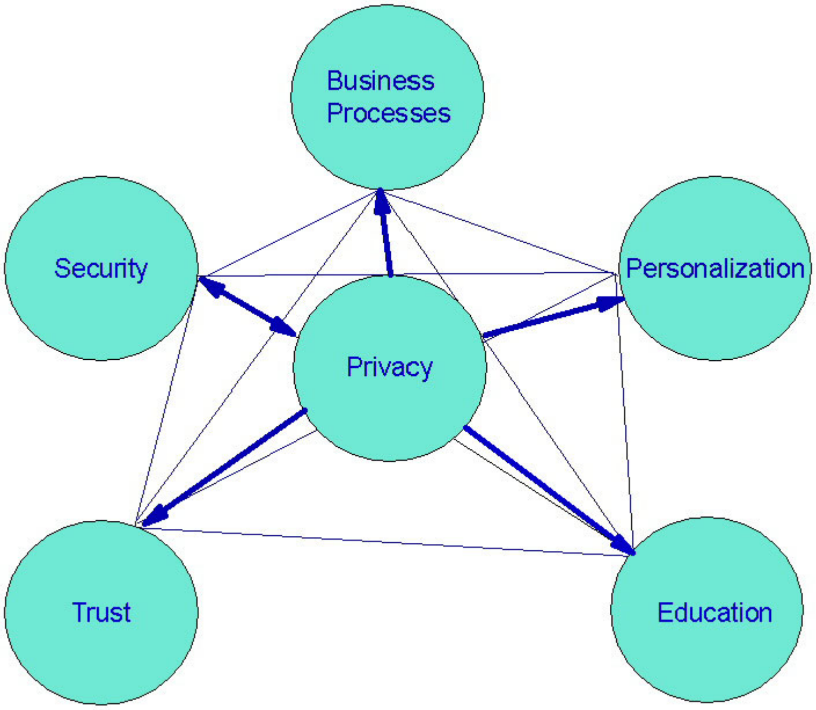

Figure 4. The Relationships between Privacy and Other Concepts.

Figure 4 shows how each of these concepts relates to privacy. It is important to note that these concepts are at different levels of maturity. For example, there has been work in the area of computer security and the development of business processes since the earliest days of business computing, but privacy, personalization, and trust from an e-business perspective are relatively new areas of research and development. The emergence of these new areas has created the need for education around them

Many of the participants that we interviewed discussed how privacy is related to many of these concepts. These individuals felt that it was necessary to explain how they saw privacy fitting into a larger framework (e.g. IT, Customer Relationship Management (CRM), marketing). From an IT perspective they felt privacy is related to security and the definition of business processes. From a CRM perspective, they felt privacy was related to personalization. From a marketing perspective, they felt privacy was related to trust and education. Each of these relationships will be explored below. For the purposes of this report, we define each of these concepts as it relates to computer infrastructure below:

1. **Security** can be defined in terms of the services it provides. These are as follows: providing *access control* to data and computer resources, *identifying and authenticating* of users of those resources, protecting the *confidentiality* of data by ensuring that it cannot be accessed by unauthorized individuals or

applications, protecting the *integrity* of data from unauthorized modification, and providing *non-repudiation* guarantees (Baker, Beere, Bogardus, Jongvattanasiri, and Seeldrayers, 2001).

2. **Privacy** depends on the security services listed above, but it adds the concept of purpose – privacy demands that we define for what purposes data can be used by particular individuals and applications or individuals and applications with certain roles.
4. **Business Processes** are the manual or automated procedures and set of steps or activities by which an organization completes necessary tasks.
5. **Personalization** is the use of information about an individual in order to tailor the experience of that individual within the context of the business or organization using the data. Personalization is often used in the context of an individual's experience on a website, however it is not necessarily limited to this use (Karat, Karat, Brodie, Vergo, and Alpert, 2003). It may be used for many purposes such as to determine what communications an individual receives from the organization through hardcopy mail.
6. **Trust** in the context of this report refers to the perception by an individual that an organization will not do anything with the PII data that the organization has about that individual that the individual did not intend.
7. **Education** in the context of this report refers to actions taken by participants to sensitize others in their organization to privacy issues.

4.1 Privacy and Security

The relationship that was most often mentioned by participants was between privacy and security. Although the idea was stated in several different ways, most of the participants we spoke with agreed, "*security is a mechanism to enforce privacy*". They all indicated that privacy overlaps security, but is not identical to it. As one person stated, "*Confidentiality is one piece of privacy (and security). Obtaining consent is not confidentiality, it is privacy*." Therefore, privacy and security overlap in the requirement that data be kept confidential, but in addition, privacy requires an organization to obtain permission for data to be used for a particular purpose.

4.2 Privacy and Business Practices

Many participants explained that the implementation of a privacy policy can affect business processes, however multiple ways in which business processes were affected were expressed. Some of the participants interviewed explained that they felt it was necessary to change their organization's business processes to protect privacy. As one participant stated, they were "*redesigning the business processes so that they don't result in the transfer of (PII) data* (unless it was determined to be necessary)".

