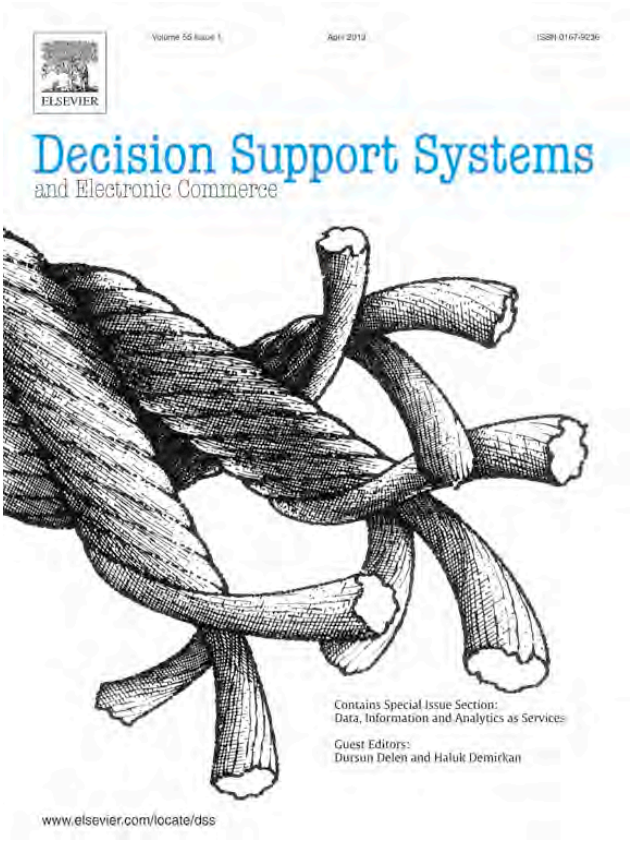


Decision Support Systems 55 (2013) 385–398

Contents lists available at SciVerse ScienceDirect

Decision Support Systems

journal homepage: www.elsevier.com/locate/dss

# Secure federation of semantic information services ☆

Benjamin Fabian [a,*], Steffen Kunz [a], Sebastian Müller [b], Oliver Günther [c]

[a] Institute of Information Systems, Humboldt-Universität zu Berlin, Spandauer Str. 1, 10178 Berlin, Germany
[b] Databases and Information Systems Group, Freie Universität Berlin, Takustr. 9, 14195 Berlin, Germany
[c] Universität Potsdam, Am Neuen Palais 10, 14469 Potsdam, Germany

ARTICLE INFO

Available online 1 June 2012

*Keywords:*
Information federation
Service orientation
Semantic web
Information security

ABSTRACT

A fundamental challenge for product-lifecycle management in collaborative value networks is to utilize the vast amount of product information available from heterogeneous sources in order to improve business analytics, decision support, and processes. This becomes even more challenging if those sources are distributed across multiple organizations. Federations of semantic information services, combining service-orientation and semantic technologies, provide a promising solution for this problem. However, without proper measures to establish information security, companies will be reluctant to join an information federation, which could lead to serious adoption barriers.

Following the design science paradigm, this paper presents general objectives and a process for designing a secure federation of semantic information services. Furthermore, new as well as established security measures are discussed. Here, our contributions include an access-control enforcement system for semantic information services and a process for modeling access-control policies across organizations. In addition, a comprehensive security architecture is presented. An implementation of the architecture in the context of an application scenario and several performance experiments demonstrate the practical viability of our approach.



## 1. Introduction

Increasing global market competition raises the pressure on corporations in collaborative value networks to cooperate more closely and enhance their interorganizational information flow to improve support services, business analysis, and decision making. In particular in product lifecycle management [3], the vast amount and detail of business information already available in manufacturing and supply chain networks is increasing with the adoption of new information technologies. Important new information sources include *Radio-Frequency Identification* (RFID) and the EPCglobal Network [63], the "Internet of Things" including sensor networks in manufacturing, operations and logistics, but also new social technologies leveraging information from customers and employees, such as blogs, online fora and social networks.

A fundamental challenge for traditional corporate business intelligence in collaborating enterprises is the collection, refinement, and comprehensive presentation of the huge amount of product information available from legacy business databases and those emerging technologies in order to improve business analytics, decision support, and processes, in particular if relevant information is distributed across multiple organizations [3]. Information federations, based on service-orientation and semantic web technologies [9], are a promising solution to this challenge [8]. Service orientation is an important paradigm in technology and management for achieving increased flexibility and business integration [19,68]. The Aletheia project [55], a collaboration of major European organizations in research and industry, targets the development of a reference semantic information federation [40,79], based on a *Service-Oriented Architecture* (SOA).

The aim of Aletheia is the creation, federation, analysis, and presentation of relevant information across all phases of the product lifecycle for decision support (Fig. 1). With the help of semantic models in the form of ontologies, information is extracted from heterogeneous sources into semantic repositories. Information sources include corporate databases, office and email documents, but also RFID and sensor networks, public Web sites, and blogs. In order to exchange information between partners in a service-oriented federation based on semantic technologies, specific services are needed that can be accessed by corporate applications, business partners, and end users of a product searching for information. We define a *Semantic*

☆ This research was funded by the German Federal Ministry of Education and Research under grant number 01IA08001E as part of the Aletheia project. The responsibility for this publication lies with the authors.

* