Policy Management of Enterprise Systems: A Requirements Study*

Pranam Kolari, Tim Finin, Yelena Yesha
University of Maryland, Baltimore County
Ebiquity Research Group
1000 Hilltop Circle, Baltimore, MD
{kolari1, finin, yeyesha}@umbc.edu

Kelly Lyons, Jen Hawkins, Stephen Perelgut
IBM Toronto Labs
8200 Warden Avenue
Markham, ON, Canada
{klyons, jlhawkin, perelgut}@ca.ibm.com

Abstract

Policy enabled applications are being increasingly employed to support responsive Information Technology services. In competitive business environments, such services increase adaptability of both software and the processes they implement through externalized business and security logic. Over the last decade this has driven both industry and academia to contribute to policy research and engineering, by developing specification languages, frameworks and toolkits. Since this work has typically been applied to and evaluated using new enterprise solutions, policy management for existing applications has been less well studied. In this paper we share our experiences on policy enabling an existing web based solution, together with identifying new policy enabling requirements from a specific class of enterprise systems.

1. Introduction

Policy enabled systems have been successful in various domains ranging from network infrastructure, multi-agent to distributed systems in general. Most of these real world systems leverage the benefits of policy based access control. However, policy management in enterprise information systems for a combination of access control and business policy management is not well studied. We hence analyze the exact requirements of an existing enterprise information system through a study of a widely used application.

CASSIS is an Information System (IS) developed by CAS (Center for Advanced Studies) [8], a collaborative research organization within IBM. CAS is spread across five countries, and as a whole, collaborates with university faculty and PhD students to apply their research to IBM products and processes. Towards managing this relationship, CASSIS is used by CAS in a year round effort from accepting proposals to project monitoring. Research proposals submitted by *Academic Researcher*s encapsulate information common to typical proposals like principal investigators, area of work, collaborators, requested funds, duration etc. CAS is responsible for these submitted proposals, primarily through the *CAS Head* and the *CRSM* (CAS Research Staff Member) who make decisions that include identifying *reviewers*, *evaluators* and sponsoring departments within IBM, to the eventual acceptance or rejection of the proposal. The overall nature of proposal content, and the workflow used by various role players, makes CASSIS an interesting case study for policy enablement.

We summarize our primary contributions as (i) Our report on discussion with users and management of an existing enterprise system is useful for business rules and organization modeling across other organizations and systems, (ii) Our completed work on privacy enablement provides a more adaptable solution with minimal changes to the existing infrastructure. (iii) Our study on business policy enablement of a specific class of enterprise IT systems formalizes the notion of policy management using knowledge on the Web and details what it entails.

2. Management and User Requirements

CAS is an organization run using a rotational management system, in which employees of IBM are assigned to CAS over varying periods of time. As a result, CASSIS requirements evolve and adapt over time, driving changes to its design and implementation. One of the primary goals of the management was to make the system easily adaptable through the specification, management and enforcement of policies, thus reducing the monetary cost and human resources required to maintain CASSIS in a competitive and increasingly stringent business environment. This primary requirement led to viewing policies from the privacy and business perspective, and constraints included providing alignment with existing efforts in the organization so

*This work is supported in part by IBM Toronto Lab Center of Advanced Studies, Canada and by NSF awards NSF-ITR-IIS-0326460 and NSF-ITR-IDM-0219649

as to maximize overlap, minimize redundancy and enable compatibility with similar systems across IBM. We first analyze privacy policy requirements before discussing business policies.

CASSIS is organized as a role-based application in which users are responsible for different aspects of the workflow based on their current roles. All fields in the proposal are visible to the primary role players, *CRSM* and *CAS Head*. *Reviewers* and *Evaluators* do not get to see reviews and evaluations of each other, and reviews and evaluations are not open to *Academic Researchers*. The management did not have any new requirements for making certain fields accessible to specific role players. However, the priorities of the management, in the order of precedence, included: (i) privacy requirements should be easily adaptable, (ii) privacy practices must be transparent to users, (iii) the application should enable automated compliance verification to relevant privacy rules within IBM as well as laws of the host country (as and when they are made statutory). The first of the three requirements had to be addressed immediately from the policy enablement perspective. The last two requirements would lead to enabling translation of privacy policies to P3P (or a similar vocabulary [6, 3]) and an enterprise specific vocabulary or ontology [4] for accountability to statutory and organization specific laws and rules.