Others indicated that they thought that the implementation of practices that ensure privacy benefited people in other business areas or their organizations as a whole. One participant explained, "*when tools and procedures for protecting customer privacy are put in place, the PII data about the customers is more accurate and available*." Therefore, business people can rely on the validity of the data. A number of the participants we spoke to told us that they saw the potential for cost savings by removing redundancies found while mapping data flows. In this case, the implementation of privacy protections was seen as a potential cost savings for the organization.

4.3 Privacy and Personalization

A number of the participants felt that personalization and privacy are "*definitely linked*." Two of the participants we interviewed discussed how their organizations collect PII data to provide personalized services to their customers. As one stated, "*We personalize everything we can. It is important to (our customers) to protect personalized data*."

4.4 Privacy and Trust

Many of the participants expressed concerned about trust. One respondent told us that her organization "*had a 185-year culture of confidentiality throughout the enterprise*." She went on to say that it was important to educate the organization about "*privacy, trust and confidentiality … as opposed to just confidentiality*."

4. 5 Privacy and Education

Because privacy as it relates to accessing data stored on computer systems is a relatively new concept, several participants told us that they spend a large percentage of their time educating others in their organizations about privacy in addition to developing new processes and installing new technologies. They told us that part of their job is to make sure that employees that handle PII data understand the organization's privacy policies and why they are important. One respondent summed up these ideas with the following statement. "*It is not so much of a technical issue as an ongoing educational issue – building sensitivity for business and IT people*." He continued, "*They have to become part of the culture and DNA of privacy. The human element drives human behavior. You can't solve this through technology alone*."

5.0 Business Drivers for Implementing Privacy Policies, Procedures and Technology

5.1 Current Participant Status from a Privacy Point of View

From our interviews, we learned that all of the participants we spoke with were concerned about protecting the PII data they had about their customers, constituents, and employees. However, it was also clear that they all understood that their organizations existed to provide services or products and for the businesses "*to return money to (their) shareholders*". The participants told us that they believe that collecting and using PII helps them to provide better services and therefore to make more money. Privacy is important to them because in many geographies and in many industries there are legal regulations that must be complied with and because it is important to their customers and constituents. One example of this idea came from a respondent who said, "*We want to make maximal use of data while in compliance with privacy legislation*." Similarly, another respondent explained, "*We offer great services which lead to customer loyalty and trust. It is appropriate to use (the PII) data we have to offer great services and make money, but the data is private. We don't violate (our customers') trust by selling it*."

Figure 5 (below) shows how participants view the value of implementing privacy policies, procedures and technology. To summarize this flow, the participants explained that they need to use PII data in order to provide high quality services to their customers. However, the use of this data forces them to comply with privacy regulations in the jurisdictions in which they operate. It also leaves them vulnerable to security breaches that could harm their reputation with their clients. For these reasons, they are adopting strategies such as the creation and implementation of privacy policies within their organizations. Finally, they do believe that there are advantages to implementing privacy solutions that will save them money in addition to mitigating the risks involved with the use of PII data. This flow will be explained in more detail in the rest of this section.

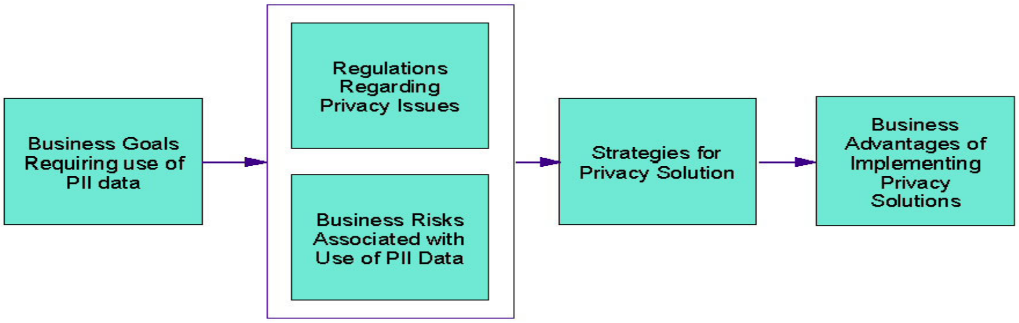
Business Goals Requiring use of PII data
Regulations Regarding Privacy Issues
Business Risks Associated with Use of PII Data
Strategies for Privacy Solution
Business Advantages of Implementing Privacy Solutions

Figure 5. Participant Motivations for Privacy.

5.1.1 Legislative requirements regarding privacy

International Organizations

Every participant told us that they were concerned with some or all of the privacy legislation that has been passed and may soon be passed for different geographies and industries. There is a wide discrepancy in laws around the world and the correct interpretation of the laws is not always clear. In some geographies and industries, "*any breach (in protecting privacy) is a breach of the law*." In other situations, there is almost no legal protection for PII data. Organizations that cross national boundaries must deal with a wide range of different and sometimes conflicting regulations. In fact, the majority of the concerns raised by participants are related to the complexity of dealing with privacy issues that cross national boundaries. Many of the respondents were from multi-national organizations. This means that they must have privacy policies that comply with the laws in each geography in which they operate. A participant comment that is representative of many of the concerns we heard is as follows, "*There are regulations from the European Union, Australia, Hong Kong, and new laws in Canada as of 1/1/2004. We want to be in regulatory compliance (with all of them)*." Sometimes this is very challenging. For example, although an agreement has been reached between the US and the EU, earlier this month a conflict between new US security regulations regarding airline passengers flying into the US and EU privacy laws developed. Airlines flying between the US and the EU were required by the US to turn over passenger data that potentially violated EU privacy regulations (Knight, 2003). The respondents we spoke with are clearly concerned about this type of issue. One expressed it well by saying that they have a desire for "*harmonization around the world on (privacy) regulations*."

US-only Organizations

Even participants who are from companies or organizations that currently only operate within the United States have cross-jurisdictional concerns. One respondent told us that although his company currently only does business in the US, they often acquired new businesses. He said, "*I care (about EU and Australian laws) simply because if they (company executives) came to me and said they were buying something in Europe or Australia, I would like to be able to say 'ok' – not –' ok we will not be sleeping for the next year and a half' (while working on compliance issues in these new geographies)*." Other organizations that do business only in the US worry about the difference in state laws (see www.caprivacyprotection.org/laws_california.html). One participant told us, *"The problem in the US is that different jurisdictions are putting in place different privacy laws that will be difficult to implement."* Additionally, participants are concerned about changing laws. For example, one respondent was concerned about one of the new California privacy laws that states that "*users must be notified when we even suspect a leak*." These concerns are driving participants to look at technological solutions. As one participant told us, "*We need tools that have built-in functionality to recognize any accidental disclosure or change to data – any unusual attempts to access data*."

Move from Opt-out policies to Opt-in policies
A few participants whose organizations have already begun to implement new tools and processes to ensure privacy discussed up-coming legal changes that would cause some industries to ask their customers to opt-in to personalization and rewards programs rather than opt-out. One participant told us that he felt that in the future "*any use of PII data would have to be approved by the customer*". The move towards opt-in policies has been driven in part by criticism of opt-out policies. Critics point out that because many users do not read the privacy policies they don't know they have the option of opting-out, and even people who do know they have the option of opting-out sometimes choose not to because of the time and effort required. However, changing from opt-out policies to opt-in policies may have unintended human-computer interaction consequences. Recent studies involving the creation of privacy profiles for the AT&T Privacy Bird have shown that it is easier for individuals to indicate what they don't find acceptable rather than what they do think is acceptable regarding the privacy of their PII data (Cranor, 2003). It may be that easily recognized and used opt-out policies are more effective. Additional research is required to determine the best solution.

5.1.2 Current Strategies for Privacy

With the increasing requirements from new legal regulations in many geographies and industries and the recognition that their customers are concerned about the privacy of the data they share, the participants we interviewed have adopted many strategies for addressing privacy concerns in their businesses.

Creation of Organizational Privacy Policy
Many of the respondents told us that creating an organizational privacy policy or set of policies was the first step in addressing privacy issues. Although most of the organizations to which participants in this study belong already have a privacy policy, a few told us that they do not feel that their organizations current privacy policies are adequate or that the members of their organizations are sensitive enough to privacy issues. Several participants told us that while their organizations did not currently have a formal CPO, they thought that one was needed and would be hired soon. In several cases, the individual we spoke with told us that although privacy was not officially their job, they had become the "*privacy evangelist*" for their organization. These privacy evangelists told us that they work with the legal department and the business areas to interpret the laws for their organizations. One participant described the "*privacy committee*" that consists of technical security people and lawyers that was used within his organization to address privacy issues. He told us that it worked well for his organization.

Creation and Documentation of Business Processes to Implement Privacy Policy
Once the organization has an enterprise-wide privacy policy, the next step is to create business processes and automated tools to ensure that policy is enforced. One respondent summed this up, "*We meet privacy needs with technology and procedural controls*." However, in many of the interviews we conducted, it was clear that although there was a desire for technology to help enforce privacy, many organizations now using mainly manual procedures as a first step. One participant told us that given the deadlines for

compliance with parts of the Graham-Leach-Bliley regulations, they were now concentrating on "*manual processes, policies and documenting both of these"*. He went on to say that they were also putting in place risk mitigation plans for areas in which they knew they were weak. Finally, one participant we spoke with explained that her organization felt that "(*although) sometimes there is no control over privacy breaches, having a corporate policy, educating employees about that policy, and having manual and automated safeguards would make breaches less frequent*."

Data Classification

While some participants told us that they are working on manual processes at this point, others reported that they are making use of automated tools or are re-working their systems so that they will be in a position to make use of automated tools. Some of the respondents we spoke with discussed their efforts to create and apply data classifications to all of the data they store about people. These classifications will be tied to privacy policies so that anytime the data is accessed, it will be clear what privacy rules apply to that data.

The strategy of using data classifications is designed to address some of the privacy concerns related to the use of enterprise-wide data warehouses. Several participants told us that as business processes for handling PII data within the organization have become more automated and the data is stored in a consolidated manner, the problem of protecting privacy has become more difficult. One participant told us that after the data was consolidated in a data warehouse, people who could not access a particular piece of data within the original silo could access it in the data warehouse and that this was a problem. Another told us that "*everyone and their cousin*" had access to data in the data warehouse and he wanted to understand why "*each person needed access*". Using a data classification strategy, organizations hope to control access to individual pieces of PII data by labeling it with a particular category and then tying access control roles to that category.

Privacy Data Flows

Another challenge when implementing a privacy policy is to identify all of the "touch points" for PII data. This requires that the organization understand how all PII data is collected, used, and destroyed. Participants told us they were working on data flows to document how PII flowed through the organization. Some participants explained that these had never been created before, but would be kept up to date from now on.

Customer Profiling

A number of the participants told us that they are working on profiles for their customers so that they can capture how each of their customers is willing to have their PII data used. One participant told us that when their new rules engine is complete, every business system will access the profile each time it is going to access PII data.

5.1.3 Technology Solutions in Use by Participant Organizations

In addition to business strategies, the participants we spoke with told us about some of the technologies that they have created and / or applied to the problem of enforcing privacy within their organization. It is important to note that only a small number of the respondents discussed particular applications with us by name. A number of the participants discussed technology at a higher level of abstraction. These will be discussed in section 6 of this document.

RACF

Three of the participants we spoke with told us that they use RACF to enforce security and privacy within their systems. They each mentioned aspects of RACF that they particularly liked. One told us that he used RACF for one-way encryption. Another of the participants discussed how he felt that access control was a much better way to protect privacy for his business rather than storing encrypted information. He went on to say that the RACF user registry worked well for employees, but not for customers because of the problem in registering very large numbers of individuals. The third participant praised RACF's built in logging and auditing functionality, especially as it concerned the actions by users with SPECIAL access. He liked the fact individuals who were given open access to the system could be held accountable for their actions.

Tivoli Privacy Manager

Tivoli Privacy Manager was mentioned by one participant we spoke to as something they have purchased and are considering rolling-out. He told us that they are experimenting with using Tivoli Privacy Manager as part of a new, overall On-Demand solution.

Data Retention

Many of the participants told us that although there are legal requirements regarding how long PII data is stored for each of their given industries, they store the data virtually forever. One respondent explained that the reason they stored data forever, was that the applications they used did not have an easy way to define retention dates when data should be deleted. He told us, "*most systems are data hogs*".

Data Collection Issues

One participant discussed issues regarding data collection as being a privacy concern. He pointed out that data can be collected in one of three ways: 1) explicitly by asking the customer; 2) it can be collected anonymously and then linked to other data to identify the data subject; or 3) it can be collected and used anonymously. There are additional privacy concerns related to the second method in that the individual does may not realize that PII data is being stored about them without their consent.

5.2 Risks regarding Privacy

All of the participants we spoke to discussed the privacy risks that concern them the most about their organizations. Although there are many more risks that could be raised, in this section we highlight the risks that the participants discussed during the interviews.

5.2.1 Brand Damage

The risk of harm to the brand name was one of the risks most often named by the participants we spoke with. The participants explained that customer trust and loyalty are important to their businesses and that if customers felt their trust was violated, they might not be loyal and might take their business elsewhere. Many respondents felt that this was a bigger risk than fines by government agencies for violating regulations. As one participant put it, "*Protecting privacy is important to preserving loyalty. This is an issue for multi-national companies. Customers expect it and they have a right to expect that we will protect their privacy*."

5.2.2 Protect PII Data from IT Staff

Another risk that was named by many respondents was the danger of exposing customer PII data to their IT staff who had no business need to see it. As one participant told us, "*Administrators don't need access to individual fields to understand what they need to know*." Another participant told us that he was concerned about the harm that an "*occasional disgruntled employee*" could do. To address these concerns, the possibility of encrypting data stored on servers was discussed, although at least one participant was concerned about the performance impact of this strategy. Other respondents discussed the need for logging. One wanted keyboard logging of everything that was typed into the server consoles so that he "would know what had happened." Along with logging, participants raised the issue of the lack of privacy auditing tools. Many participants wanted privacy auditing tools. Respondents told us that they do manual privacy audits currently but that "*(they) need automated tools to help review the logs and find process violations."*

5.2.3 Risks from Automation and Consolidation

Several participants told us that as business processes for handling PII data within the organization have become more automated and the data is stored in a consolidated manner, the problem of protecting privacy has become more difficult. One respondent told us that currently most PII data is handled manually and that the advantage of the manual process is that "*it is not privacy invasive*". However, there is clearly a move to consolidate data so that it can be used more efficiently by the organization even though this leads to privacy concerns. For example, one participant involved with a health care organization told us, "*From a treatment and care perspective the benefits of the electronic patient record are so huge that they outweigh the risks to privacy*." While everyone seems to agree that this is the direction most organizations are heading and a direction that has definite advantages for most businesses, the privacy concerns must be addressed. One person we interviewed summed this up by saying, "*Everyone is suffering from the same problem. Users love the data warehouse, but it is a little more open than we would like. Many people have access to the data and it has the least security*." Participants told us that they recognize both this business direction and the risks associated with it and are looking for ways to mitigate the risks.

5.3 Advantages of Implementing Privacy Solutions

A few of the participants we spoke to are starting to recognize that the implementation of privacy solutions has the potential to be a business advantage rather than just a way of preventing harm.

5.3.1 Remove Redundancies from Business Processes

The process of identifying data flows that is necessary to implement effective processes for a privacy policy can identify redundancies in the existing business process. Fixing these redundancies can lead to cost savings. As one participant said, *"Instead of collecting the same PII data several times in different applications as we used to, we now collect it just once. There's a cost saving here as well as better customer service".*

5.3.2 Privacy as a Differentiator

As organizations' customers come to recognize that some organizations work harder than others to protect their privacy, it can be a differentiator. One participant we spoke to told us that after considering that his company's brand value was based in part on customer loyalty and that was based in part on customer trust which was based in part on the protection of customer privacy, he "*had done some fourth grade math and came up with a number greater than zero for the value of privacy*."

5.3.3 Facilitate the Flow of Information Across Boundaries

Another area in which privacy technology can make a major impact on a range of organizations is by facilitating the flow of information across system and organizational boundaries. If technology can automate the flow of information across system boundaries, while preserving privacy policy information associated with data, there is a potential for greatly improved efficiencies in organizations. At the current time, participants are concerned about damage (both direct financial and to company image) that might be caused by breaches in privacy. This concern appears to encourage a cautious approach which leads to inefficient data transfer across systems. For example, in medical settings obtaining anonymous data related to a disease like cancer is extremely burdensome. Similarly, data is often not freely shared across an enterprise's organizational boundaries for fear of violating privacy policies that might actually allow such sharing.

6.0 Possible Privacy Solutions Identified

In the interviews, participants provided a view of some of the requirements for technology to become valuable in the management of data privacy. We interpreted their comments as addressing architectural directions for privacy technology (Section 6.2.1 below), along with a range of privacy functionality enablers.

6.1 Architectural Directions

Our analysis of the participant interview data produced five key concepts. These were:

1. Importance of one integrated solution
2. Separation of privacy functionality from application code
3. Appropriate granularity for applying privacy policy
4. Simple and flexible privacy functionality
5. Ability to work with structured and unstructured information

We consider each of these in turn in the sections below.

6.1.1 Importance of an Integrated Solution

Participants expressed a great desire to have privacy technology that worked consistently across a number of platforms and systems. Most of the participants we interviewed came from organizations that included both legacy and web-based data. They felt that separate solutions and approaches would only contribute to additional costs, higher risks, and decreased efficiency realized through privacy technology. One respondent said, *"The value of one integrated solution is in the consistency across the enterprise and its effectiveness."* Another noted the desirability of having the enterprise look like a single entity, rather than like a collection of unrelated silos. They noted that privacy concerns were ultimately not bound by the channel.

6.1.2 Privacy Functionality Separated from the Application Layer

"We want to have privacy and security functionality separated from applications so that there are not cases where some applications are more stringent and some are less stringent in their enforcement of security and privacy policy."

The participants in this study have generally done some amount of development of technology to implement privacy functionality. A number of them mentioned that this experience had led them to the belief that privacy functionality – like security functionality – should not be left to application control. Some were rewriting applications to remove privacy related functions, while others were working on architectures that placed security and privacy outside of the application layer. For cost and consistency reasons, participants felt that they needed to have functionality that worked across applications and systems in a centralized manner. One said, *"We don't want to have to change our applications every time regulations or our business processes change."*

Taking this a step further, participants also mentioned the need for automated tools for managing the privacy related flow of information – from creation to removal of the data from a system. Clearly ease of privacy policy development and administration are concerns. Means of easily authoring and applying privacy policies across the enterprise are a major concern for respondents to whom we spoke.

6.1.3 Appropriate Granularity for Applying Privacy Policy

Participants seemed to strongly favor technology that would enable privacy policy information to be associated with specific data items (the field-level in a database was the commonly mentioned level of desired granularity) so that the policy could remain associated with the data as it moves through the enterprise. As one participant noted, *"The personal record can't be treated as a whole. I would like to have the privacy policy attached to the data elements so that everyone who accessed the data would understand the privacy policy associated with it."* This requirement does not only apply to data in business silo applications, but includes data stored in a data warehouse. A number of respondents were concerned that there were different levels of security and privacy associated with data warehouses than with other data stores.

