Privacy & Technology

- **Premise:** *Technology can mitigate privacy risks, not just create them.*

- PIAs are required to identify privacy risks, but PIAs alone don’t mitigate risks.

- Policy and procedures are essential for general direction
  - but unless they’re machine readable they’re of limited use in IT applications.

- Privacy issues are made more complex by increasingly sophisticated information management. Effective privacy protection demands more technical solutions.

- **Privacy by Design** requires consistent design standards across the enterprise. Privacy architecture provides such standards.
“Architecture is defined by the recommended practice as the fundamental organization of a system, embodied in its components, their relationships to each other and the environment, and the principles governing its design and evolution.”

ANSI/IEEE Standard 1471-2000,
Recommended Practice for Architectural Description of Software-Intensive Systems
Architecture Relationships

Business | Data | Application | Technology

Privacy

Security
Privacy Architecture Topics

- Horizon 1 (fully addressed)
  - **Terminology** – a common language for discussing privacy requirements, issues and solutions
  - **Identification Keys** - how will data subjects be uniquely identified?
  - **Data Classification** - how should personal information or its uses be classified?
  - **Data Sharing, Re-Use and Placement** – to what extent can personal information be shared between departments and where should it be stored?
  - **Data Transformation** - guidance for rendering data anonymous, at varying levels of anonymity
Privacy Architecture Topics

Topics for future development

- User Interface
- Data Subject Access to Data
- Software Acquisition Criteria
- Consent and Choice
- Access Control

- Use of Technology to Enforce Privacy Rules
- Use of Technology to Monitor Privacy Compliance
Key Concepts: *Identity Domains*

- **Public Body**
  - Program A
    - IID-1
  - Program B
    - IID-2

**Identity Domains**
Key Concepts: Data Architecture
Terminology: Privacy Glossary

- Defines terms and phrases relevant to PA development
- Integrated with broader GAEA glossary
- Legal and policy definitions take precedence
  - The term ‘record’ has different meanings in the FOIP/HIA and ICT worlds. Both are recognized, but the FOIP meaning is adopted because it is defined in law.
Identification Keys:  
**ID Key Scheme**

Features:

- Meaningful identifiers are separated from personal attributes
- Application databases contain only depersonalized information
- Identifiers are stored separately under high security
- Repersonalization is subject to both policy and software controls

- “Federated identifiers” (FIDs) provide a control point for data sharing
  - FIDs must be enabled to share any given data element between any given pair of identity domains. Authorization is required.
- Thereafter, authorized data sharing by applications can occur automatically, under the control of the ID Protection Component.
ID Key Scheme

- **Public Identifier** (PID): used by individuals to identify themselves.
  - E.g., user ID, driver’s license number, personal health number

- **Internal Identifier** (IID): used by applications *within a domain* to attach depersonalized individual information to an identity. (MBUN – Meaningless But Unique Number)

- **Federated Identifier** (FID): used by applications *in different domains* to exchange or share information about an individual, *if authorized to do so under FOIP or HIA*. (MBUN)
### ID Key Scheme

**Separation of identity and attributes**

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

<table>
<thead>
<tr>
<th>Individual</th>
<th>Employee #</th>
<th>Internal ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homer Simpson</td>
<td>123 456 789</td>
<td>40596342</td>
</tr>
<tr>
<td>Joe Quimby</td>
<td>987 654 321</td>
<td>61036229</td>
</tr>
<tr>
<td>Semour Skinner</td>
<td>827 364 591</td>
<td>00836291</td>
</tr>
<tr>
<td>Sally Skinner</td>
<td>837 234 003</td>
<td>39570093</td>
</tr>
<tr>
<td>Ralph Wiggum</td>
<td>007 461 101</td>
<td>39608801</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Salary</th>
<th>Rating</th>
<th>Internal ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>$45K</td>
<td>4</td>
<td>40596342</td>
</tr>
<tr>
<td>$250K</td>
<td>2</td>
<td>61036229</td>
</tr>
<tr>
<td>$95K</td>
<td>1</td>
<td>00836291</td>
</tr>
<tr>
<td>$97K</td>
<td>2</td>
<td>39570093</td>
</tr>
<tr>
<td>$2K</td>
<td>3</td>
<td>39608801</td>
</tr>
</tbody>
</table>
Use of PID, IID and Protected Mapping Service unhook direct connection of PI with the externally known ID.

