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. We first detail policy enablement requirements and constraints from the perspective of management and users of the application. We then present completed and ongoing work, our observations, and future directions.

1. Introduction

Abstracting out the policy related constraints and aspects of services, grid infrastructures and software agents has been an area of active research in the past decade. As theory, tools and frameworks develop and mature, IT consultants are increasingly employing policy management in enterprise systems, and referring to it with terms like *Autonomic Computing*, *On Demand Computing* and *Adaptive Enterprise*. What underlies all these IT solution infrastructures is the use of a policy driven control loop, either synchronous or asynchronous, which ensures the overall adaptability of the deployed system.

*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

Policy Managed Systems provide several advantages over traditional enterprise systems. Even though the exact benefits depend on the application domain, they can be broadly characterized into the following, listed in decreasing level of research maturity:

- **Adaptability** of the enterprise system as a whole to changes in the business policy, privacy laws or the more complex dynamics in the consumer environment through externalized business and security logic.
- **Centralized audit** of enterprise wide policies automatically to enable accountability, through merging policies employed at the (micro) department level with (macro) enterprise policies and privacy laws.
- **Accountability** to end users, of the enterprise system's privacy practices as a whole, enabled by translation to standard privacy vocabularies like P3P [12].
- **Negotiation**, primarily inter-enterprise where externalized dynamic policies are created and enforced on the fly through interaction between cross enterprise software agents.

Such systems have been deployed in various domains that includes network infrastructure, coordination in multiagent systems and the management of distributed systems in general [24]. 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 has not been well studied. In view of this, we take an approach of analyzing the exact requirements of an existing enterprise information system from both these perspectives. Such systems come with various constraints and requirements from users and management, which needs to be formalized prior to policy enabling the solution. We summarize our primary contributions as follows:

1. 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.

2. Our completed work on privacy enablement provides a more adaptable solution with minimal changes to the existing infrastructure.

3. 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.

4. Our discussion on experiences, observations and challenges from a practical perspective is useful to drive requirements from the policy tooling standpoint.

The rest of the paper is organized as follows. In section two we describe the specific enterprise system we used as a use-case for requirements gathering. Section three describes the policy enablement requirements along two axes, and defines the scope of our work. In section four we introduce PMAC, a realization of IBM's Autonomic Computing initiative. Section five reports on our completed work on privacy policy enabling an existing application. In section six we discuss Business Policy Enablement in detail and show how it is directing our future work in this area. Finally in section seven, we report on our observations on policy enabling an existing enterprise solution and present conclusions in section eight.

2. CASSIS

CASSIS is an Information System (IS) developed by CAS (Center for Advanced Studies) [21], 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 problems in IBM. CASSIS is a web based application used to manage this relationship with universities worldwide. It provides a solution for soliciting research proposals, reviewing them and monitoring accepted ones - typically a year round effort. Though the information system as a whole is complex and consists of many independent and interacting workflows on different artifacts, we consider CASSIS from a single artifact standpoint, that of a research proposal and an independent workflow it goes through during its lifecycle. We first introduce the role players (actors) in this workflow and then detail key fields in the artifact.

2.1. The Actors

The following actors work on (and are responsible for) the artifact at various points in its lifecycle:

Figure 1. *The CASSIS view for CAS Research Staff Members listing submitted proposals for the current financial year.*

- **CAS Research Staff Members** are key players in CAS, responsible for a proposal from its initiation to its final inclusion in an IBM product or solution. In addition to interacting with developers and managers inside IBM to gather requirements for the academic community, the CRSM is also responsible for identifying research in the academic community of overlapping interest. The CRSM either initiates a proposal or is contacted by academic researchers interested in applying theory they have developed into real world practical problems. The CRSM is then responsible for all aspects of the proposal which includes requesting formal proposal submission from the researcher, finding reviewers and evaluators for the proposal within IBM, identifying sponsoring IBM departments, to the eventual approval/rejection of the proposal and its management thereafter. Figure 1 depicts one of the views used by the CRSM, which lists submitted proposals for the current financial year.

- **Academic Researchers** submit proposals, and see value in working with IBM, both from the perspective of IBM's products as well as the applicability and visibility of their own research. Once a proposal is accepted, they are responsible for supervising students and interacting with the primary point of contact, the CRSM towards gathering practical requirements and scheduling time lines.

- **CAS Reviewers** are usually picked by the CRSM based on the value of the proposal to the reviewer's department. Typically, they are developers and architects of IBM products and solutions who clearly understand IBM's requirements from the academic community.

- **CAS Evaluators** vouch for the credibility of reviewer comments, and are typically higher level management in charge of the specific products. Evaluators are responsible for meta-reviews to guarantee that existing

				Servers		
116	1	2004	John Doe	Towards an Effective with IT Infrastructure	rejected	John Johny
117	1	2004	John Doe	Towards an Effective with IT Infrastructure	rejected	John Johny
118	1	2004	John Doe	Towards an Effective with IT Infrastructure	rejected	John Johny
119	1	2004	John Doe	Towards an Effective with IT Infrastructure	rejected	John Johny
120	1	2004	John Doe	Towards an Effective with IT Infrastructure	rejected	John Johny
121	1	2004	John Doe	Towards an Effective with IT Infrastructure	accepted	John
121	2	2005	John Doe	Towards an Effective with IT Infrastructure	rejected	John Doe
122	1	2004	John Doe	Towards an Effective with IT Infrastructure	approved	John Johny
123	1	2004	John Doe	Towards an Effective with IT Infrastructure	rejected	Mike Doe
124	1	2004	John Doe	Towards an Effective with IT Infrastructure	rejected	John Johny
125	1	2004	John Doe	Towards an Effective with IT Infrastructure	approved	John Johny
				Database Management		

Figure 2. *The CAS Head Proposal View listing all present and past projects (both approved and rejected) managed by CAS.*

reviews are unbiased and not prone to conflict of interest. In addition, they also provide inputs from the perspective of long term IBM strategy for their product and solution offerings.

- **CAS Head** makes the final decision as to whether a proposal has to be approved or rejected based on inputs from the CRSM, the Reviewers and Evaluators spread across IBM. Approved proposals are referred to as CAS Projects. Such approved proposals are usually funded over a period of three years, reviewed yearly and support the graduate studies of a student. Figure 2 depicts one of the views used by the CAS Head listing all present and past projects (both approved and rejected) managed by CAS.

- **Public Internet.** As the project shows results through publications, patents and reports, IBM CAS makes certain aspects of the project public, so that it is readily accessible for general use in projects outside IBM including other academic research groups. Such public reports encourage collaboration with new academic research groups, a key competency requirement of CAS.

2.2. The Artifact

The artifact is the research proposal, similar in content to proposals submitted by researchers to other funding agencies. If the proposal is accepted, the project report becomes the artifact in subsequent years. The project report has fields similar to that of the proposal. We mention only the salient attributes of the artifact here to introduce the readers to the domain we are dealing with:

- **Title and Year** are self-explanatory. The **Year** field specifies the year during which the proposal/project is going through a review process.

- **CAS Location and Country** specifies the CAS Location, and the country the researcher submitting the proposal is located in. It is very much possible that a researcher in one country submits a proposal to a CAS location in a different country.

- **PI's and their Affiliation** stand for the Principal Investigators of a proposal, a list of academic researchers who collaboratively submit the proposal.

- **CAS RSM** is the primary point of contact for the project – a research staff member who works for IBM CAS.

- **Participants**, with name, affiliation and contact information is a list of all people involved in the project - academic researchers, students and IBM employees.

- **Proposal Content** consists of project objectives, relation to other funded projects, proposed work, participation of individuals, reason for funds and recent publications. The academic researcher is expected to make a convincing argument of applicability of their work to IBM in this section.

- **Funding Requested and Approved** are based on the participation of individuals, hardware requirements and other logistics.

- **Reviewers, Evaluators and Comments** consists of the key reviewers and evaluators of the proposal as identified by the CRSM and their comments.

The complex workflow involving this artifact can be abstracted into proposal submission by academic researchers, initial screening by CRSM, identifying reviewers and evaluators by CRSM, the review and evaluation process, the final approval or rejection by the CAS Head followed by its continual monitoring.

3. Management and User Requirements

The management and users of CAS put forward certain requirements and constraints for policy enablement. 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 as to maximize overlap, minimize redundancy and enable compatibility with similar systems across IBM. It also had to be scalable and efficient so as to be useful in a production environment.

3.1. Privacy Policy Specification

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 artifact, 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 more transparent, (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. However the management made it clear that any such policy enablement ensure that it eases addressing the latter requirements eventually. The last two requirements would lead to enabling translation of privacy policies to P3P (or a similar vocabulary [17, 13]) from the user perspective and an enterprise specific vocabulary or ontology ([1, 2, 14]) for accountability to statutory and organization specific laws and rules.

3.2. Business Policy Specification

Since *business policy* is a relatively broad term, we clarify this aspect in detail. From the perspective of CAS and CASSIS, business policies map down to policies governing the very running of CAS, as driven by IBM's current policies as an organization and the business environment in general. Though CASSIS consists of many role players, we preferred to consider one user type (CRSM) and one management type (CAS Head) and their perspective of business policies important to CAS.

From the CRSM perspective, business policies influence the way a proposal is viewed during its lifecycle in the CASSIS workflow. Such policies will eventually be used to consider the proposal with additional interest. Some of the important facets of the business policy include:

- **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.)
- **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.
- **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.

An important additional responsibility of the CRSM, as shown in figure 3, is identifying key reviewers (assigned reviewers) for a proposal, a two step process. In the first step, prospective new reviewers are identified based on existing policies. This is followed by the CRSM directly contacting some of the prospective reviewers to confirm their availability for the review process, followed by moving them to the available reviewers list. The key reviewers are then identified based on various conditions that are again driven by the business policies in effect at that time. Some aspects of the business policy for identifying prospective and key reviewers include at least some of (but not limited to):

- **Domain Expertise**, of the reviewer so that it aligns with the topic of the proposal. A reviewer need not necessarily be in a department working on that area.
- **Management Function**, which decides the seniority of the employee in their department.
- **Department**, which identifies the broad area of the reviewers responsibility. This would most likely be a department close to the area of the proposal and the sponsoring IBM group.
- **Social Network**, the context of the reviewer in the social network comprising of the CRSM, PI and proposal participants.
- **Existing Key Reviewers**, which ensures that all reviewers are not picked from the same department, geography etc.

From the CAS Head perspective, business policies drives at the very least, **Identification of interesting new collaborations**, **Distribution of research topics** and **Past collaboration experiences**. In the interest of not further divulging

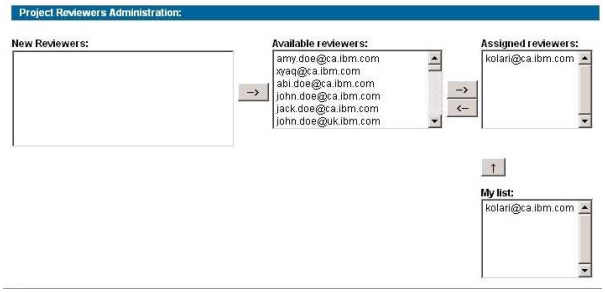

Figure 3. *Using the CRSM view, the CRSM identifies and then assigns key reviewers for a proposal.*

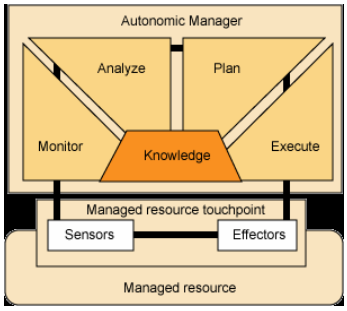

Figure 4. *IBM's Autonomic Manager model includes the MAPE (Monitor, Analyze, Plan, Execute) control loop with managed resource touch points interfacing with the policy managed application.*

the complete running of CAS, we only mention some of the criteria used by the CAS Head without going into particular details. It should however be clear to the reader that, all the above criteria are driven by the business policies currently in place at IBM. These policies when made explicit act as recommendations for role players, enabling the actors 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.

3.3. Scope

To provide an initial proof of concept of the advantages of policy enablement we decided to limit ourselves to certain aspects of CASSIS.

From the security and privacy perspective, we abstracted ourselves from the role creation process itself. This was mainly driven by business requirements, as it had historically been a static set of features of CASSIS. Hence, our privacy policy enablement approach can be viewed as effective once users are assigned to roles.

Though CASSIS as an application involves many role players we limited our study of business policy to two of the key actors - the CRSM and CAS Head. This was partly due to business demands and partly due to time and resource constraints. Our business policy enablement should be easily extensible to other role players and will be used as a guide, as the need arises.

4. Policy Enablement Tooling

The policy enablement of CASSIS required that it align and make use of ideas within IBM's vision for Autonomic Computing [16, 15]. In the autonomic computing framework, the policy enabled application is managed by one or more autonomic managers which handle externalized business and security logic. Figure 4[1] depicts one such autonomic manager constituted by the MAPE (Monitor, Analyze, Plan, Execute) control loop. Managed Resource touch points are the interaction points between the policy managed application and the autonomic manager. This interaction can be initiated by either the autonomic manager (unsolicited) or the managed resource (solicited) and can result in either synchronous or asynchronous policy guidance. Policies are based on knowledge available for the autonomic manager and almost always includes the monitor, analyze and execute steps, but could also additionally include a plan phase.

4.1. PMAC

PMAC stands for Policy Management Autonomic Computing and is a policy driven framework aimed at realizing IBM's Autonomic Computing vision. PMAC provides a toolkit complete with a policy language, editor, analysis system and deployment capabilities. It uses ACPL (the 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. Interested readers are directed to the IBM PMAC web site[2] for more information.

[1] Image courtesy of http://www.research.ibm.com/autonomic/
[2] http://www.alphaworks.ibm.com/tech/pmac/

4.2. Developer Framework

PMAC is implemented in Java and provides APIs that enable applications to request for policy guidance based on deployed policies. PMAC also has web service capabilities enabling its use by applications written across disparate platforms, development frameworks and languages.

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.

Having detailed the core requirements and introduced the domain, we next report on our work towards realizing these requirements by policy enabling CASSIS. Given the needs of the business, the management and the users we viewed policy enablement from two independent perspectives, privacy and business, and hence two independent autonomic managers (policy enablement mechanisms) address them. We first detail our completed work on privacy enablement.

5. Privacy Policy Enablement

In an earlier section we put forward the privacy policy enablement requirements from the CAS management. Since the primary focus of this paper is on requirements gathering, we detail privacy policy enablement only briefly.

Clearly, privacy specific requirements were geared towards long-term goals. In order to policy enable these privacy aspects, we employed policies encoded in ACPL and enforced using the PMAC framework. We used solicited synchronous policy guidance from the Autonomic Manager using the provided Java API. The entire architecture of privacy policy enablement is depicted in Figure 5.

The management specifies privacy policies of CASSIS offline and deploys them to the Autonomic Manager 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 actors work on the artifact using various views. A partial snapshot of one such policy is shown in figure 6. It shows an individual privacy rule for the *reviewer* role specifying constraints on viewing the fields of the artifact (*C*assisProposalScope scope) through the *read* and *mask* string constants. Similar rules can be encoded for all fields of the artifact for different role players.

The current implementation of CASSIS, based on servlets and JSPs (Java Server Pages) made interfacing the application with the Autonomic Manager relatively easy. The changes to the JSPs were trivial and excerpts of code

```
<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 6. *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.*

changes are listed for completeness in figure 7. *DecisionInput* is part of the developer API which enables requesting for policy guidance from the autonomic manager. Request includes role information of the actor and the current scope (artifact) and response is in the form of all fields viewable by the user in that context.

The JSPs implementing views for various role players allowed us easy access to user and role information without requiring hits to the database for additional contextual information. This in turn enabled us to satisfy another constraint from the management, that changes to the existing implementation be kept to a minimum. It should also be noted here that in a typical Java based web based application policy guidance can be implemented at either the servlet or JSP level. We used JSPs in our policy enablement since it appeared to be a more convenient option for this application.

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 immedi-

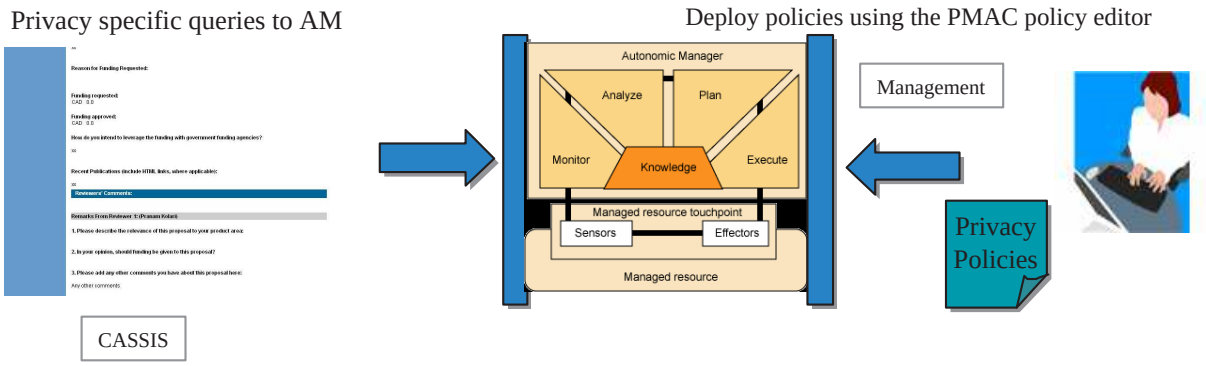

Figure 5. *The overall architecture for privacy policy enablement uses policies encoded in ACPL and enforced using the PMAC framework through guidance from the Autonomic Manager.*

