Policy Authoring

Matthew Dunlop
Policy Authoring

What is privacy?

- The right of an individual to control information about themselves
Do people care about privacy?

- In 1999, 78% of people surveyed refused to provide personal information due to concern of misuse.
- In 2000, 50% of people surveyed routinely provide false personal information.
- In 2004, 94% of people surveyed believe the benefit gained does not outweigh the cost of sharing personal information.
Objectives

- Identify organizational privacy requirements
- Identify approaches that address privacy requirements
- Design and validate a prototype for flexible and simple privacy policy creation
Key privacy design concepts

1. One integrated solution for an organization
2. Privacy functionality separated from application code
3. Support an appropriate level of granularity
4. Work with both structured and unstructured information
5. Simple and flexible privacy functionality
Abstract privacy architecture

- Privacy Policy Authoring Utility
- Machine Readable Privacy Policy
- Audit Mechanism
- Policy Enforcement Mechanism
- Applications
- Privacy Log
- Data Stores
Integrated solution techniques

- Creation of a common set of privacy utilities
- Creation of a single system that acts as a personal information “vault”

Not Met Yet
Separated functionality techniques & techniques for maintaining granularity

- Hippocratic database
- Tivoli Privacy Manager

Met
Structured/unstructured info handling & simple and flexible functionality

- No approaches exist

Not Met
Expanded abstract privacy architecture

[Diagram showing privacy policy creation utility, machine readable privacy policy, internal privacy audit utility, policy enforcement engine, and privacy log.]
SPARCLE

- **Server Privacy ARchitecture and CapabiLity Enablement**

- **Goals**
  - Create understandable privacy policies
  - Link written privacy policies with their implementation
  - Monitor enforcement of policies
Natural language policy creation

- **Policy Name**: Customer Template - Finance
- **Domain**: Finance
- **Created Date**: 2005-08-23
- **Last Modified Date**: 2005-10-08

**Rule Guide**

Example Rule Guide:

1. Financial consultants can collect and use customer name for the purpose of confirming identity.
2. Financial analysts can use customer accounts to make loan decisions.
3. Management can report customer transactions if required by law.

**Policy Information**

- **Policy Text Editing Area**
- **Invoke Parser**
- **Save and Continue**
Structured policy creation

- Original Policy Rule Text
- Policy Rules Reconstructed From Elements Identified by Parser
- Policy Elements For Selected Rule
Viewing privacy policy rules

<table>
<thead>
<tr>
<th>Customer accounts</th>
<th>Financial analysts</th>
<th>Financial consultants</th>
<th>Management</th>
</tr>
</thead>
<tbody>
<tr>
<td>Customer name</td>
<td>No Access Allowed</td>
<td>No Access Allowed</td>
<td>No Access Allowed</td>
</tr>
<tr>
<td>Customer transactions</td>
<td>No Access Allowed</td>
<td>No Access Allowed</td>
<td>No Access Allowed</td>
</tr>
</tbody>
</table>

- 2. Can use for the purpose of making loan decisions
- 1. Can collect or use for the purpose of confirming identity
- 3. Can report for the purpose of None Selected if required by law
Natural language parsing

- Uses a shallow parser
  - Identify syntactic structures (e.g. nouns, verbs)
  - Use grammars to choose desired text based on speech patterns
- SPARCLE defines five grammars
  - User categories
  - Actions
  - data categories
  - Purposes
  - Conditions/obligations
Natural language parsing example

Pharmacists can use social security number

User Category  Action  Data Category

To confirm identity if the customer registers.

Purpose

Condition or Obligation
Parsing Accuracy

- Conservative
  - 86% precision
  - 88% recall

- Liberal
  - 95% precision
  - 97% recall
Prototype testing

- Initial 2004 tests of SPARCLE were favorable
  - High to very high on their seven point scale

- Improvements after testing:
  - Import pre-existing privacy policies
  - Use privacy policy templates as a starting point
  - Improved readability of the table view
Implementation testing

- 2005 implementation tests were mostly favorable
  - High to very high on their seven point scale

- Participants were less favorable about pre-processing time
  - Desire for “no additional work” before inputting rules
  - Need better language parser
Subsequent Improvements

- In 2008, the authors looked at assisting policy authors in writing policies
  - “Evaluating assistance of natural language policy authoring” (SOUPS 2008)
- No improvement was made to the language parser
- Hypothesized that syntax highlighting would improve authoring
New Authoring Page

Rule Workspace

Customer Service Representatives and Pharmacists can only read social security number for the purpose of verifying customer's identity.

Example: [User Category(ies)] can [Action(s)] [Data Category(ies)] for the purpose(s) of [Purpose(s)] if [(optional)] Condition(s).

1. Customer Service Representatives can collect full name, shipping address, telephone number, date of birth, gender, marital status, Social Security Number, and current medications information to fulfill customer orders.

2. Customer Service Representatives and Pharmacists can only read social security number for the purpose of verifying customer's identity.

3. Pharmacists or Pharmacy Technicians can use the customer's current medication information to identify possible drug interactions if they are processing a new order.
Why syntax highlighting failed

- Immediate feedback caused users to stop mid-process to correct mistakes
  - Interrupted verbalization
  - Interrupted recording of ideas

- Recommended fix
  - Move syntax highlighting to the translation page
## Summary

<table>
<thead>
<tr>
<th>Good</th>
<th>Improve</th>
</tr>
</thead>
<tbody>
<tr>
<td>- Improves on P3P by guaranteeing policy enforcement</td>
<td>- Accuracy of policy parsing</td>
</tr>
<tr>
<td>- Provides interface usable by both IT and non-IT professionals</td>
<td>- Preprocessing time</td>
</tr>
<tr>
<td>- Good policy visualization</td>
<td>- No results on accuracy of machine-translated policies</td>
</tr>
<tr>
<td></td>
<td>- and preventing or granting access</td>
</tr>
</tbody>
</table>
Discussion

- Will the overhead of designing policies using SPARCLE discourage its adoption?
- Will organizations want privacy policies with guaranteed enforcement?
- Is 86-88% policy translation accuracy okay?
- What about policies that can’t be defined within the context of SPARCLE?