Though CASSIS consists of many role players, we considered business policies as viewed by one user type, the CRSM. Business policies influence the way a proposal is viewed during its lifecycle in the CASSIS workflow. Such policies will eventually be used to consider proposals with additional interest. Some of the important facets of the business policy include: (i) **Closeness to IBM Participants**, in which past collaboration experiences of participants (IBM Employees) in the proposal are analyzed semi-automatically (based on co-published papers, personal communication etc.), (ii) **Goals of the sponsoring group**, wherein the sponsoring groups specify certain areas of research collaboration as important to IBM. For instance, IBM's push towards Autonomic Computing would consider Policy related research to be of high short-term interest, and (iii) **Connections and Collaborations**, of principal investigators with other IBM employees, or collaboration history of the principal investigators in academia and their results (through publications, patents etc.) are considered.

These policies when made explicit act as recommendations for role players, enabling them to comply with CAS business policies during their actions in the workflow. Policy enabling the business hence requires addressing certain important issues. The primary challenge is to identify what part of the business policy can be made explicit, what knowledge bases do they operate on and what this entails, and how such explicit business policies can be incorporated through externalized business logic for CASSIS.

3. Privacy Policy Enablement

PMAC[1] stands for Policy Management Autonomic Computing and is a policy driven framework aimed at realizing IBM's Autonomic Computing vision [5]. PMAC provides a toolkit complete with a policy language, editor, analysis system and deployment capabilities. It uses ACPL (Autonomic Computing Policy Language) as the underlying policy language. A key feature of ACPL is its notion of using *business value* to specify priorities between policies. In order to policy enable the privacy aspects, we employed policies encoded in ACPL and enforced using the PMAC framework. The primary reasons behind our use of PMAC as the policy enabling toolkit were its efficiency, compatibility with the existing production environment of CASSIS, excellent internal support, simplicity of the policy language and clear development architecture.

We used solicited synchronous policy guidance from the Autonomic Manager using the provided Java API. The management specifies privacy policies of CASSIS offline and deploys them on PMAC using the policy editor. This enables independent changes to the privacy policy without mandating changes to the actual implementation, one of the key advantages of policy enablement. Such deployed policies are enforced by the CASSIS web application when role players work on the proposal using various views. A partial snapshot of one such policy is shown in Figure 1. It shows an individual privacy rule for the *Reviewer* role specifying constraints on viewing the fields of the proposal (*CassisProposalScope* scope) through the *read* and *mask* string constants. Similar rules can be encoded for all fields of the proposal for different role players.

In addition, ACPL's ability to use business value as a means of specifying priorities between rules (in this case - access control) could also have been exploited. While this was not a current requirement, the management immediately noticed the usefulness of using such prioritized policies to override defaults in their system, by selectively enabling access to certain fields for important reviewers and evaluators.

4. Business Policy Enablement

Business Policy Enabling CASSIS put forward new challenges, which in turn made the exercise of requirements gathering particularly interesting. Two of the key issues included (i) functioning and decision-making process in CASSIS is influenced not only by local knowledge but also by knowledge external to the CASSIS (ii) unlike privacy (security) policies and action-reaction policies employed in other domains, the result of policy rules are recommendations which guide actions of the actors.

[1] http://www.alphaworks.ibm.com/tech/pmac/

```
<acpl:Policy
decisionName="CassisProposalPermission"
      policyEnabled="true"
policyName="ProposalPermissionPolicy"
...
<acpl:Condition>
<exp:Equal>
  <exp:PropertySensor propertyName="role" />
<exp:StringConstant>
  <Value>reviewer</Value>
</exp:StringConstant>
</exp:Equal>
</acpl:Condition>
<acpl:Decision>
<acpl:Result>
<acpl:Property propertyName="permissions">
<exp:CompositeDataConstant>
<permissions>
  <fundrequest>read</fundrequest>
  <fundapprove>mask</fundapprove>
</permissions>
</exp:CompositeDataConstant>
</acpl:Property>
</acpl:Result>
</acpl:Decision>
<acpl:BusinessValue>
  <Importance>6</Importance>
</acpl:BusinessValue>
<acpl:Scope>
<acpl:StringScope>
  <Value>CassisProposalScope</Value>
</acpl:StringScope>
</acpl:Scope>
</acpl:Policy>
```

