GOVERNMENT OF ALBERTA

PRIVACY ARCHITECTURE OVERVIEW

Government of Alberta Enterprise Architecture

Recipient of the

HP Privacy Innovation Award

PrivacyCon 2003, Columbus, Ohio

"This privacy architecture places the Government of Alberta in the forefront of public sector organizations in responding to privacy requirements."

Award news release, October 2, 2003, www.privacycon.org.

Final as of: **May 21, 2003**

Author: **IBM Global Services**

Alberta INNOVATION AND SCIENCE

TABLE OF CONTENTS

TABLE OF FIGURES

INDEX OF TABLES

1. Executive Summary

The Privacy Architecture was conceived both as a logical next step for completing the ICT Privacy Framework, and as a needed extension to the existing Government of Alberta Enterprise Architecture (GAEA). Both requirements had the same objective, namely to provide specific guidance on the intersection of technology with privacy obligations. Recent estimates have placed the proportion of databases within the Government of Alberta (GoA) that will contain personal information to be about 57% once current development initiatives are accounted for. This statistic alone underlines the importance of having clear structured guidance for addressing privacy in an information technology context.

The specific requirements for the Privacy Architecture were established in detail via a set of IBM facilitated cross-government workshops held in October 2002 with representatives from business, policy and IT communities. The result of these workshops was the following list of twelve requirements, which are detailed in the GAEA Privacy Architecture Requirements document. The requirements were divided into three horizons with the closer horizons reflecting requirements expected to have more fully formed output and immediate applicability.

Horizon 1 Requirements:

1. **Terminology** – a common language for discussing privacy requirements, issues and solutions
2. **Identity Keys** - how will data subjects be uniquely identified?
3. **Data Classification** -- how should personal information or its uses be classified?
4. **Data Sharing, Re-Use and Placement** to what extent can personal information be shared between departments and where should it be stored?

Horizon 2 Requirements:

1. **User Interface** - what privacy related features are required and what should they look like?
2. **Data Transformation** - guidance for rendering data anonymous
3. **Data Subject Access to Data** – how should Data Subjects be provided with access to their own data?
4. **Software Acquisition Criteria** – privacy criteria for both privacy-enhancing and general software
5. **Consent and Choice** - rules for what consents and choices are to be offered
6. **Access Control** – expression of “need to know” in a privacy context

Horizon 3 Requirements:

1. **Use of Technology to Enforce Privacy Rules** - where should technology be used to enforce privacy rules vs. using processes and procedures?
2. **Use of Technology to Monitor Privacy Compliance** - where should technology be used to monitor privacy compliance vs. using processes and procedures?

IBM was engaged again to facilitate the building of a Privacy Architecture against these twelve requirements and began by building a Straw Model based on experience, best practices, research, and available reference models. This Straw Model was then reviewed and modified through a series of cross-government workshops held in March 2003, again with representatives from business, policy and IT communities. Time spent on each topic was roughly proportional to its position in the horizon scheme.

The specific output of the workshops was an initial Privacy Architecture consisting of eight **Guidance Elements** plus a set of **implementation recommendations** for implementing the Architecture. The remainder of the Executive Summary provides a high-level view of these Guidance Elements and implementation recommendations, while the body and appendices provide the detail.

The Privacy Architecture Guidance Elements are:

1. A **Privacy Glossary** that provides a common language for business, policy and information technology communities to discuss privacy requirements, issues and solutions effectively and without ambiguity.
2. A **Privacy Taxonomy** which provides a comprehensive scheme to consistently label privacy-relevant objects and actions in an IT environment to increase the speed and strategic alignment of both design and operational decisions. It is based on recognized industry standards and directions but with extensions to allow customization with specific definitions at the cross-government, departmental levels and beyond. It offers near term benefits in terms of design and operational decisions in such areas as the placement, security, handling and audit of personal information. It also builds a foundation for the longer term adoption of active privacy technology which enforces privacy in a real-time manner
3. An **Identity Key Scheme** based on hidden Meaningless But Unique Numbers (MBUNs) which are used to reference all personal information instead of using publicly known identifiers. It provides a simple way for personal information to be decomposed into separately accessible pieces that align with the minimum personal information needs of various users. Some of these keys are used to define “islands” where personal information is normally shared (e.g. within a Program) and others define “bridges” to selectively allow islands to exchange or share personal information without revealing their own keys to each other.
4. **Privacy Design Guidance** consisting of discreet pieces of privacy wisdom that can be applied as a checklist either during software design or as part of software acquisition requirements. This guidance maps to GAEA and ICT Privacy Framework Principles, and contains specific direction on “static” design features that can improve both privacy protection and perception.
5. A process for **Privacy Transformation** that incorporates techniques for transforming personal information into less identifiable forms. A method is proposed for optimizing the privacy aspects of storing personal information and making it available to users in the least sensitive form appropriate to meet their needs. This is achieved by defining different identity levels for information about individuals, providing procedures for moving between the levels, and procedures for testing the ability to re-identify information.
6. An **Active Privacy Architecture** that defines a future state view of how specialized technology components and services can provide real-time privacy decision-making and

transaction processing. Based on the best available industry references, this provides both direction and near-term design and acquisition considerations that can help move GoA towards the future state. This includes a rules structure for understanding the additional parameters required for making a data access decision regarding personal information that takes privacy policy into account.

7. A process for **Data Placement** which details a method for optimizing the placement of personal information into the data sharing bands and security zones defined by GAEA. The process leverages the Privacy Taxonomy and Design Guidance elements and integrates them with the existing Privacy Impact Assessment (PIA) process in an iterative fashion. Precedent-setting decisions are captured to improve the efficiency of future iterations.
8. A process for facilitating **Private Access** (providing an individual with access to his or her own personal information). The process leverages the Privacy Taxonomy, Identity Keys Scheme and Privacy Design Guidance to assist with the response to an access request that involves provision of structured electronic data. This may be either a routine request or a request under FOIP or HIA.

The guidance elements do not always correspond one-for-one with the twelve requirements and Table 1 in the next section shows the association between them.

The key recommendations for immediate action are:

1. Adopt and promote the use of the Privacy Glossary and integrate it into the existing GAEA Glossary
2. Adopt the Identity Key concept and identify an early opportunity to pilot it, including the development of the necessary identity protection component
3. Adopt the Privacy Taxonomy and promote its use in building “metadata” descriptions of all databases containing personal information
4. Adopt the Privacy Design Guidance and integrate its use into existing software design and acquisition processes
5. Adopt the Privacy Transformation process and promote its use in application and database design to reduce identity level of information the minimum necessary
6. Find an existing data sharing arrangement involving personal information and pilot the process for Data Placement

The near-term benefits of adopting these recommendations include:

- More effective communications on requirements, issues and solutions between business, policy and information technology communities.
- Greater speed and success moving IT initiatives through the PIA process through pro-active use of Privacy Architecture guidance elements.
- Better and more consistent software design and acquisition decisions that ensure all realistic privacy protection (and perception) measures are considered at the beginning of the cycle.

- Greater speed, efficiency and consistency in executing manual privacy processes such as Private Access by leveraging Privacy Architecture guidance elements.
- Improved ability to audit that privacy policies, principles and practices are being observed in an IT context.
- Encourage progress from current manual privacy processes towards the eventual use of active privacy technologies and services (which could be shared across the GoA).

At the same time as accruing these near term benefits, adopting these recommendations also positions the Government of Alberta (GoA) in the longer term for quicker and less disruptive adoption of active privacy technology as it becomes appropriate.

Note: The Government of Alberta has succeeded in developing a Privacy Architecture that is among the first of its kind. As such, the GAEA PA presents many new and innovative concepts, techniques, and approaches for implementing "Privacy by Design". As with any first-of-a-kind product, some refinement may be required before it can meet all of the demands that will eventually be placed upon it. For example, some design concepts (e.g. Identity Keys) may have performance and/or capacity issues to be addressed through design elaboration, testing, and enhancement. In this respect, it is suggested that the Privacy Architecture be considered "under construction" – and that users proceed accordingly.

IBM would like to recognize the excellent sponsorship and participation within the GoA, which has resulted in a focused, usable architecture with broad-based support. Although elements of the architecture have been seen before in other organizations and industries, GoA is one of the first organizations to recognize the value of compiling privacy guidance into a structured asset that connects its privacy obligations with its existing enterprise architecture for IT.

2. Privacy Architecture Overview

The Privacy Architecture was developed following the GAEA method as illustrated by the simplified process shown in Figure 1. An initial Requirements Workshop was conducted to establish the 12 key requirements. These requirements were subsequently organized into 3 horizons reflecting the relative focus to be placed on each. The core team then built a Straw Model based on extensive research including an in-depth review of FOIP and HIA legislation as well as industry-leading thought on privacy in a technology context. The Straw Model was then thoroughly reviewed and enhanced by a cross-functional team through a series of workshops to produced the 8 distinct Guidance Elements that make up the Privacy Architecture.

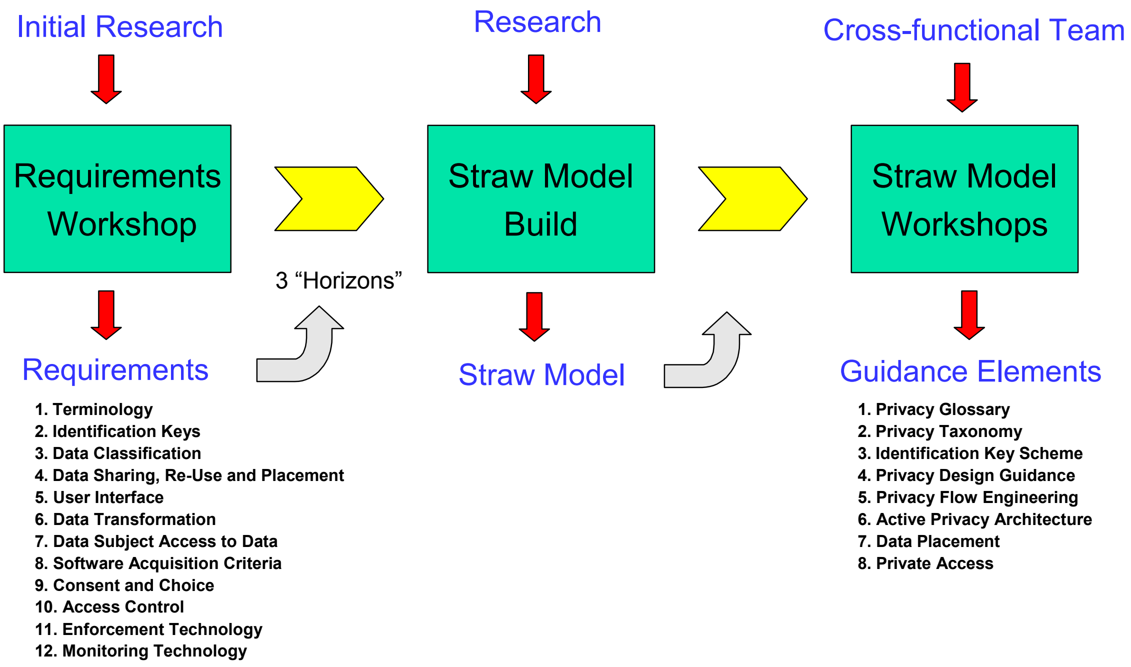


Figure 1: Privacy Architecture Development

As suggested above, there is not a one-to-one correspondence between the original 12 requirements and the resulting 8 Guidance Elements. Some requirements did directly drive out a Guidance Element (for example the Privacy Glossary derived from the Terminology requirement). In other cases, a single Guidance Element addressed many requirements (as is the case for the Privacy Design Guidance). The association between requirements and Guidance Elements is shown in Table 1. Note also that some of the naming conventions were changed as a result of workshop discussions. So for instance, the requirement for Data Classification resulted in a Guidance Element called the Privacy Taxonomy, rather than a Privacy Classification.

Table 1: Association between Privacy Architecture Guidance Elements and Requirements

Requirement Guidance Element	1. Terminology	2. Identity Keys	3.Data Classification	4. Data Sharing, Placement	5. User Interface	6. Data Transformation	7. Data Subject Access	8. Acquisition Criteria	9. Consent and Choice	10. Access Control	11. Privacy Enforcing Tech.	12. Privacy Monitoring
1. Privacy Glossary	X											
2. Privacy Taxonomy			X	X		X	X					
3. Identity Key Scheme		X	X	X								
4. Privacy Design Guidance		X	X	X	X	X	X	X	X		X	X
5. Privacy Transformation				X		X						
6. Active Privacy Architecture		X	X				X			X	X	X
7. Data Placement				X								
8. Private Access							X					

One of the fundamental general requirements of the Privacy Architecture was that it align with, and support the GAEA Privacy Principles and Table 2 shows how these same Guidance Elements align to the principles.

Table 2: Mapping of Guidance Elements to GAEA Privacy Principles

GAEA Privacy Principle Guidance Element	1. Collection Limitation	2. Data Quality	3.Purpose Specification	4. Use Limitation	5. Security Safeguards	6. Openness	7. Access	8. Accountability
1. Privacy Glossary			X			X		
2. Privacy Taxonomy				X			X	
3. Identity Key Scheme				X	X		X	
4. Privacy Design Guidance	X	X	X	X	X	X		
5. Privacy Transformation				X				
6. Active Privacy Architecture			X	X			X	X
7. Data Placement		X						
8. Private Access							X	

2.1 Privacy Architecture Guidance Elements

The following sections provide a high-level view of each Privacy Architecture Guidance Element.

2.1.1 Privacy Glossary

The Privacy Glossary addresses the Terminology requirement and provides a foundation for the Privacy Architecture that will help bridge the potential communication gap between business, policy and information technology communities on the topic of privacy. Use of the Glossary should result in a **common language** for these communities **to discuss privacy requirements, issues and solutions**.

The Glossary is **consistent with FOIP and HIA** and will also be integrated into the GAEA Glossary of Terms. The Glossary has been delivered as a stand-alone item of some 80 plus terms and has a format as illustrated by the sample in Table 3. Note that it is cross-referenced, indicates the source for each definition and is mapped to the Privacy Taxonomy that will be described shortly.

Table 3: GAEA Privacy Glossary - sample

Term	Definition(s)	Notes	Source Details	Taxonomy
Access	The ability and means to communicate with a system in order to use its resources to either handle information or to gain knowledge of the information it contains. Access is a specific interaction type that results in the flow of information.	This would normally refer to an authorized employee (Requestor) gaining access to data as part of a job responsibility but could include Private Access.	GAEA Glossary of Terms	Action
Action	A policy dimension of the Privacy Taxonomy that is used to describe the basic form of privacy-relevant manipulation the Data User intends to take on the Personal Information in question.	For example, provide **Error! Reference source not found.**, obtain Consent, perform a Disclosure	Created for the Privacy Architecture	Action

2.1.2 Privacy Taxonomy

The **Privacy Taxonomy** directly addresses the requirement for data classification and goes beyond to provide a comprehensive scheme to consistently label privacy-relevant objects and actions in an IT environment. Implementation of the Privacy Taxonomy will help to **increase the speed and strategic alignment of both design and operational decisions**.

The Taxonomy is based on recognized **industry standards and directions** but with extensions to allow **customization** at the cross-government and departmental levels. It offers near term benefits in terms of design and operational decisions in such areas as the placement, security, handling and audit of personal information. It also builds a foundation for the longer-term adoption of specialized privacy technology, which enforces privacy in a real-time manner.

The Taxonomy has several dimensions, as illustrated in Figure 2, that allow different privacy-relevant attributes to be expressed as required. It has a **Data Dimension,** which expresses attributes that are properties of the personal information itself. It also has a **Policy Dimension**, which expresses attributes that are needed to describe the policies that apply to the data. These policy dimensions are organized into **Intent**, **Conditions** and **Consequences** groupings, which prepare the way for policy to be described in a format that can be interpreted by technology at some future point.

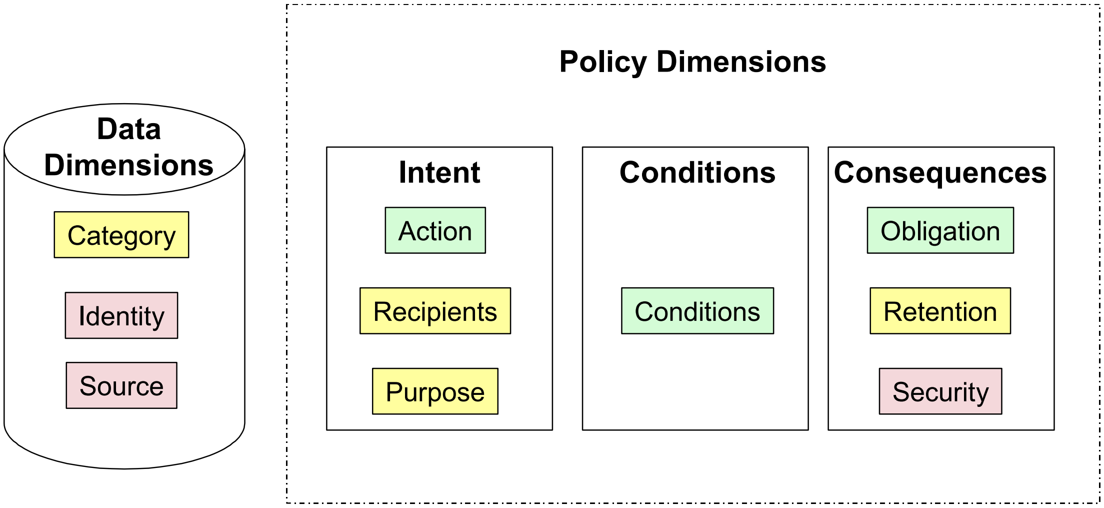


Figure 2: Privacy Taxonomy Dimensions

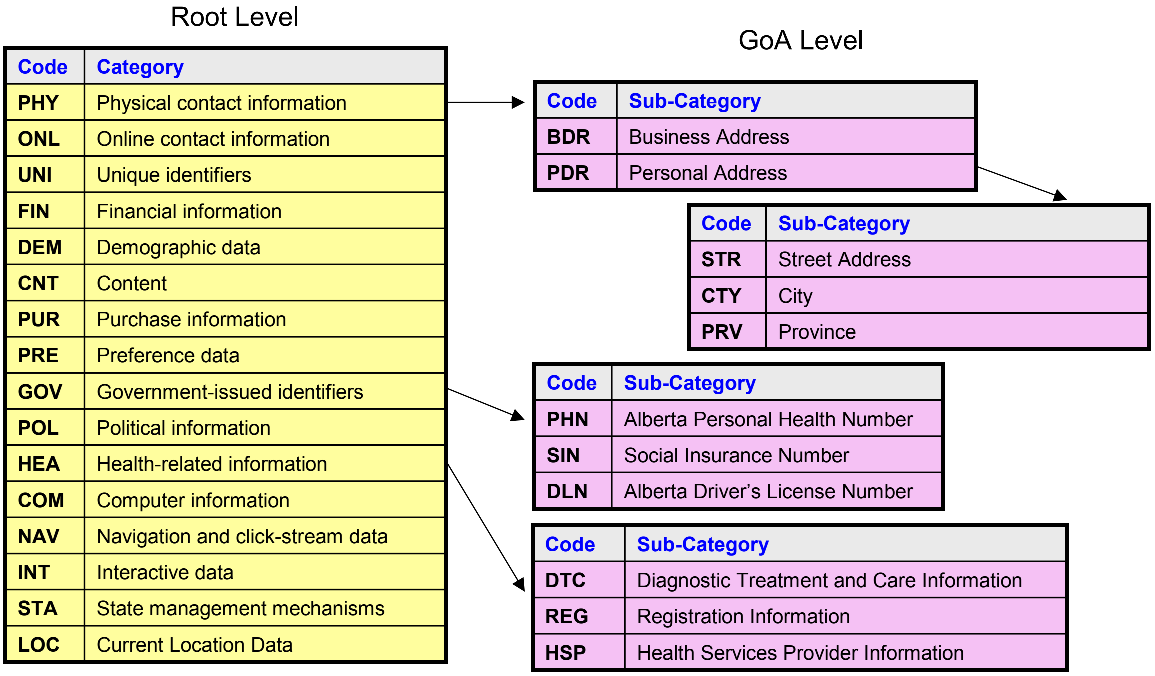
The specific dimension elements are the following:

- **Category** – the type of personal information, for example contact data, health data etc.
- **Identity** – the degree to which the information is anonymized
- **Source** – the source of the personal information, for example it could be collected from the individual or it could be from a third party etc.
- **Actions** – the action intended to be taken on the information, for example it may be updated, deleted or disclosed etc.
- **Recipients** – the party(s) that will be receiving the information, for example, the Privacy Commissioner, a law enforcement agency etc.
- **Purpose** – the reason for taking action on the information, for example, to provide health services, to conduct research, to pursue law enforcement etc.
- **Conditions** – conditions that must be met before the action is allowed, for example, providing proof of authority or obtaining individual consent
- **Obligations** – obligations incurred as a condition of being allowed to take the action such as informing an individual of their right to appeal a decision
- **Retention** – the length of time the information can be kept
- **Security** - the security level required to protect the information

Each of the above has a specific set of pre-defined values, which allow personal information and privacy policy to be fully defined. These values are arranged in a hierarchy:

- **Root Level** – contains "universal" dimensions that reference outside standards wherever possible
- **GoA Level** – contains GoA specific dimensions that will be common across GoA
- **Ministry Level** – contains Ministry-unique dimensions common within a Ministry

By way of illustration, Figure 3 shows how this hierarchy applies to the Category dimension. Note that the examples shown for the GoA level are not exhaustive, for instance, there are other categories of Health information outside of those defined in the Health Information Act.



Root Level

Code	Category
PHY	Physical contact information
ONL	Online contact information
UNI	Unique identifiers
FIN	Financial information
DEM	Demographic data
CNT	Content
PUR	Purchase information
PRE	Preference data
GOV	Government-issued identifiers
POL	Political information
HEA	Health-related information
COM	Computer information
NAV	Navigation and click-stream data
INT	Interactive data
STA	State management mechanisms
LOC	Current Location Data

GoA Level

Code	Sub-Category
BDR	Business Address
PDR	Personal Address

Code	Sub-Category
STR	Street Address
CTY	City
PRV	Province

Code	Sub-Category
PHN	Alberta Personal Health Number
SIN	Social Insurance Number
DLN	Alberta Driver's License Number

Code	Sub-Category
DTC	Diagnostic Treatment and Care Information
REG	Registration Information
HSP	Health Services Provider Information

Figure 3: Sample Root and GoA Level values for the Category Dimension

When used in a standard compact notational format, the Privacy Taxonomy can be applied to create consistent "meta-data" descriptions of database content and the policies that apply to it.

2.1.3 Identity Key Scheme

The Identity Key Scheme addresses the Identity Key requirement directly and provides a means for connecting an individual's identity with information stored about them in a privacy-preserving fashion. When compared with existing Identity key techniques within the GoA, it is expected that the proposed Identity Key Scheme will improve both privacy protection and the ability to share information when that sharing is authorized. Following are some key advantages of this scheme:

- Meaningful public identifiers (e.g. SIN) are separated from PI – making information "depersonalized" and therefore lowering privacy risk by making information less easily identifiable.
- Only the most sensitive data (e.g. ID linkage functions and tables) need to be kept in the highly secure zone – thus making control points easier to manage
- Promotes an efficient and effective means of sharing by enabling depersonalized PI to be placed in Band 1 for cross-government sharing if appropriate (with some conditions)
- ID key scheme allows for separate storage of different aspects of identity for greater security
- Authorized data sharing by applications can occur readily, under the control of the identity protection component

The Identity Key Scheme is based on several types of **Meaningless But Unique Numbers** (MBUNs), which are used to reference all personal information instead of using a publicly known identifier such as a Social Insurance Number. Individuals continue to use existing Public IDs (PIDs), such as conventional user IDs to identify themselves from the outside, but special Identity protection components map these to the appropriate MBUNs for internal use.

A hidden MBUN called an **Internal ID or IID** is used internally by applications and databases to index all personal information within a chosen identity domain, such as a Program, within which personal information is normally shared. This in effect creates an "**island**" outside of which the Internal ID does not work, as shown in Figure 4. Management of the Internal IDs is performed by a new software component called an **Identity Protection Component**. This component generates new Internal IDs when required, registers new IDs and provides protected mapping of Public IDs to Internal IDs.

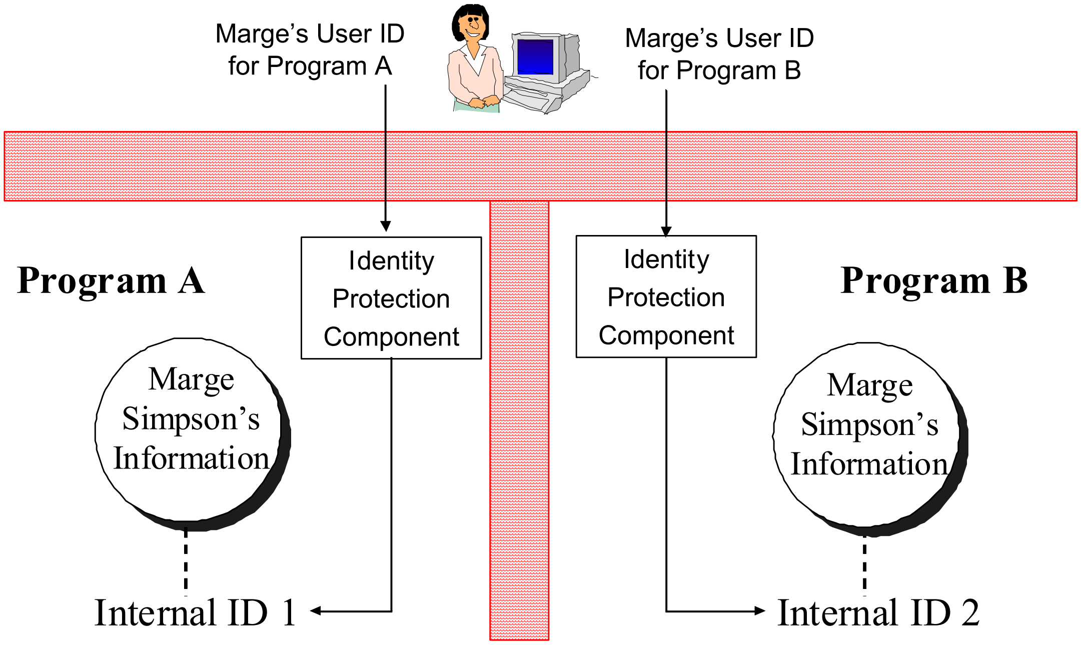

Marge's User ID for Program A
Marge's User ID for Program B
Program A
Identity Protection Component
Identity Protection Component
Program B
Marge Simpson's Information
Marge Simpson's Information
Internal ID 1
Internal ID 2

Figure 4: Internal ID's create "islands" of personal information

Use of the Internal ID provides a simple way for personal information to be broken down and stored in separately accessible parts that align with the minimum personal information needs of various users and are less sensitive if inadvertently disclosed. This is illustrated in Figure 5.

Individual	Employee #	Salary	Rating
Homer Simpson	123 456 789	$45K	4
Joe Quimby	987 654 321	$250K	2
Semour Skinner	827 364 591	$95K	1
Sally Skinner	837 234 003	$97K	2
Ralph Wiggum	007 461 101	$2K	3

Individual	Employee #	Internal ID
Homer Simpson	123 456 789	40598342
Joe Quimby	987 654 321	61036229
Semour Skinner	827 364 591	00836291
Sally Skinner	837 234 003	39570093
Ralph Wiggum	007 461 101	39608801

Salary	Rating	Internal ID
$45K	4	40598342
$250K	2	61036229
$95K	1	00836291
$97K	2	39570093
$2K	3	39608801

Figure 5: Separation of Personal Information Using Internal IDs

A shared MBUN called a **Federated ID or FID** is used as a "**bridge**" to selectively allow islands to exchange or share personal information without revealing their Internal ID's to each other. As with Internal Ids, these are only used internally by applications and databases and are never seen by humans. These bridges must be explicitly set up and so provide a control point for allowing sharing only in approved circumstances as shown in Figure 6. The Identity Management Component also manages Federated IDs in a similar manner to Internal IDs.

As a matter of law, federated IDs can only be enabled if the related data sharing, within or between public bodies, is allowed by the applicable privacy legislation (the FOIP Act or the Health Information Act).

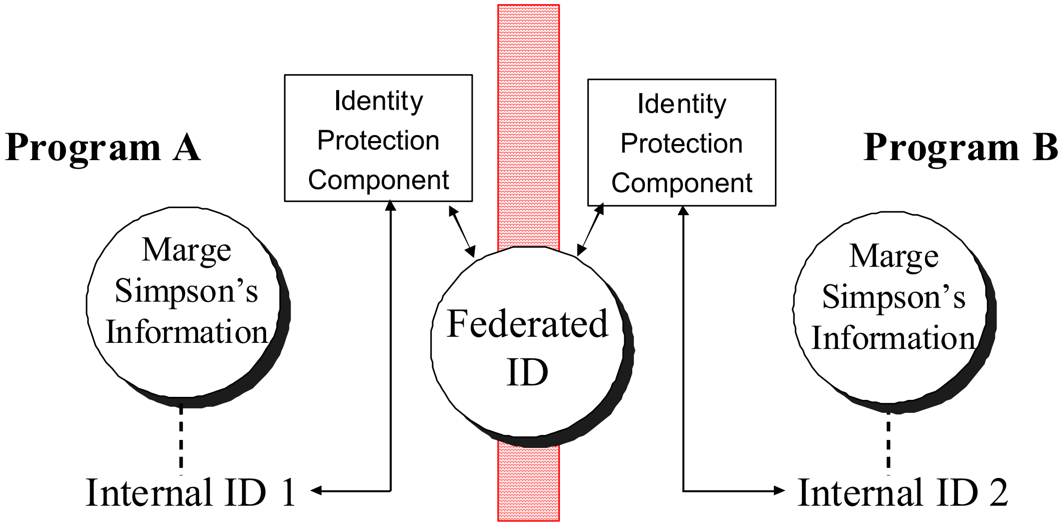

Program A
Identity Protection Component
Identity Protection Component
Program B
Marge Simpson's Information
Federated ID
Marge Simpson's Information
Internal ID 1
Internal ID 2

Figure 6: Federated IDs selectively link information between islands

Federated ID's are created through a secure registration process that cross-references them to a Public ID (or combination of Public Ids) that is common to all the participating local identity domains. This ensures that each Federated ID created will uniquely map to the same person's information in each of the participating local identity domains. However, each local identity domain continues to store personal information based on its Internal ID (not the Federated ID) and maintains its own unique mapping of the Federated ID to the Internal ID. In this manner, knowledge of a Federated ID can only be used to access personal information within a local identity domain by those already authorized to use it in the domain. This protects privacy by avoiding the need for a single key that connects data belonging to an individual across all participating domains.

Existing Public IDs, such as conventional user IDs, are the only element in the Identity Key Scheme seen by human eyes and they can be linked or unlinked to the underlying Federated/Internal ID scheme in a number of flexible ways. This allows a separation between how external public registration schemes are managed and how personal information is stored and managed internally. For instance, individuals could be given a single user ID that works across many Programs and yet each program would retain independent control of the personal information it manages for that individual.

Figure 7 summarises the use of Internal and Federated IDs, Identity Protection Components and their relationship to existing Public IDs.

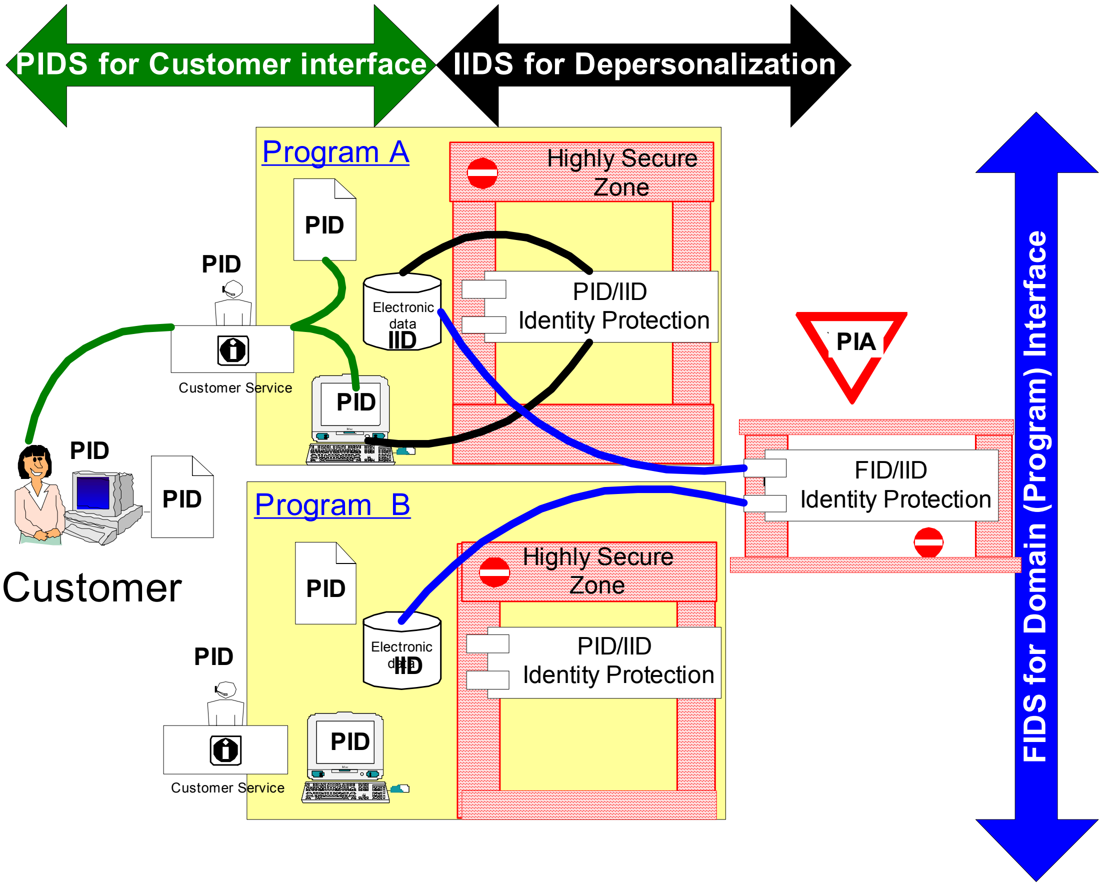

PIDS for Customer interface
IIDS for Depersonalization
Program A
Highly Secure Zone
PID
PID
Electronic data
IID
PID/IID Identity Protection
Customer Service
PID
PIA
FIDS for Domain (Program) Interface
PID
PID
Customer
Program B
FID/IID Identity Protection
Highly Secure Zone
PID
PID
Electronic data
IID
PID/IID Identity Protection
Customer Service
PID

Figure 7: Identity Key Scheme Summary

The recommendation from the Straw Model Workshops was that the default identity domain for an “Island” (i.e. an Internal Identifier) be a government defined Program, operating within a single government department, which typically represents the boundaries within which personal information would be shared under routine circumstances. Note that the preceding descriptions have shown an implementation with multiple copies of the Identity Protection Component, but it could also be implemented via a single cross-government component if desired.