- PID = a Public ID (ex: a userid)
- IID = an Internal ID (an MBUN)
ID Key Scheme

- ID Protection Component (IDPC):
  - Issues and controls Internal IDs, Federated IDs
  - Enables/disables Federated IDs for sharing
  - Controls linkages between Public IDs, Internal IDs and Federated IDs
  - Stores and maintains ID tables
  - Is located in the restricted zone of the Security Architecture (highest security)
  - May be one or more ID protection components
    - One is better for control and administration
    - Multiple (e.g., one per public body) may be better for performance & scalability
    - Multiple would require central service for federated ID management across GoA
IDPC: Layered Protection

External Destinations

ID Protection Component (GoA)
- Depersonalization
- De-Identification Services
- Re-Identification Services
- Privacy Linking Services
- ID Linking Controls
   (Highly Secured Services)

ID / Data Linking
## ID Key Scheme

### ID Protection Component Example: Public ID

**Utility functions:**
- **IID creation**
- **IID linkage**
- **FID enable/disable**
- **Administration functions**

<table>
<thead>
<tr>
<th>PID (e.g., Driver's License Number)</th>
<th>Internal ID #1</th>
<th>Internal ID #2</th>
<th>Federated ID</th>
</tr>
</thead>
<tbody>
<tr>
<td>01234-010</td>
<td>234561789092</td>
<td>183927038109</td>
<td>982401765204</td>
</tr>
<tr>
<td>12345-123</td>
<td>97870734</td>
<td>087439874298</td>
<td>774374387428</td>
</tr>
<tr>
<td>23456-001</td>
<td>198987437474</td>
<td>210943723987</td>
<td>906347864207</td>
</tr>
</tbody>
</table>
ID Key Scheme Overview

PIDS for Customer interface → IIDS for Depersonalization

Program A
- PID
- Electronic data
- IID
- Identity Protection

Program B
- PID
- Electronic data
- IID
- Identity Protection

Customer Service

Customer

FIDS for Domain (Program) Interface
### Privacy Taxonomy

*(Personal Information Metadata Standard)*

<table>
<thead>
<tr>
<th>Name</th>
<th>SIN</th>
<th>Address</th>
<th>Age</th>
<th>Salary</th>
</tr>
</thead>
<tbody>
<tr>
<td>Homer Simpson</td>
<td>123 456 789</td>
<td>12 My Street</td>
<td>47</td>
<td>$60K</td>
</tr>
<tr>
<td>Seymour Skinner</td>
<td>372 809 875</td>
<td>245 Elm Avenue</td>
<td>52</td>
<td>$92K</td>
</tr>
<tr>
<td>Ned Flanders</td>
<td>375 059 354</td>
<td>11 My Street</td>
<td>45</td>
<td>$85K</td>
</tr>
</tbody>
</table>

In English this might say

“this table contains names, social insurance numbers, addresses, ages and salaries obtained from the individual. . . ”

In the notation of the taxonomy it might be something like:

“SRC=IND”, “CAT=NAM,SIN,ADD,AGE,SAL”, . . .

*To be implemented via metadata at the table and column levels.*
Privacy Taxonomy

Policy Dimensions

Intent
- Action
- Recipients
- Purpose

Conditions
- Conditions

Consequences
- Obligation
- Retention
- Security

Data Dimensions
- Category
- Identity
- Source

P3P Derived
EPAL Derived
Created for the Privacy Architecture
Subject to <conditions> and <security> requirements, allow <action> for <purpose> where data are <category> about <identity> from <source> to <recipients>, with <obligations> and <retention> requirements.
1. Data Sharing Analysis

Initial data sharing analysis based on business requirements, to create a potential sharing status.