Figure 1. *This ACPL description encodes an individual privacy rule for the reviewer role specifying constraints on viewing the fields of the artifact through the 'read' and 'mask' string constants.*

Even though policy enabled software systems have been engineered for over a decade, knowledge driving these policies is local, readily available from the application standpoint, and is often not considered as a first class citizen i.e. at the same importance level as policy specification. Recent research [10][7] has initiated a focus on this aspect by enriching the existing knowledge base through reinforcement learning and mining. The business requirement of CASSIS revealed another interesting dimension, where knowledge openly available on the World Wide Web, in combination with Intranet knowledge, and that of the application drives business decisions through recommendations. This is different from knowledge useful for privacy enablement that involved only role of the actor and fields of the proposal. The challenge then is to identify useful knowledge sources, integrate them, analyze properties of integrated knowledge to clarify what sets it apart, and finally explore how this effects business policy enablement, from the tooling, policy language and engineering perspective. We first discuss the nature of existing knowledge sources before analyzing their implications for policy enablement.

IBM, being an organization of over three hundred thousand employees, has a fairly sophisticated and well-organized Intranet. One of the popular services on the Intranet is the IBM Blue Pages [2], a directory of all employees. In addition to listing contact details, location, designation and other details it also lists employee hierarchy in the management, people managed, department, skills, projects and experiences. IBM also has useful knowledge stored in databases maintained by different departments like redbooks[2], patents etc. published as different services either in the Intranet or Web scope. In addition to the Intranet, the World Wide Web is a large knowledge base by itself. For CASSIS, some of this available knowledge does have a potential impact on the business. In particular, the domain of interest for CASSIS is the large wealth of information on the academic community, specifically their ongoing research and collaborations, that is readily available through Citeseer[3] and DBLP[4].

The combined knowledge of the Intranet and the Web is highly important from the CAS stand point. What is lacking however is a way of using them together through integration to a common data model, followed by identifying what this would entail for policy enablement. It turns out that the former aspect of data integration is another current thrust within the organization. IBM is adopting ideas of the Semantic Web vision [1] by RDF (http://www.w3.org/RDF/) enabling information on their Intranet. RDF (Resource Description Framework) is the graph based data structure that underlies the Semantic Web. IBM is using RDF as a content management mechanism through the development of Luna, a framework for knowledge integration of data sources within IBM. We are currently working on integrating this knowledge with those available on the Web.

From the policy specification stand point, aggregated knowledge of the Intranet and the Web taken together has certain important characteristics: (i) Unbalanced, in that the dataset is biased towards specific academic communities, author types or geographic locations [9], (ii) Conflicting, where DBLP and Citeseer present mutually conflicting information for the same publication or author, (ii) Mutually Reinforcing, i.e. the datasets together reinforce their knowledge about a publication or author, (iv) Not completely trustworthy, by which content extraction techniques from publications sometimes make author attributions that are not necessarily correct.

While the mutually reinforcing characteristic is useful, the other attributes mandate that policy recommendations be accompanied by information as to why such a recommendation was made, and why a recommendation was ac-

[2] http://redbooks.ibm.com
[3] http://citeseer.ist.psu.edu/
[4] http://dblp.uni-trier.de/

cepted by a CAS role player. These specific characteristics of the underlying knowledge and requirements of CASSIS lead to a new model for business policies, slightly different from the well known *ECA, Event, Condition, Action* model. We hence were required to consider policies as constituting {*Event, Condition, Recommendation, Justification*}, abbreviated as *ECRJ*. We briefly introduce each entity in the context of CASSIS: (i) *Event*, is when a business policy guidance is solicited. The *Event* model in CASSIS is similar to the one found in traditional *ECA* and is fired when role players in CASSIS are involved in operations requiring policy guidance. (ii) *Condition*, is a subgraph match as knowledge driving CASSIS policies is encoded in RDF. Conditions can readily make use of graph matching constructs provided by SPARQL[5], a query language for RDF (iii) *Recommendation*, is typically a recommended action for the role player. These recommendations enable role players to take actions in the workflow that confirms to business policies. For untrustworthy knowledge bases, they are accompanied by matching conditions from (ii), and (iv) *Justification* to the policy system i.e. if a role player accepts a recommendation, the condition that best justifies the action is indicated to the policy system through a feedback loop.