The identity key scheme could also be the basis for other useful functions such as a “Personal Information Finder” which locates all of an individual’s personal information within a given identity domain for the purposes of responding to a private access request.

2.1.4 Privacy Design Guidance

Privacy Design Guidance consists of a series of **discreet pieces of privacy ‘wisdom’** representing existing best practices, some of which are already in use in various areas across government. The Guidance does not address one particular requirement, but rather it addresses many requirements at the same time. The Guidance also directly maps to GAEA and ICT Privacy Framework Principles in many instances.

Much of the Guidance contains specific direction on "static" information technology design features that can **improve both privacy protection and perception**. Most of the concepts can be **applied immediately** and can serve as a pro-active checklist either during **software design** or as part of **software acquisition requirements**.

The Guidance applies both to general software and to privacy-enhancing software. Much of the guidance focuses on "static" design features that can improve both privacy protection and perception – especially at the user interface. It is organized into the following topic areas:

- Architectural Concepts
- Application Design
- User Interface Design
- Database Design
- Logging, Retention and Audit
- Classification Schemes
- Authentication, Authorization and Identity Management

This guidance is intended to provide a basis for evolving a privacy compliant IT infrastructure and for ensuring that components added to the infrastructure either enhance privacy management, or at the very least, do not compromise it. The guidance can be useful both from the perspective of developing IT components like applications and for establishing criteria for procuring IT components.

The Privacy Design Guidance consists of some 50 plus individual elements of guidance separated into "Concepts" which should be followed under normal circumstances, and into "Considerations" which really represent privacy decision points where the facts should be evaluated to yield the optimal decision under the circumstances:

- **Concept:** Based on best practice or thought leadership. Once accepted, these become the default guidance position for IT design and implementation. At some later point, a second pass can be made to determine the strength of this guidance for each concept. The strongest form might be a Standard, variance from which might require an exception process
- **Consideration:** Represent a decision point where the situation should be evaluated to decide on a course of action for design or implementation. Pros and cons are articulated to assist with the decision. If it is determined that the evaluation will almost always point to the same decision for a particular environment, the status could be changed from consideration to Concept.

The complete set of Privacy Design Guidance is provided in the Privacy Architecture Detailed Report, however a sample of the structure is shown in Table 4.

Table 4: Privacy Design Guidance sample

ID / Category / Related Requirements	Guidance	Comments and Rationale
UI1 User Interface	Consideration: Collect PI in context – collect only the PI needed at the point-in-time that it is needed	• Incremental collection of PI whenever it is needed is perceived as less privacy invasive than collection of a larger set up front • This also supported the principle of Notice since by matching up collection with granular purposes, the uses are much clearer to the Data Subject • Clearly this is most applicable when there are multiple optional application/process paths that a Data Subject may take – with correspondingly different PI requirements. • If all PI is ultimately going to be required, even if not immediately, and if collection closer to the time of use is not convenient for the individual or the program, then up-front collection is acceptable.
UI2 User Interface	Concept: Clearly distinguish PI collection fields that are optional from those that are required.	• Supports the principles of openness and limited collection • Optional PI is not so common in a GoA context but may occur where service choices are provided (ex: “Please supply your email address if you wish to be notified by email”). • Needs to be accompanied by a corresponding description of the additional services or benefits that the individual will realize if they provide the additional PI.
UI3 User Interface	Concept: Employ validation checks when collecting PI that are commensurate with the consequences to the Data Subject of processing inaccurate data	• Supports the principle of accuracy. • If consequences of inaccuracy are severe, then validation checking should be extensive. • Examples of validation techniques include format checking (ex: telephone numbers), confirmation (ex: “enter new password twice”) or checks against normal value ranges or existing data

The recommendation is for the Privacy Design Guidance to be adopted for immediate use in software design and acquisition processes and to be maintained via the GAEA Vitality Process using the Privacy Framework Advisory Committee to review and approve changes

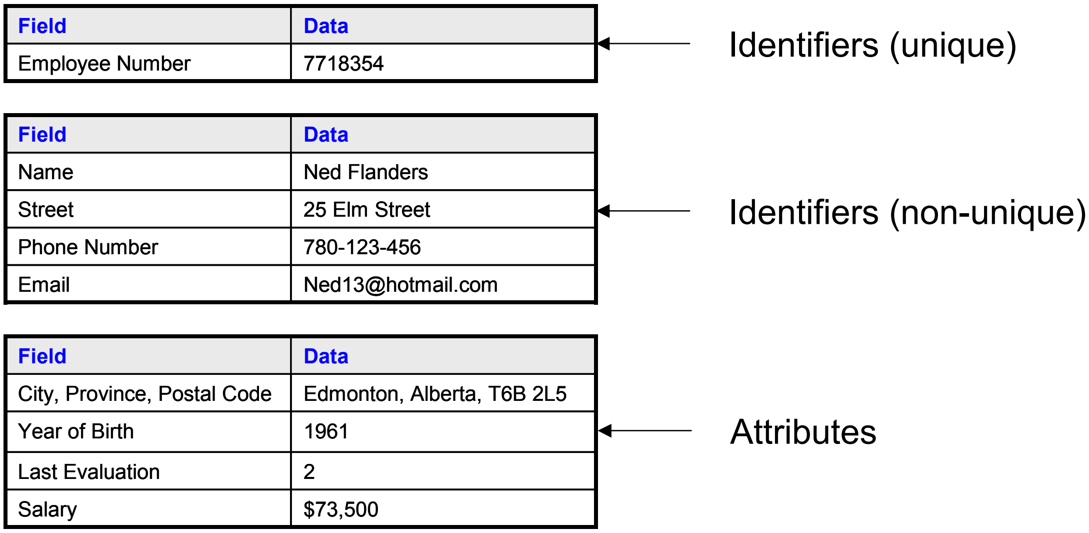
2.1.5 Privacy Transformation

One of the main tenets of good privacy practice is that personal information should be **transformed into the least sensitive form possible** that will still allow the legitimate purpose for use to be met. In fact, this is often required by legislation, but the challenge is that there are rarely any specifics on how such transformations should be made.

Privacy Transformation provides such guidance by describing a number of transformation techniques to render personal information into less sensitive forms and combines these into a structured process for optimizing the privacy aspects of storing personal information and making it available to users in the least sensitive form appropriate to meet their needs.

“Least sensitive form” translates into two imperatives. The first is simply providing the least amount of information that will allow the job to be done. The second is to modify the remaining information so that it has the least connection back to identifiable individuals as possible, while still allowing the purpose to be achieved. To address these imperatives, the Privacy Architecture defines different **identity levels** for information about individuals and provides techniques for moving between the levels. It also provides a procedure for testing the ability to re-attach identify to information.

The first element of guidance is simply to break personal information into its constituent parts prior to storing it. As previously described in the Identity Key Scheme, this is achieved by using MBUNs to index the separated information and to re-assemble it for legitimate use. In order to know how to separate the information, we turn to the Privacy Taxonomy, which includes an Identity Dimension that describes how personal information is made up of identifiers (like a Personal Health Number) and attributes (like age or salary). A number of separate **storage states** are defined which describe progressively finer breakdowns of personal information into its constituent identifier and attribute elements, which provides progressively more privacy protection. A Privacy Impact Assessment, or similar process, can determine the degree of breakdown required. Figure 8 shows an example of a basic breakdown of personal information for storage based on identifiers and attributes.



Field	Data
Employee Number	7718354

Field	Data
Name	Ned Flanders
Street	25 Elm Street
Phone Number	780-123-456
Email	Ned13@hotmail.com

Field	Data
City, Province, Postal Code	Edmonton, Alberta, T6B 2L5
Year of Birth	1961
Last Evaluation	2
Salary	$73,500

Figure 8: Simple Breakdown of Personal Information for Storage

Guidance is also provided on identifying and addressing situations of “**Meta Context**” where the label on the personal information container itself may attach personal information attributes as illustrated in Figure 9.

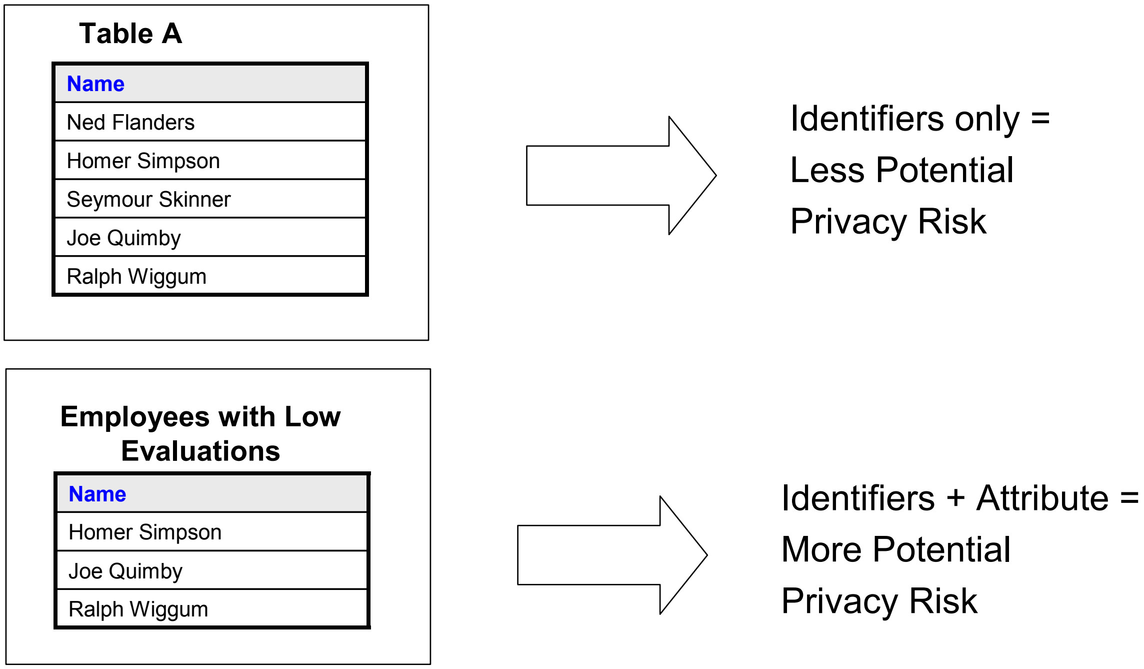


Table A

Name
Ned Flanders
Homer Simpson
Seymour Skinner
Joe Quimby
Ralph Wiggum

Identifiers only = Less Potential Privacy Risk

Employees with Low Evaluations

Name
Homer Simpson
Joe Quimby
Ralph Wiggum

Identifiers + Attribute = More Potential Privacy Risk

Figure 9: An example of Meta Context

Next, a number of **transformation techniques** are defined, which describe how personal information can be transformed between various levels of identity (i.e. the degree to which the information can be associated with an individual). These techniques are:

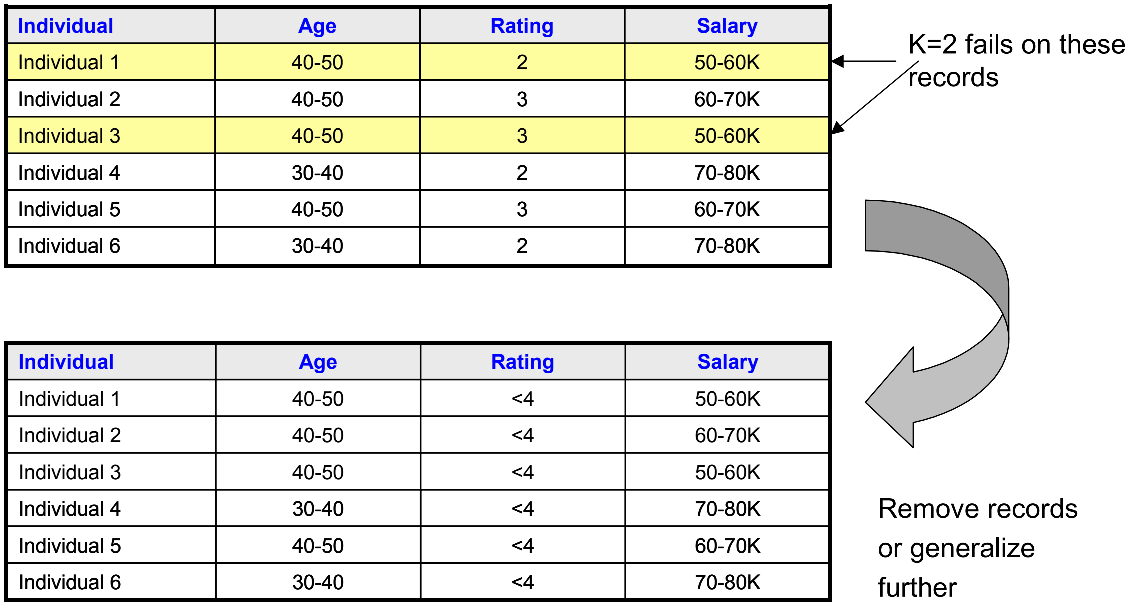
- **Reduction**: Fields are removed that are not required for the purpose
- **Generalization**: specific fields or field values are replaced with generalized fields or values. Examples include replacing city names by province or country, replacing age by an age range, replacing job description with a job category.
- **Suppression**: parts of a field are suppressed or removed (e.g., removing the last three digits of a postal code).
- **Perturbation**: field values are perturbed or blurred using a statistical technique (e.g., adding a random number between -10 and +10 to an age),
- **Aggregation**: aggregate statistics are computed and published (e.g., average of 10 consecutive records).

Table 5 illustrates some of these techniques:

Table 5: Transformation Technique Examples

Field	Data	Technique	Result
City	Mount Albert	Reduction	-
Province	Alberta		
Postal Code	T6B 2L5	Suppression	T6B ***
Date of Birth	14/06/1961	Generalization	Age: 40-50
Last Evaluation	2	Generalization	OK
Salary	$73,500	Generalization	> 50K

Additionally, a simple **uniqueness test** is defined, which allows de-identified data to be assessed with respect to how easy it might be for a "curious" user to re-attach identity using other sources of information as a cross-reference. This technique essentially looks at the number of identical records in a collection of de-identified data and if the number falls below a threshold then the uniqueness test is failed. Figure 10 shows an example where the threshold ("K" value) is set to 2, which means that the data must have no unique records. Intuitively, the more unique the records are in a database, the more potential there is to cross-reference them with outside information and find a unique match with a given individual.



Individual	Age	Rating	Salary
Individual 1	40-50	2	50-60K
Individual 2	40-50	3	60-70K
Individual 3	40-50	3	50-60K
Individual 4	30-40	2	70-80K
Individual 5	40-50	3	60-70K
Individual 6	30-40	2	70-80K

K=2 fails on these records

Individual	Age	Rating	Salary
Individual 1	40-50	<4	50-60K
Individual 2	40-50	<4	60-70K
Individual 3	40-50	<4	50-60K
Individual 4	30-40	<4	70-80K
Individual 5	40-50	<4	60-70K
Individual 6	30-40	<4	70-80K

Remove records or generalize further

Figure 10: Uniqueness test example

Finally, these techniques and tests are incorporated into the Privacy Transformation process illustrated in Figure 11. The concept is based on analysis of user needs and a progressive application of transformation techniques to achieve a minimum usable identity level. It also includes using the uniqueness test and other measures to check that this minimum level is not readily re-identifiable. This process provides a consistent approach to ensuring that users of information about individuals receive the data with the **minimum level of "identity"** required to do their job.

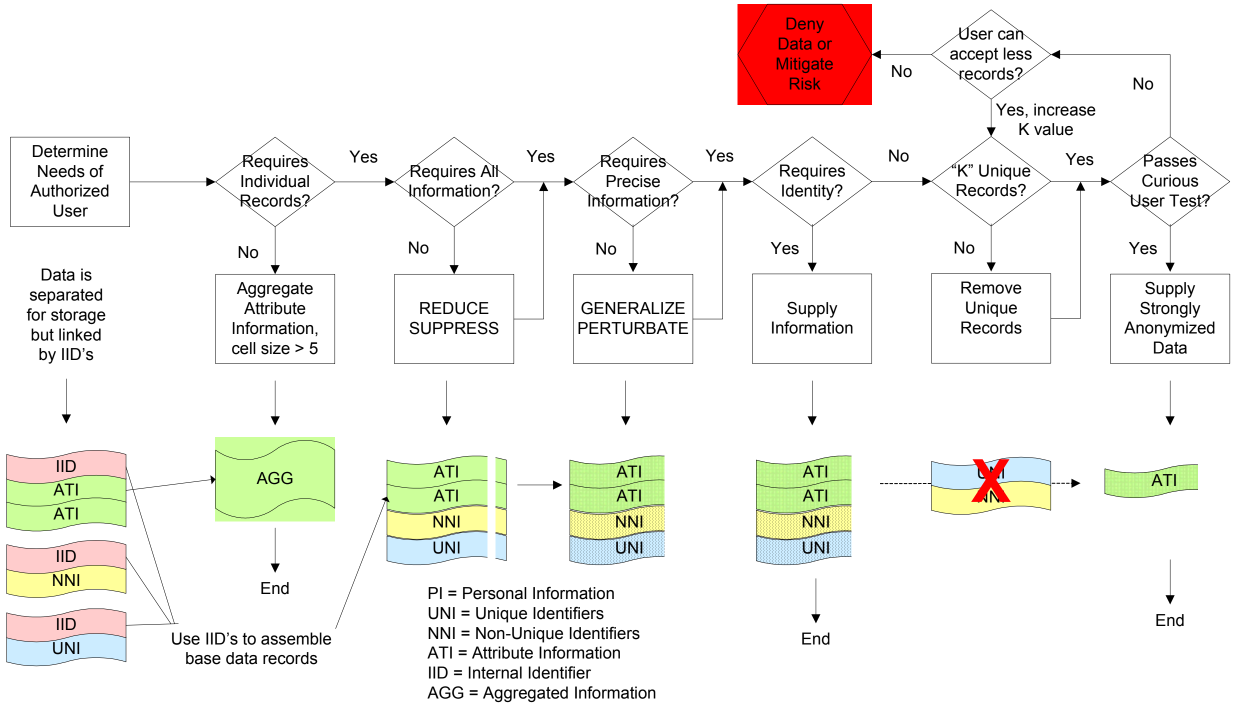

Deny Data or Mitigate Risk
No
User can accept less records?
No
Yes, increase K value
Determine Needs of Authorized User
Requires Individual Records?
Yes
Requires All Information?
Yes
Requires Precise Information?
Yes
Requires Identity?
No
"K" Unique Records?
Yes
Passes Curious User Test?
No
No
No
Yes
No
Yes
Data is separated for storage but linked by IID's
Aggregate Attribute Information, cell size > 5
REDUCE SUPPRESS
GENERALIZE PERTURBATE
Supply Information
Remove Unique Records
Supply Strongly Anonymized Data
IID
ATI
ATI
IID
NNI
IID
UNI
AGG
End
Use IID's to assemble base data records
ATI
ATI
NNI
UNI
ATI
ATI
NNI
UNI
ATI
ATI
NNI
UNI
End
UNI
NNI
X
ATI
End
PI = Personal Information
UNI = Unique Identifiers
NNI = Non-Unique Identifiers
ATI = Attribute Information
IID = Internal Identifier
AGG = Aggregated Information

Figure 11: Privacy Transformation Process

2.1.6 Active Privacy Architecture

The privacy protection incorporated into most existing information technology infrastructures can be characterized as "static". That is to say that the protection consists of fixed design features or implementation decisions that improve privacy in a fixed way. The Privacy Design Guidance element contains many such static protection measures such as designing user interfaces to collect only the minimum necessary personal information.

By contrast, the term "**Active Privacy Architecture**" defines a future state view of how specialized technology components and services can provide real-time privacy decision-making and transaction processing. Such technology may become a requirement to support business strategies that require more complex sharing of personal information and increased flexibility based on individual preferences. Defining such a vision for GoA now provides both direction and near-term design and acquisition considerations that can help move GoA towards the future state with a minimum amount of retrofitting.

The Active Privacy Architecture for GoA is based on the best available industry references at the current point in time including IBM's Enterprise Privacy Architecture and ISTPA's Privacy Framework. Both of these references contain important design concepts that have been adopted as part of GoA's roadmap to active privacy and include:

- The ability to virtually associate all personal information with the rules that govern its collection, use, disclosure and retention
- The use of **specialized privacy components or services** that minimize the impact of implementing Active Privacy into an existing IT infrastructure
- The ability to **manage policy centrally** for consistency but **enforce locally** for performance
- The ability to **choose integration points** that best fit with the existing infrastructure (application versus. database etc.)

The specialized privacy components or services referred to above can be grouped into a number of functional areas, which provide a sense of what Active Privacy technology can do as illustrated in Figure 12.

1. **Access.** Controls access to personal information consistent with privacy policy. Also provides automated functions for transforming personal information to less sensitive forms.
2. **Management.** Supports the Access and User Access functions by providing privacy-specific services for things like policy management, audit, and handling non-routine events.
3. **User Interaction**. Provides specialized communication functions between users and applications that wish to submit or access personal information. This can include functions such as allowing individuals access to their own information.
4. **Identity and Security.** Provides the services required to identify and authenticate users and to securely exchange personal information with other enterprises. Also includes specialized functions such as the ability to allow individuals to interact without revealing identity
5. **Support Tools.** Supports administration functions required to optimize the operation of the active components and to ensure they are operating against up-to-date information.