6.1.4 Simple and Flexible Solution

"I would hope that privacy solutions would be easy to use."

Participants mentioned that from a business point of view, the simpler the better. They mentioned a range of function that they would like to be able to employ – including field-level encryption (a function seen as key in enhancing privacy protection involving internal staff to an organization). Some participants mentioned difficulties associated with maintaining rules-based systems, and preferred a privacy solution that either eliminated the need for a complex rules engine or provided one that business staff (rather than IT staff) could maintain. A number of participants explained that cost and error concerns were a factor in their thinking about complex technology. Simplicity and consistency were viewed as a potential product differentiator.

Since privacy policies are seen as relatively early in development, the participants felt that flexibility – the ability to make changes and adjustments in policies – was critical. In general, respondents did not feel that niche systems tuned to specific current regulations (such as HIPAA) would be very valuable because they would be quickly outdated.

6.1.5 Support for Both Structured and Unstructured Information

In general, the participants discussed privacy considerations for structured data stored in databases when discussing their privacy needs. They mentioned that access to field level data within an individual or organizational record was important, and that the technology should support this granularity. Clearly this is important, but structured data is not the only data source requirement for an integrated solution. Several participants made note of the fact that privacy technologies needed to address the identification and coding of data in sources other than databases. They mentioned that there are substantial amounts of data in unstructured text documents such as email messages, government documents, or patient health records. While it might seem straight forward to simply treat such documents as a single entity and apply a single policy to the whole document, there were needs and benefits mentioned for being able to automatically de-identify documents (removing or hiding PII). This was seen as having clear application in health care – where the difficulty of obtaining de-identified medical history data is currently seen has a major obstacle to effective medical care and research. It was also seen as potentially valuable wherever documents were stored which contained important data for analysis imbedded in free text that might also include PII. This presents a case for utilizing text analytic function in the development of privacy technology.

6.2 Privacy Functionality Enablers

There were a number of details that emerged in the discussion as critical to the success of privacy technology. These included:

1. Technology to apply privacy policy
2. Technology to track and audit use of PII data
3. Secure data transport and email
4. Data tagging and classification mechanism
5. Customer preference profiles
6. Access control and encryption
7. Identity management

We consider each of these in turn in the sections below.

6.2.1 Technology to Apply Privacy Policies

Participants were concerned about the privacy policy authoring process. One respondent summarized the ultimate goal as: *"It would be great if I could take an English statement of what we wanted to do and convert it to a working system."* Participants seem to need some help in understanding how to go from a legislative requirement (like HIPAA) to an implementation which appropriately provides a data user the right amount and level of data for an approved purpose, and monitored and audited access to it.

Participants also drew a distinction between privacy policy monitoring and enforcement – stating in several cases that they felt that the technology needed to move toward actual enforcement.

On a detail level, one participant noted that privacy policy management should be similar to access control management. Some participants felt that they already had access control that worked across systems, and that this was the way privacy control should behave. Participants were very clear about one requirement: that privacy policy should be able to move with data.

There were mixed feelings about implementing privacy policies as a series of rules. Some participants seemed to assume that such rules would be a necessary part of a privacy policy-authoring environment, and focused on making sure that they were easy to understand and write. One participant expressed the desire to have "*business people and not IT staff manage it."* Another expressed concern over the manageability of enterprise privacy policy through rules and preferred a system that was *"as static as possible."* In general, we observed some fairly complicated policy specification requirements. For example, one health care participant expressed a desire to have a system that was flexible enough to allow access to patient data to be highly context dependent – allowing access by certain staff in certain medical situations.

6.2.2 Privacy Policy Auditing and Tracking Tools

Fundamental to privacy technology is the ability to track data access. However, some participants mentioned that they felt this was currently a very weak part of their processes. One respondent said: *"The weakest part of our process is analyzing audit logs and inappropriate disclosures."* Part of what they seem to be asking for here is the ability to trust the system to increasingly "do the right thing". That is, it shouldn't just monitor access, but should prevent violation of policy. But beyond that there was an expression of the need for better tools to help in tracking and auditing privacy data – to remove some of the need to rely on a manual system of tracking that the participants felt was not really up to the requirements.

Respondents gave a broad picture of just what the tracking requirements ultimately were. One participant said *"I want to be able to understand from inception to grave how every bit of PII is collected, used, and destroyed."* This needs to be done in a way such that organizations do not have to devote a great deal of resource to the auditing task. The view was expressed that auditing should help in producing better business processes: meaning that audit tools should be useful in understanding how privacy policy might be better positioned within the business process. For example, periodic reviews of audits should help business people better understand changes necessary to seamlessly integrate privacy policy and business process.