```
DecisionInput   di = ...
di.putString ( "role", "reviewer");
String  fundrequest   = "mask";
Decision d =   Decider.getDecider  ()
                    .requestGuidance(di  );
if( d.noGuidance  ()) {
        ...
} else {
        ...
<% if ( fundrequest.equals("read   ")) {%>
<tr>
  <td  colspan ="2"><b>< br>
  Funding requested:</b><   br>
  <%= segmentAllInfo.getCurrency   ()%>
  ;<%= segmentAllInfo.getFundreq   ()%>
  </td>
</tr>
...
```

Figure 7. *Requests include role information of the actor and the current scope (artifact) and response is in the form of all fields viewable by the user in that context.*

ately 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.

6. 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.

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 the policy specification. Recent research [25][19] 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.

Noteworthy is the fact that this is different from knowledge useful for privacy enablement that involved only role of the actor and attributes of the artifact. 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.

6.1. Existing Intranet Knowledge

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 [11], 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 red-

books[3], patents etc. published as different services either in the Intranet or Web scope.

6.2. Web Knowledge

The World Wide Web is a large knowledge base by itself. From the CASSIS perspective some of this available public 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. To begin with we identified three such information sources - Citeseer[4], DBLP[5] and FOAF[6]. Each one of these sources [22] have their own bias and are useful in specific domains. We discovered during our analysis that both DBLP and Citeseer had significant information that the management found useful for evaluating submitted proposals and identifying new collaborations. Our study also showed that there is no sufficient overlap with the FOAF dataset (gathered using Swoogle[10]), and hence we eventually dropped it from consideration. In addition, both Citeseer[7] and DBLP[8] metadata are public datasets, readily available in structured formats and not prone to privacy issues common with data analytics.

It should be evident from our earlier discussion on CASSIS business policies that the combined knowledge of the Intranet and the Web is highly important from the CAS perspective. 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.

6.3. Luna RDF Graph Infrastructure

IBM is adopting ideas of the Semantic Web vision [3] 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. RDF enables data sharing, integration and interoperability through the use of popular vocabularies, and in addition provides a rich data source for the use of graph based mining and search techniques on enterprise wide information.

IBM is adopting RDF as a content management mechanism through the development of Luna, a framework for knowledge integration within IBM. Luna not only incorporates the IBM Blue Pages, but also additional information found in their patent servers, publication repositories and mailing listing to create a social network which complements the employee hierarchy available from Blue Pages. Luna is based on the Jena [6] toolkit[9], an open source implementation, and encourages the use of IBM Intranet data in new applications within IBM. One such application is Relescope [9], which fosters collaboration between IBM employees by exploiting social networks within IBM. From the perspective of CASSIS, we have readily available IBM Intranet knowledge and have to use it in combination with information from the Web, also available in structured formats.

Complete integration[28][8][20] is a tough problem and has been an active area of research in both the Semantic Web and Database communities. Hence an integration of Intranet and Web knowledge is to be done to an extent where it is acceptable by the CAS management and useful for CASSIS. In the interest of space and the scope of this paper we do not detail how we are currently integrating knowledge important for CASSIS.

6.4. Aggregated Knowledge & Business Policy

Aggregated knowledge of the Intranet and the Web taken together has certain important characteristics:

- Unbalanced, in that the dataset is biased towards specific academic communities, author types or geographic locations [23].
- Conflicting, where DBLP and Citeseer present mutually conflicting information for the same publication or author.
- Mutually Reinforcing, i.e. the datasets together reinforce their knowledge about a publication or author.
- 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 justifications. In other words, the management and users of CASSIS cannot completely trust recommendations without a cursory analysis of why such a recommendation was made. 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 from the perspective of {*Event, Condition, Recommendation, Justification*}, abbreviated as *ECRJ*.

[3] http://redbooks.ibm.com
[4] http://citeseer.ist.psu.edu/
[5] http://dblp.uni-trier.de/
[6] http://www.foaf-project.org/
[7] http://citeseer.ist.psu.edu/oai.html
[8] http://dblp.uni-trier.de/xml/
[9] http://jena.sourceforge.net/

6.5. Extending ECA

The *ECRJ* model for business policies consists of the following core entities. We briefly introduce each entity in the context of CASSIS:

- *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.
- *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[10], a query language for RDF. Elsewhere on the Semantic Web, similar methods of condition matching have been used in TriQ.L [4], [5] for filtering information based on trust policies.
- *Recommendation*, is typically a ranking of entities which enable prioritization over many entities under consideration. In CASSIS for instance, this could be either proposals in the workflow or prospective and key reviewers. We currently consider prioritization at three levels, namely *high*, *low* and *medium*, with scope for future extension to a higher granularity.
- *Justification*, is used to validate policy recommendations. This is a combination of the specific subgraph match for a condition and English language text associated with the fired policy rule. Such information is displayed to the actors to justify recommendations since the underlying knowledge is not completely trustworthy. The actions of the actors are driven by how convinced they are by these justifications. A similar mechanism, in the context of trust policies is used in TriQ.L for justifying filtered knowledge.