<table>
<thead>
<tr>
<th>Personal Information Category</th>
<th>Program A</th>
<th>Program B</th>
<th>Program C</th>
<th>Potentially Shareable Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
<td>x</td>
</tr>
<tr>
<td>C2</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>✓</td>
</tr>
<tr>
<td>C3</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>x</td>
</tr>
<tr>
<td>C4</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td>✓</td>
</tr>
<tr>
<td>C5</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>✓</td>
</tr>
</tbody>
</table>

2. Privacy Impact Assessment

Because so many factors impinge on data sharing, potential sharing status must be confirmed by a PIA to ensure compliance with privacy law and policy.

Standard PIA template for ICT projects may require supplementary items to address band placement.
3. Data Placement Decision

If supported by a PIA, shareable data might be placed in Band 1** of the Data Architecture for use across government.

FIDs would be enabled to implement sharing, on a pair-by-pair basis.

**Band 1 = shared across GoA; Band 2 = ministry specific**
Data Transformation: Privacy Transformation Techniques

- Applications for privacy transformation (all related to PI disclosure):
  - Researcher requests
  - Data warehousing
  - Routine disclosure & active dissemination
  - Reporting functions
    - Screen display
    - Print functions
  - Role-specific access & display
Privacy Transformation Techniques:
Transformation Decision Process Summary

1. Don’t need individualized PI:
   - **Aggregation**: Aggregate statistics are computed and published (e.g. average of 9 consecutive records). Recommend minimum cell size of 5 as a default.

2. Need individualized PI, but don’t need complete PI:
   - **Reduction**: One or more fields are removed entirely (usually identifiers)
   - **Suppression**: Parts of a field are removed (e.g., removing the last three digits of a postal code or the day and month of a birthdate).

3. Need complete PI, but don’t need precise PI:
   - **Generalization**: Specific fields or field values are replaced with generalized fields or values (e.g. replacing an age with a range of ages or an assertion such as > 19 years old or even just a category description)
   - **Perturbation**: Field values are perturbed or blurred using a statistical technique (e.g., adding a random number between -10 and +10 to an age)

4. Need precise PI (no transformation):
   - **PIA**: Do PIA with intent to release identifying information.
# Privacy Design Guidance

<table>
<thead>
<tr>
<th>ID / Category / Related Requirements</th>
<th>Guidance</th>
<th>Comments and Rationale</th>
</tr>
</thead>
</table>
| 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 supports the Privacy 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 Privacy Principles of openness and limited collection  
• Optional PI is not so common in a GxA 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 Privacy 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. |

© Copyright Government of Alberta 2003
The Future: Horizon 3
Active Privacy

- Rule based
- Metadata dependent
- Real time
- Transparent

Privacy User Interaction
- Send requests to data subjects
- Provide access for data subjects
- Present and negotiate policies

Privacy Specific communications

Privacy Directory and Security
- Manage Identity Protection
- Manage Pseudonyms and Credentials
- Manage Attribute Exchange

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

Privacy Support Tools
- Policy Editors
- Log Analyzers
- Vulnerability Checkers
- Inventory Tools

Privacy Services
- Manage policy
- Record requests and decisions
- Manage privacy obligations

Data Users

Individuals

Personal Information

Policy Management and Tracking
Overview Report

Available at

www.sharp.gov.ab.ca/ppa/
Alec Campbell
Executive Director, Privacy Policy & Assessment
Office of the Chief Information Officer
Government of Alberta

9th floor, John E. Brownlee Building, 10365 – 97 Street
Edmonton, AB, Canada  T5J 3W7
Tel: 780-427-3986  Fax: 780-422-0956
Email: alec.campbell@gov.ab.ca

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.”

Privacy Architecture on the Web:

http://www.sharp.gov.ab.ca/ppa
http://www.sharp.gov.ab.ca/itm_projects/gaea