6.2.3 Secure Data Transport and Email

A number of participants expressed the view that privacy technology had a big role to play in the movement of data between systems (including in applying privacy policies to email, which they saw as possibly containing information that should have the same attention as structured data or file transfer). Within a single system, participants tended to see privacy needs as being addressed by security technology. For enterprises in general though, there were concerns that nothing (short of only communicating with someone on the same email platform) really ensured privacy.

6.2.4 Need for Standard Data Tagging and Classification Mechanisms

If we are going to have systems that enable privacy polices to be associated with data elements, it seems to require that data have the ability to be tagged with a classification and/or additional information. Participants discussing enterprise data warehouses expressed a desire for such tagging to enable *"PII data to become invisible when someone looks at a record"*. Similarly, they described scenarios where PII data would not be allowed to flow beyond contexts where privacy policy could be maintained. Data tagging could be used for reasons other than tagging with privacy information of course. The real issue for participants seems to be that there is not a standard for associating data elements with privacy policy information. Until there are standards in this area, participants must create their own and work in a fragmented environment of different systems.

6.2.5 User Interface for Customer Preference Profiles

Fundamental to privacy technology is the need to allow individuals to create policies to be used with their data. One participant summarizes this by saying *"We need a way for customers to indicate who they want data to go to without having to call the help desk to do it."* This goes beyond simply talking about opt-in or opt-out conditions. At a high level, participants indicated that thy want to enable their customers to easily indicate *"how they want to be treated"* and to have this converted to appropriate privacy policy.

6.2.6 Access Control and Encryption

For the participants, privacy policy was largely about controlling who has access to information. In general they felt that access from outside the organization was adequately covered by security and access control mechanisms. There was concern that existing access control mechanisms might not be sufficient to handle the range of *"Need to know"* requests typical for large organizations, and that they might have to be augmented in order to handle Privacy policy. In short, there was a sense that more than role-based access was needed – that other elements of purpose and context might be necessary.

A primary mechanism in which technology was seen as having a role particularly for internal access is through providing easy encryption functionality. Some participants also felt that encryption was an important additional component of physical security (and privacy). They suggested that information should be encrypted, not just in transport, but also in storage for PII data.

6.2.7 Identity Management

Several participants mentioned problems associated with "multiple identities" for their individual or institutional customers. This can create difficulties in determining the data subject for information and relating a specified privacy policy to that information. For example, in health care, physicians can have 50-75 "*identities*" to cover all of the sub-organizations that they have to deal with. There were two topics discussed by participants; federated identity management and trusted third-parties as a data source. Federated Identity Management (FIM) involves managing an identity that might involve a number of components distributed across systems - is a topic that is receiving increased attention. Here the emphasis is on being able to bring the appropriate pieces of an identity together while respecting privacy policies. There is also some interest in having trusted third party for providing identity information (e.g., having someone authorized by a data subject to provide their data to others).

6.3 Conclusions and Directions

All of the participants to whom we spoke made it clear that they consider privacy to be an important issue to their organizations. Looking across the interviews we conducted, we identified a series of activities that organizations execute while addressing privacy concerns. Although the activities varied from organization-to-organization, there seemed to be general types of activities that are often addressed, and that these activities are often performed in the order shown below. Although the list is presented serially, the activities are re-visited as the organization's experience in dealing with privacy increases and as technological solutions are introduced.

1. Create privacy policies and education programs to sensitize all members of the organization to privacy concerns.
2. Create data flows to document the flow of PII data through the organization's business processes and develop manual procedures to operationalize the privacy policy.
3. Develop and / or acquire technology to automate the creation, monitoring, enforcement, and auditing of organizational privacy policies.

In this report, we have presented a summary of the privacy business drivers, strategies for dealing with privacy issues, and organizational and technical directions regarding privacy that we identified from the interviews. During the remainder of this year, the user-experience based privacy research team will investigate the issues raised during these interviews.

7.0 References

Baker, G., Beere, J., Bogardus, B., Jongvattanasiri, Seeldrayers, I. (2001). IBM Host Integration in a Secure Network: A Practical Approach. *IBM Redbook*, IBM.

Beyer, H. and Holtzblatt, K. (1998) *Contextual Design: Defining Customer-Centered Systems*. San Francisco: Morgan Kaufmann, 1998.

Cranor, L. (2003). User Interfaces for Privacy: Design and Evaluation of the AT&T Privacy Bird P3P User Agent. Presentation at the IBM TJ Watson Research Center, March 26, 2003.

Financial Services Privacy Coalition. (2003). www.caprivacyprotection.org/laws_california.html

IBM Privacy Technical Institute. (2003). Enterprise Privacy Architecture (EPA). http://www916.ibm.com/press/prnews.nsf/jan/229860B7DD77736085256A7D0054450A.

IBM Privacy Technology Institute. (2003). Enterprise Privacy Architecture Language (EPAL). http://www.zurich.ibm.com/security/enterprise-privacy/epal/.