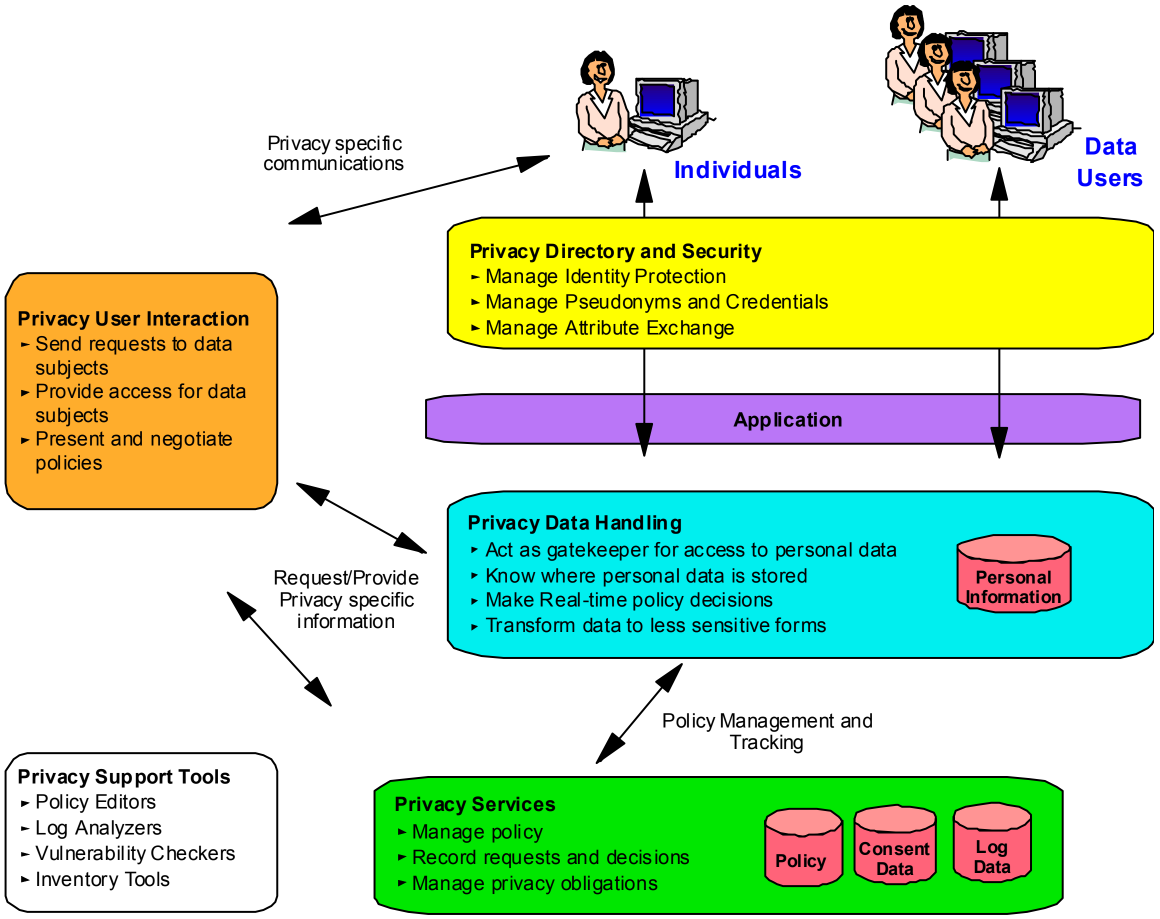

Privacy specific communications
Individuals
Data Users
Privacy Directory and Security
► Manage Identity Protection
► Manage Pseudonyms and Credentials
► Manage Attribute Exchange
Privacy User Interaction
► Send requests to data subjects
► Provide access for data subjects
► Present and negotiate policies
Application
Privacy Data Handling
► Act as gatekeeper for access to personal data
► Know where personal data is stored
► Make Real-time policy decisions
► Transform data to less sensitive forms
Personal Information
Request/Provide Privacy specific information
Policy Management and Tracking
Privacy Support Tools
► Policy Editors
► Log Analyzers
► Vulnerability Checkers
► Inventory Tools
Privacy Services
► Manage policy
► Record requests and decisions
► Manage privacy obligations
Policy
Consent Data
Log Data

Figure 12: Active Privacy Functions

As described in the Access functions above, a key aspect of Active Privacy is the ability to **automate a privacy-enhanced access to personal information decision**. Privacy is distinct from pure security in this regard as privacy introduces concepts such as “purpose” and “individual consent” as potential factors in making an access decision.

The structure of rules that enable access decisions to be made against privacy policy are important to understand, and for GoA purposes, a generic extract from IBM’s Enterprise Privacy Authorization Language is used to define the parameters necessary to construct such rules. These parameters are embodied in the previously described Policy Dimensions of the Privacy Taxonomy.

In summary having a vision of a future Active Privacy infrastructure helps align current information technology design and acquisition decisions to a target state and reduce the potential need to retrofit. In addition, it allows earlier definition and evaluation of potential automation such as:

- Privacy Enhanced Access to personal information including factors such as consent
- Privacy Vulnerability and Compliance Analysis
- Data Transformation to less sensitive forms
- Data Validation to ensure accuracy and currency
- Private Access for individuals to access their own information
- Interaction and assertion of facts without revealing identity
- Proactive Citizen contact functions

2.1.7 Data Placement

An important objective of the Privacy Architecture is to provide guidance on placing personal information into the **data-sharing bands** and **security zones** defined by GAEA. This must be done while protecting privacy and balancing the efficiency, data quality and client convenience of sharing data to the extent allowed by legislation.

The initial hope was that fixed guidance in this regard would be possible to develop using the Privacy Taxonomy. Unfortunately, it was determined that this is not viable at this time because the context required to provide this guidance is too variable and complex to be represented by a static model. However, it is viable to incorporate elements of the Privacy Architecture into a **process** by which placement decisions can be made and by which the **precedence** of such decisions can be tracked over time and used to improve the speed and consistency of future decisions. Key input factors into this process are the type (category) of personal information in question and the level of identity associated with the information.

The resulting **Data Placement** process details a method for optimizing the placement of personal information into the data sharing bands and security zones defined by GAEA. The process leverages the Privacy Taxonomy and Privacy Design Guidance and integrates them with the existing Privacy Impact Assessment (PIA) process in an iterative fashion. Precedent-setting decisions are captured to improve the efficiency of future iterations.

The process first makes use of the Privacy Taxonomy **Categories** to establish which data is **potentially shareable** between two or more programs within government as illustrated in Figure 13.

Personal Information Category	Program A	Program B	Program C	Potentially Shareable Information
C1	Yes	No	No	X
C2	Yes	No	Yes	✓
C3	No	Yes	No	X
C4	No	Yes	Yes	✓
C5	Yes	Yes	Yes	✓

Figure 13: Analysis of Potentially Shareable Information

This establishes the target data-sharing band for proposed placement, that if approved by a PIA, might result in an implementation similar to that shown in Figure 14. At this point, design guidance from the Privacy Architecture is used to identify any risk mitigation steps that would be used in conjunction with sharing (Identity Keys, Privacy Transformation etc.)

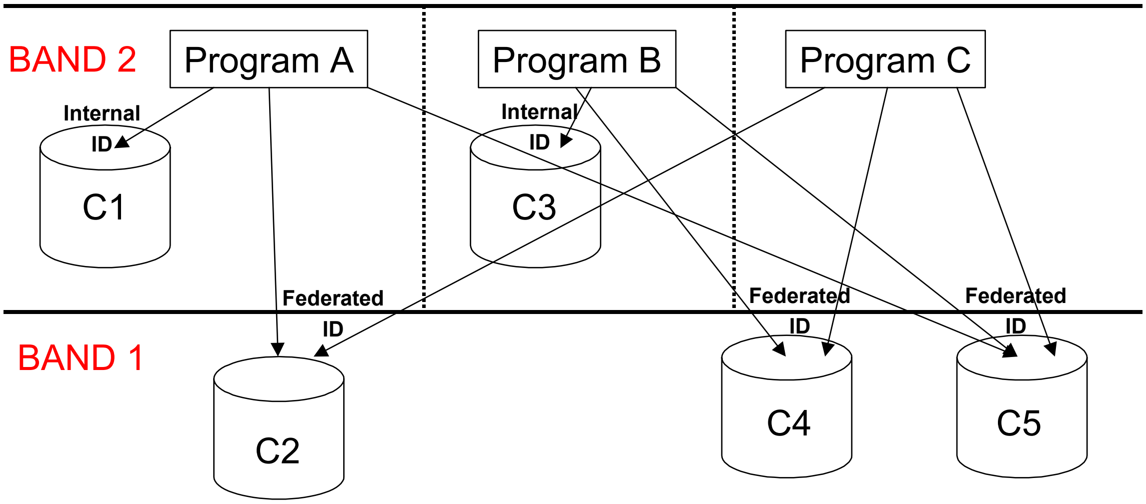


Figure 14: Placement of Shareable and Non-Shareable Data

The proposed band information, along with the **Identity Level** of the information, as established by the Privacy Taxonomy, is placed into a **Privacy Assessment Table** for input to the Privacy Impact Assessment (PIA) process. This table is customized for each Program and contains a case history reference to previous decisions, which becomes more refined over time.

	Category: medical/diagnostic	**Category:** personal
Identifiability: Personal	Band1: not allowed (exceptions possible…) Band2: HighlySecZone + privacy req. Cases: 92981, 09293, …	…
Identifiability: De-identified	Band1: HighlySecZone + privacy policy Band2: Internal Zone + privacy req. Cases: 29801, 2293, …	
Identifiability: Weakly Anonymized	…	B1: Internal B2: Internal Cases: 2293, …

Figure 15: Privacy Assessment Table

In addition, any appropriate concepts from the **Privacy Design Guidance** (for example, use of identity keys as shown in Figure 14) are also input into the design proposed to the PIA as proactive measures to mitigate privacy risk. The PIA process renders the ultimate decision if data sharing is allowed, on the band the data can be placed into, and on the corresponding required security zone.

As illustrated in figure 15, the process of evolving a data sharing idea from initial concept through to final decision made in the PIA step is often iterative. For example, a data sharing idea could be submitted for a preliminary PIA along with a proposed data placement. The PIA can provide recommendations back to the designers from which the data sharing and placement concepts are elaborated and submitted for further PIA. This process could iterate several times as the concept evolves. Once a final decision is made, it is recorded for future precedence. Once the PIA is completed, the detailed design is completed to incorporate the protection measures identified earlier as risk mitigation factors. These may include other elements of the Privacy Architecture, such as Identity Keys, or other elements of GAEA as a whole.

The difference between this process and the current situation is that this provides a consistent pro-active approach that improves in speed and effectiveness over time since it makes use of past precedence. The process is illustrated in Figure 16.