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, a need for such models will grow and their utility will become increasingly evident.

6.6. Current and Future Work

Our current focus is completely driven by business policy enablement. We are working on two areas, integration of knowledge for CASSIS and the design and implementation of a framework that enables the *ECRJ* model for business policy enablement. Though preliminary work suggests TriQ.L to be the best fit for our domain, we are also studying the feasibility of adopting other Semantic Web friendly policy frameworks [18] [7] [27][26].

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

7. Observations

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 believe such 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[11] 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, powered by syndication feeds (e.g. RSS[12]) and blogs, and the secondary text analysis services(e.g. WebFountain[13]) built on top of them. Secondly, the untrustworthy nature of such knowledge mandates that justifications accompany policy guidance, which is then analyzed by role players before taking actions. We believe that business policy management on web scale knowledge will grow in importance in the coming years as policy enablement becomes increasingly common.

8. Conclusion

We have described an initial, exploratory study on the policy management of a web-based enterprise application, driven by both business needs and constraints. Our study provides key insights into policy enablement in a specific domain. In doing so we have also reported on completed work on realizing partial solution to the problem, our con-

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

[12] http://blogs.law.harvard.edu/tech/rss

[13] http://www.almaden.ibm.com/webfountain/

tinuing work on policy enablement using web scale knowledge and related observations.

9. Acknowledgements

We would like to thank Lionel Marks, the lead developer of CASSIS at IBM for sharing development related details. Walid Rjaibi of the IBM DB2 Security Group initially pointed us in the direction of policy management for enterprise systems and the challenges therein.

References

[1] R. Agrawal, P. Bird, T. Grandison, J. Kiernan, S. Logan, and W. Rjaibi. Extending relational database systems to automatically enforce privacy policies. In *ICDE*, pages 1013–1022, 2005.

[2] A. Barth and J. C. Mitchell. Enterprise privacy promises and enforcement. In *WITS '05: Proceedings of the 2005 workshop on Issues in the theory of security*, pages 58–66, New York, NY, USA, 2005. ACM Press.

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

[4] C. Bizer, R. Cyganiak, T. Gauss, and O. Maresch. The triql.p browser: Filtering information using context-, content- and rating-based trust. In *Semantic Web and Policy Workshop at the 4th International Semantic Web Conference, Galway, Ireland*, 2005.

[5] J. J. Carroll, C. Bizer, P. J. Hayes, and P. Stickler. Named graphs, provenance and trust. In *WWW*, pages 613–622, 2005.

[6] J. J. Carroll, I. Dickinson, C. Dollin, D. Reynolds, A. Seaborne, and K. Wilkinson. Jena: implementing the semantic web recommendations. In *WWW (Alternate Track Papers & Posters)*, pages 74–83, 2004.

[7] N. Damianou, N. Dulay, E. Lupu, and M. Sloman. The ponder policy specification language. In *IEEE 4th International Workshop on Policies for Distributed Systems and Networks*, pages 18–38, 2001.

[8] A. Doan and A. Y. Halevy. Semantic-integration research in the database community. *AI Mag.*, 26(1):83–94, 2005.

[9] S. Farrell, C. S. Campbell, and S. Myagmar. Relescope: an experiment in accelerating relationships. In *CHI Extended Abstracts*, pages 1363–1366, 2005.

[10] T. W. Finin, L. Ding, R. Pan, A. Joshi, P. Kolari, A. Java, and Y. Peng. Swoogle: Searching for knowledge on the semantic web. In *AAAI*, pages 1682–1683, 2005.

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

[12] http://www.w3.org/P3P/. The w3c p3p framework.

[13] 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.

[14] 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.

[15] J. O. Kephart. Research challenges of autonomic computing. In *ICSE*, pages 15–22, 2005.

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

[17] 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.

[18] S. Lamparter, D. Oberle, and A. Eberhart. Approximating service utility from policies and value function patterns. In *POLICY*, pages 159–168, 2005.

[19] 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.

[20] N. F. Noy. Semantic integration: A survey of ontology-based approaches. *SIGMOD Record*, 33(4):65–70, 2004.

[21] 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.

[22] V. Petricek, I. J. Cox, H. Han, I. Councill, and C. L. Giles. A comparison of on-line computer science citation databases. In *9th European Conference on Research and Advanced Technology for Digital Libraries (ECDL 2005)*. Springer, 2005.

[23] 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.

[24] M. Sloman. Policy driven management for distributed systems. *Journal of Network and Systems Management*, 2:333–360, 1994.

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

[26] G. Tonti, J. M. Bradshaw, R. Jeffers, R. Montanari, N. Suri, and A. Uszok. Semantic web languages for policy representation and reasoning: A comparison of kaos, rei, and ponder. In *International Semantic Web Conference*, pages 419–437, 2003.

[27] A. Uszok, J. M. Bradshaw, R. Jeffers, N. Suri, P. J. Hayes, M. R. Breedy, L. Bunch, M. Johnson, S. Kulkarni, and J. Lott. Kaos policy and domain services: Toward a description-logic approach to policy representation, deconfliction, and enforcement. In *IEEE 4th International Workshop on Policies for Distributed Systems and Networks*, pages 93–, 2003.

[28] H. Wache, T. Vögele, U. Visser, H. Stuckenschmidt, G. Schuster, H. Neumann, and S. Hübner. Ontology-based integration of information — a survey of existing approaches. In H. Stuckenschmidt, editor, *IJCAI–01 Workshop: Ontologies and Information Sharing*, pages 108–117, 2001.