Karat, C., Karat, J., Brodie, C., Vergo, J., and Alpert, S. Personalizing the user experience on ibm.com. *IBM Systems Journal*, Volume 42 #4, in press.

Knight, V. (2003). (2003) EU Airlines in Consumer Backlash over Data Privacy. *Dow Jones International News* , March 25th, 2003.

Moore, G. (1991) *Crossing the Chasm*. NY: Harper Business.

8.0 Appendix

8.1 Pre-Session Questionnaire Form

Thank you for agreeing to be interviewed on privacy on XXX xx, 2003 at x:xx. To help us prepare for the interview, we would appreciate it if you would answer the three questions below and return this note to us before your interview. As with the information that you provide us during the interview we will keep your answers confidential and only use them summarized along with the responses from the other people we are interviewing. We will provide you a summary of the results. If you have any questions or need to re-schedule the interview for any reason, please do not hesitate to call me (914-784-xxxx) or email me at xxxxxxx@us.ibm.com.

To return your answers to me, please click on "Reply with History Attached" or the equivalent function in your email software and type your answers right onto the form. Then simply send it to me.

1. What are your top three privacy concerns regarding your business? Please rank order your top three concerns in the list below by writing a 1, 2, and 3 next to the item listed. If a top concern is not listed below, please write it in "Other" and provide its rank order.

__ Protecting the privacy of legacy data from unauthorized review or use
__ Protecting the privacy of Web data from unauthorized review or use
__ Keeping external users from violating the privacy of others' data
__ Keeping internal employees from violating the privacy of others' data
__ Keeping business partners from violating the privacy of customer data
__ Violating the privacy laws of the different countries within which the organization does business
__ Being sued by individual customers for loss of privacy
__ The economic harm that would result to this company if a privacy breach regarding customer data became public
__ Being taken to court by a governmental organization for violating privacy laws
__ Staying up-to-date on changes in privacy laws around the world
__Other: (Please describe)
__Other: (Please describe)

2. Please place the numbers 1, 2, and 3 next to the top three types of privacy functionality you would like to have available to address your privacy concerns regarding your business.

__ Privacy protection for data stored on servers from IT staff with no need to view data content through:
Key management for encryption to protect encryption keys.
Data-element level encryption of data
Other: ____________________________________
__ Application-specific privacy policy authoring, implementation, auditing, and enforcement
__ The ability to associate privacy policy information with individual data elements in a customer's file.
__ One integrated solution for all legacy data
__ One integrated solution for all Web data
__ One integrated solution for all legacy and Web data
__ Other: (Please describe)
__ Other: (Please describe)

3. At this time, what action is your organization preparing to take to address the top privacy concerns you listed above? Check all that apply.

__ We will purchase a privacy solution to address these concerns.
__ We will develop a privacy solution in-house to address these concerns.

__ We are going to bring privacy consultants in to advise our organization on how to proceed.
__ We have made no plans to take action at this time.
__ We have begun to implement privacy solutions, but more work is needed.
__ We have already implemented privacy solutions that address our concerns.
__ Other: (Please describe)

8.2 Interview Questionnaire Form

Thank you for agreeing to talk with us today. We would like to understand your perspective on protecting the privacy of your customers' personally identifiable information (PII) within your organization and on your systems. Your answers to our questions will be kept absolutely confidential and the data will be summarized across the small group of participants we are talking to. We will provide you with a summary of the results.

1. We would like to follow-up on a couple of your pre-interview questions.

2. How would you describe your role regarding privacy in your organization?

3. Can you give us an example scenario of what happens to a piece of Personally Identifiable Information (PII) as it passes through your business from the time it is first collected until you dispose of it? We will use your answer as the context for a few additional questions.

4. What are the strengths and weaknesses of your company's current processes (manual or automated) and tools regarding privacy in the example scenario you described?

5. What additional privacy functionality does your business need for your example scenario and how would you like this privacy functionality to fit into your business process?

6. Are there different privacy issues for Web and Legacy data?

7. Do you have any other privacy requirements that you would like to tell us about?

8.3 Privacy Legislation Mentioned in Interviews

Europe
EU Data Protection Directive - http://www.privacy.org/pi/intl_orgs/ec/eudp.html
Australia
http://www.privacy.gov.au/
Canada
Personal Information Protection and Electronic Documents Act - http://www.privcom.gc.ca/legislation/index_e.asp

United States - http://www.ftc.gov/os/statutes/fcra.htm

- Fair Credit Reporting Act - http://www.ftc.gov/os/statutes/fcra.htm
- Gramm-Leach-Bliley Act - http://www.ftc.gov/privacy/glbact/
- Health Insurance Portability and Accountability Act (HIPAA) - http://www.hhs.gov/ocr/hipaa/finalreg.html
- USA PATRIOT Act - http://www.epic.org/privacy/terrorism/hr3162.html
- eCommerce / Internet Privacy bills

State initiatives

- California
  - www.caprivacyprotection.org/laws_california.html