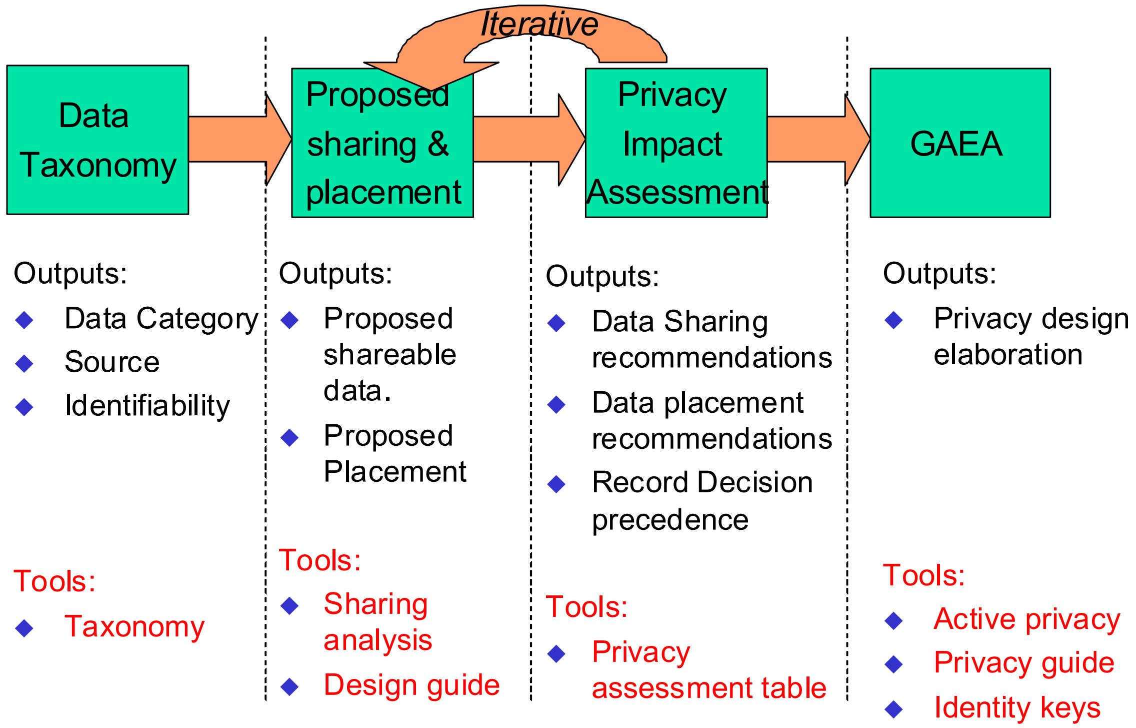


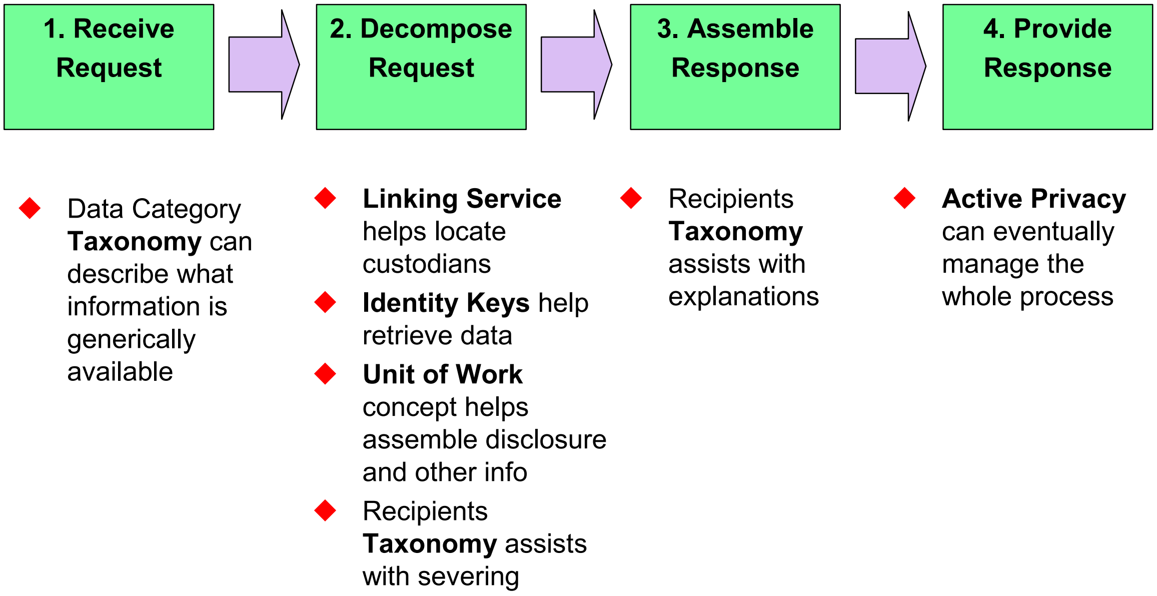
Figure 16: Data Placement Process

2.1.8 Private Access

Providing an individual with access to his or her own personal information, and potentially also processing subsequent requests to make corrections to that information, can be very slow and time-consuming process. Accordingly, one of the objectives of the Privacy Architecture is to provide guidance that may make this process more effective.

The resulting **Private Access** process is illustrated in Figure 17. It leverages the Privacy Taxonomy, Identity Keys Scheme and Privacy Design Guidance to assist with the response to an access request that involves provision of structured electronic data (e.g.. not email or off-line archived information). This may be either a routine request (ex: request for a student transcript) or a request under FOIP or HIA.

Note that the Privacy Architecture will only fully apply in the case of structured, coded personal information, i.e., that stored in database structures. For electronic documents it can only apply partially, at the document level, and then only if the documents are managed by an electronic data management system. For paper records it does not apply at all.



1. Receive Request

- Data Category **Taxonomy** can describe what information is generically available

2. Decompose Request

- **Linking Service** helps locate custodians
- **Identity Keys** help retrieve data
- **Unit of Work** concept helps assemble disclosure and other info
- Recipients **Taxonomy** assists with severing

3. Assemble Response

- Recipients **Taxonomy** assists with explanations

4. Provide Response

- **Active Privacy** can eventually manage the whole process

Figure 17: Private Access Process

The **Privacy Taxonomy** can facilitate the Private Access process by using:

- The **Category** dimension to describe what personal information is potentially held and where
- The **Recipient** dimension to capture conditions and reason codes for personal information that will not be released or updates in response

The **Identity Keys Scheme**, once implemented, can provide an interactive list of all databases containing personal information relating to the individual and subsequently providing links to the individual's information in all the databases in which it exists (provided the right authorization is provided).

The **Privacy Design Guidance** includes a "**Unit of Work**" concept that if implemented could tie together any histories of data usage or disclosure that are required to be released as part of the response.

Eventually, **Active Privacy** components could be used to manage the entire process for structured, available electronic information.

In the short term, it is likely that the Private Access process may play an assist role only for many requests because much of the target information is not in electronic form or is not structured or available online. However, even in this limited role, the process illustrates how founding Guidance Elements such as the Privacy Taxonomy start to have many different applications once they are implemented.

3. Summary of Recommendations

The following chapter provides a summary of recommendations shown against the original 12 requirements broken down by the planning horizons. Within each requirement, the recommendations may be segregated into phases if there is an obvious distinction between immediate and future actions.

3.1 Horizon 1 Requirement Recommendations

3.1.1 Terminology

The following are recommended immediate implementation steps and considerations for the Privacy Glossary:

1. Append the Glossary to the **GAEA Glossary of Terms**
   - Definitions that were copied from the GAEA Glossary of Terms can then be replaced with pointers
2. Request that the structure of the GAEA Glossary of Terms be modified to match that of the Privacy Glossary and include usage notes and specification of sources for each term
3. Maintain the Glossary via the existing **GAEA Vitality Process**
   - The **Privacy Framework Advisory Committee** should play a key role in approving changes to the privacy architecture made through the GAEA vitality process.

3.1.2 Identity Keys

The following are recommended implementation steps and considerations for the Identity Keys Scheme:

Phase 1 Recommendations:

1. Adopt the concept of using hidden MBUN's to index personal information
   - In particular, the concept of Internal, Federated and Public IDs
   - The government Program should be the default level for the identity domain of an Internal ID. For the purposes of the privacy architecture, a 'program' is a collection of functions or services, operating in a single government department, within which the exchange of personal information is required for its continued operation and which would be unlikely to be divided in a reorganization of government departments.
   - Make the use of identity keys a direction for new software development and acquisition involving the management of personal information.
   - **Reconcile the identity key scheme with the current Authentication and Authorization project to ensure compatibility – this is a high priority item due to the timetable for the Authentication and Authorization project.**

- Integrate the identity key scheme into the GAEA Data, Application, Security and Technology domains. Since the Privacy Architectures cuts across the four primitive GAEA domains, a core team discussion needs to occur as part of the transition plan which should consider factors such as:
    - o Usability: Should a user receive ALL relevant guidance (including Security and Privacy) when reviewing a primitive domain (e.g. Application Architecture)?
    - o Integration integrity: If not completely integrated, how are the separate pieces kept in synch on an ongoing basis?
    - o Maintenance. How difficult is it to maintain if something an element found in many places in GAEA?
    - o Tools. EA repository tools have the potential to provide the best of all worlds (e.g. usability, integration, maintenance). Since the tools are changing over time, so might the approach to integration.
- Request that GAEA Business Architecture follow-through on inventory of Programs and services to support the definition of the different “islands”. This will be important when defining “bridges” since bridges should only be allowed between recognized islands.

2. Conduct a one-time departmental survey to determine:
    - Existing use of unique identifiers and their individual/combined use as identity keys (e.g. SIN, Health ULI, Student Number etc.)
    - Existing use and format of MBUN’s in applications and data bases so that an MBUN format can be chosen that will minimize implementation impact to existing systems
    - Leverage the existing PI Data Stores Inventory and use as a basis for the survey
3. Make the following implementation design decisions:
    - MBUN format (e.g. for Internal and Federated IDs)
    - Central versus local MBUN issuing services (central service can provide globally unique MBUN’s and is recommended)
    - Single vs. Multiple Identity Protection Components
4. Begin detailed functional and operational design:
    - MBUN creation rules
    - Table structure, placement and protection of mapping and type tables
    - Linking and registration protocols (i.e. define exactly what parameters are passed between the Protected Mapping Services and Registration Services and how they are passed)
    - Begin the process to investigate building or acquiring the identity protection component (including key creation, mapping, registration, and retrieval services)

5. Choose a project to pilot Internal ID use
   - Pick a new initiative and/or and existing system that already uses MBUN's for other reasons
   - Pick an area that does not have access to personal information as fine-grained as they would like (i.e. there will be a benefit to decomposing the personal information using Internal ID's)
   - Try to choose an area that will also be suitable as a Federated ID pilot
   - Also pilot the co-requisite PIA with a view to using it as a template for future projects
6. Choose a project to pilot Federated ID use
   - Try to bridge off the Internal ID pilot
   - Choose an area where personal information sharing is already being done so that just the concept is being piloted and not the sharing precedent
   - Also pilot the co-requisite PIA with a view to using it as a template for future projects
   - Place a high focus on communicating key privacy-protecting design points so that there are no perceptions that the implementation will reduce privacy

Phase 2 Recommendations:

1. Adopt the concept for production use and set up processes to:
   - Register identity domains centrally for maintenance of a common list of authorized identity domains each with a unique identifier (from the GAEA Business Architecture) and a list of codes for approved EID types
   - Registration of Online Services (applications) within an identity domain so that only authorized Online Services may use the Protected Mapping services and Registration services
   - Registration of Identity protection components within an identity domain so that the Registration Service only accepts requests from authorized Protected Mapping Services

3.1.3 Data Classification

The following are recommended implementation steps and considerations for the Privacy Taxonomy:

Phase 1 Recommendations:

1. Adopt the Taxonomy as a government-wide standard under GAEA
   - Create a value proposition as part of the communication so that the short-term benefits of adoption are understood such as:
     - Facilitates separation of data for storage and transformation
     - Provides a basis for identifying data sharing opportunities
     - Can facilitate the processing of Private Access requests
     - Provides consistent input to PIA's for data sharing and security decisions
     - Provides a basis for auditing proper handling of personal information

- Create guidance for applying the Taxonomy that includes:
  - Suggestion for initially applying to databases at the highest level using meta-data descriptions (for both data dimensions and any policy dimensions that are practical)
  - Use in the design stages of new ICT initiatives
- Adopt the subset of pre-defined defaults for personal data elements as part of the GAEA data standard (ex: Name = a non-unique identifier)

2. Make the following implementation design decisions:
   - Notation format
   - Uniqueness of codes by level
3. Create/identify a simple tool (e.g. adding columns to the CITE Personal Information Datastores tables) that departments can use to quickly create specific static meta-data files for describing database content and policy in accordance with the Privacy Taxonomy. It is expected that this simple tool will suffice to support the primary use of the taxonomy metadata during Phase 1 – which is to support static privacy design decisions and processes, and additionally to support the use of P3P.
4. Create a starter-set for the GoA level of the Taxonomy prior to initial release
   - Conduct a one-time departmental survey to build on the Taxonomy definitions included in this report and create a starter set that includes a first cut at populating all the key high level categories.
   - Conduct a one-time departmental survey to determine if there are any other Taxonomies outside of GAEA in general user (Content Management etc.) that might need to be reconciled with the Privacy Taxonomy
5. Maintain the Taxonomy via the existing **GAEA Vitality Process**
   - The **Privacy Framework Advisory Committee** should play a key role in approving changes

Phase 2 Recommendations

1. During Phase 2, the Privacy taxonomy and associated meta data may begin to be used for Active Privacy purposes (e.g. Access Control). At this time, it will be necessary to acquire meta-data tools that will allow the privacy taxonomy to be “attached” to database tables and columns, and to allow active privacy components to access the meta-data for real-time privacy decision-making.
2. The focus for the Taxonomy, and indeed the entire Privacy Architecture, has been on structured, on-line electronic information. It is recommended in the Horizon 2 timeframe that investigation is started into handling of unstructured information (ex: email) and off-line information (ex: archive – tape backup). This may also extend to hardcopy information and the possibility of making electronic “metadata” descriptions of hardcopy information.

Note, these recommendations will have to be tackled in conjunction with GoA-wide Electronic Information management (EIM) initiatives - meta-data will be required for many purposes, and standards will have to be consistent. Timing may be an issue.

3.1.4 Data Sharing, Re-Use and Placement

The following are recommended implementation steps and considerations regarding the Data and Security Zone Placement Process:

Phase 1 Recommendations:

1. The proposed process should be piloted with groups that are sharing PI through a Data Sharing agreement but which are not currently using the band 1 sharing zone to share (i.e. they are sharing by copying within band 2). This will test the principles of the process vs. the precedence of sharing.

Phase 2 recommendations:

1. Once the process has successfully been piloted for an existing data sharing arrangement, the process can be proposed for adoption as a mandatory step for all programs contemplating a Data Sharing Agreement
2. Once the process begins to be exercised, the output should be monitored to see if a requirement for privacy-enhanced security and encryption starts to emerge for band 1 sharing of more sensitive PI. This will essentially be a requirement for Active Privacy in access control.
3. A cross-government initiative might be considered to establish the possibility of sharing basic contact and preference information. This is likely to be one of the few multilateral sharing needs that would require a centralized effort. Most other sharing arrangements are likely to be bilateral in nature.
4. In the longer term, reviewing a case history of decisions may supply enough information to allow the question of a static model for placement to be revisited.

3.2 Horizon 2 Requirement Recommendations

3.2.1 User Interface

The following are recommended implementation steps and considerations regarding guidance for the User Interface

Phase 1 Recommendations:

1. The proposed User Interface design guidelines (part of the overall Privacy Design Guidance) should be adopted and incorporated into both the development and software acquisition processes. Applying a common set of consistent User Interface design guidelines will result in better privacy protection and perception across government.
2. The Cross-Government Internet Standards Committee should be asked to:
   - Adopt the Internet related guidance in the Privacy Design Guidance (mostly relating to the User Interface)
   - Consider a general P3P implementation recommendation, especially for new Service Alberta initiatives

3.2.2 Data Transformations

The following are recommended implementation steps and considerations regarding guidance on Data Transformations

Phase 1 Recommendations:

1. The proposed Transformation techniques and process should be adopted as a government-wide basis upon which ITC solutions should be designed so as to provide minimum required personal information at each processing point.
   - This is especially pertinent for workflow applications that pass PI between processors or business-to-business applications that exchange PI.
2. Develop a reference list of common cross-reference resources that a "curious user" could use to re-identify data inappropriately.

Phase 2 Recommendations:

1. Adopt the guidance for separation of personal information for storage as the MBUN concept is put into production.
2. Monitor the effectiveness of the default "K=5" value for the K-anonymity test and change if appropriate.

3.2.3 Data Subject Access to Data

The following are recommended implementation steps and considerations regarding guidance on Data Subject Access to Data:

Phase 1 Recommendations:

1. Pilot a "front counter" and or FOIP/HIA coordinator assist process for Private Access once the Privacy Taxonomy recommendations have been implemented.

Phase 2 Recommendations:

1. Ensure that Active Privacy solutions support or provide Private Access functionality

3.2.4 Acquisition Criteria

The following are recommended implementation steps and considerations regarding guidance for Software Acquisition and Software Development

Phase 1 Recommendations:

1. Adopt the Privacy Design Guidance for use as a checklist in Software Acquisition
   - For both general and privacy-enhancing software.
   - Incorporate as a subset of existing selection criteria under the GAEA Technology Standards.
   - Especially focus on compatibility with Identity Key and P3P compatibility guidelines where appropriate.

2. Adopt the Privacy Design Guidance for use as a checklist in Software Development
   - Incorporate into checkpoints in the current software development life cycle – especially in the planning and design phases, through the GAEA Alignment & Compliance Process.
   - Especially focus on compatibility with Identity Key and P3P compatibility guidelines where appropriate.
3. Maintain the Privacy Design Guidance via the existing **GAEA Vitality Process**
   - The **Privacy Framework Advisory Committee** should play a key role in approving changes

Phase 2 Recommendations:

4. Over time, concepts can be strengthened into standards as it becomes apparent that particular guidance elements are critical in a GoA context and are also realistic either from a development or acquisition perspective.

3.2.5 Consent and Choice

The following are recommended implementation steps and considerations regarding guidance on Consent and Choice

Phase 1 Recommendations:

1. Ensure Consent and Choice options are comprehensively covered in the first published version of the Privacy Taxonomy
   - Both situations where consent is required and situations where it is explicitly not required can be covered under the **Condition** dimension of the taxonomy. This dimension can also capture explicit authorities to collect personal information
   - Adoption of the Privacy Taxonomy through metadata tables should ensure that consent conditions are associated with each database containing personal information
2. Adopt the Consent and Choice measures specified in the Privacy Design Guidance for electronic capture and storage of consent values
3. Take a pro-active approach to defining how service preferences will be handled
   - The Privacy Architecture contains some initial thoughts on standardizing collection and storage of service preferences, which should be developed into firm guidance. This will become increasingly important as Electronic Service Delivery becomes more pervasive.

Phase 2 Recommendations:

1. As electronic signature regulations unfold, integrate the specific steps required for electronic proof of consent into the Privacy Architecture (will likely be an additional element of the Privacy Design Guidance).
2. Evaluate the applicability of specialized consent handling components that fit with the conceptual Active Privacy architecture as they become available

3.2.6 Access Control

The following are recommended implementation steps and considerations regarding guidance for Access Control:

1. Use the identified access control rule elements as a check list item in Software Design and Software Acquisition (note these are already covered under the Acquisition recommendations but are included here again for consistency and to provide more detail):
   - Key rule elements such as Purpose, Recipient etc. should be possible to be uniquely derived from application transaction information (ex: function x is equivalent to a disclosure to a collection agency)
   - Applications should write this same information into audit logs so that privacy compliance may be manually verified if needed (note, this would not be necessary once Active Privacy is in place).
2. Use the identified access control rule elements to create a set of pseudo rules for inclusion in the metadata for every database containing PI. These rules would essentially describe the privacy policy for describing PI collection, use, disclosure and retention. (Note these are already covered under the Data Classification recommendations but are included here again for consistency and to provide more detail)

3.3 Horizon 3 Requirement Recommendations

3.3.1 Use of Technology to Monitor Privacy Compliance & Enforce Privacy Rules

The following are recommended implementation steps and considerations regarding guidance for Use of Technology to Monitor Privacy Compliance and Enforce Privacy Rules:

Phase 1 Recommendations:

1. Adoption of the Identity Keys scheme, the Privacy Taxonomy and the Privacy Design Guidance are the prerequisites for the eventual implementation of monitoring and enforcement technology. Adoption of P3P would also be a supportive step.

Phase 2 Recommendations:

1. Evaluate and prioritize the potential capabilities that Active Privacy may offer and begin preliminary investigation of available technology:
   - These include Privacy Enhanced Access, Privacy Vulnerability/Compliance Analysis, Data Transformation, Data Validation, Private Access, Pseudonymous Interaction, Assertion Credential Support, and Citizen Contact.
2. Select and pilot technology for high priority capabilities

3.4 Summary of Horizon 1 Recommendations

A summary of Horizon 1 Recommendations is supplied in Table 6 for convenience

Table 6: Summary of Horizon 1 Recommendations

Requirement	Summary of Recommended Actions
Terminology	**Phase 1** 1. Append the Glossary to the GAEA Glossary of Terms 2. Modify the GAEA Glossary of Terms to match that of the Privacy Glossary 3. Maintain the Glossary via the existing GAEA Vitality Process
Identity Keys	**Phase 1** 1. Adopt the concept of using hidden MBUN's to index personal information 2. Conduct one-time departmental survey to determine existing use of internal IDs and MBUNs 3. Make implementation design decisions on MBUN format and issuance 4. Begin detailed functional and operational design (table structure, protocols etc.) 5. Begin process to investigate building/acquiring the Identity Protection components 6. Choose a project to pilot Internal ID use 7. Choose a project to pilot Federated ID use **Phase 2** 1. Adopt the ID keys concept for production use and set up processes to register identity domains, Online Services (applications) and Identity protection components
Data Classification	**Phase 1** 1. Adopt the Taxonomy as a government-wide standard under GAEA 2. Make implementation design decisions on notation and uniqueness of codes by level 3. Fully populate the GoA level of the Taxonomy for initial release 4. Create a simple tool (e.g. a Spreadsheet) that departments can use to quickly create specific meta-data files for describing database content and policy 5. Maintain the Taxonomy via the existing GAEA Vitality Process **Phase 2** 1. Begin to use the Taxonomy as a basis for implementing Active Privacy 2. Begin looking at the inclusion of unstructured and offline information
Data Sharing, Re-Use and Placement	**Phase 1** 1. Pilot proposed process with groups sharing PI through a Data Sharing agreement **Phase 2** 1. Adopt the process mandatory for all programs contemplating a Data Sharing Agreement 2. Monitor process output to see if a requirement for privacy-enhanced security and encryption starts to emerge 3. Consider a cross-government initiative to establish the possibility of sharing basic contact and preference information 4. Review case history of decisions to revisit the question of a static model for placement