As readily available structured knowledge becomes common in many other domains and the responsiveness of enterprise systems to customer environment dynamics becomes an important requirement, the utility of this new policy model will become increasingly evident.

5. Discussion

Our work on eliciting requirements for, and adding explicit policy management to, an exiting, well-used applications leads to several observations. We believe these observations and practical requirements are valid and applicable for policy enablement of applications in other, related domains.

Privacy policy enablement of CASSIS required that certain long-term goals be also considered. Specifically, this involved translation across policy languages which will hold the key in further automating the business. From the perspective of most enterprise systems, translation is required along two dimensions: to user-specific privacy promises through P3P and to an enterprise-wide common policy vocabulary. Our encoding of data fields in ACPL makes no attempt to connect data fields to well known schemas syntactically or to connect associated rules to other rule languages semantically. Consequently, the privacy policy is only understood by the application at the level of CASSIS. We hence believe that mapping (by the policy specification engineer/management when policies are created) has to be mandated to enable better interoperability. Similar recent efforts by the World Wide Web Consortium[6] show the importance of this issue in the broader context.

Policy enabling the business raised some important considerations. Firstly, the knowledge driving business policies is becoming increasingly important. The broader area of business policy management driven by Web knowledge is becoming feasible, thanks to the exponential growth in structured knowledge publicly available on the Web. Secondly, the untrustworthy nature of such knowledge requires that aspects like recommendations and justifications be considered as entities of the policy model. We believe that business policy management on web scale knowledge will grow in importance in the coming years as policy enablement becomes increasingly common. Our current focus is completely driven by business policy enablement, and hence on formalizing the *ECRJ* model.

References

[1] T. Berners-Lee, J. Hendler, and O. Lassila. The semantic web. *Scientific American*, 279(5):34–43, May 2001.

[2] http://www.intranetjournal.com/articles/200209/ij_09_25_02a.html. Ibm's own intranet: Saving big blue millions.

[3] L. Kagal, T. W. Finin, and A. Joshi. A policy based approach to security for the semantic web. In *International Semantic Web Conference*, pages 402–418, 2003.

[4] G. Karjoth, M. Schunter, and E. V. Herreweghen. Translating privacy practices into privacy promises -how to promise what you can keep. In *POLICY*, pages 135–146, 2003.

[5] J. O. Kephart and D. M. Chess. The vision of autonomic computing. *IEEE Computer*, 36(1):41–50, 2003.

[6] P. Kolari, L. Ding, S. Ganjugunte, A. Joshi, T. W. Finin, and L. Kagal. Enhancing web privacy protection through declarative policies. In *IEEE 6th International Workshop on Policies for Distributed Systems and Networks*, pages 57–66, 2005.

[7] T. Li, F. Liang, S. Ma, and W. Peng. An integrated framework on mining logs files for computing system management. In *KDD '05: Proceeding of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining*, pages 776–781, New York, NY, USA, 2005. ACM Press.

[8] S. G. Perelgut, G. M. Silberman, K. A. Lyons, and K. L. Bennet. Overview: the centre for advanced studies. *IBM Systems Journal*, 36(4):474–488, 1997.

[9] V. Petricek, I. J. Cox, H. Han, I. G. Councill, and C. L. Giles. Modeling the author bias between two on-line computer science citation databases. In *WWW (Special interest tracks and posters)*, pages 1062–1063, 2005.

[10] J. Strassner and B. J. Menich. Philosophy and methodology for knowledge discovery in autonomic computing systems. In *DEXA Workshops*, pages 738–743, 2005.

[5] http://www.w3.org/TR/rdf-sparql-query/

[6] See http://www.w3.org/2005/rules/wg/charter/.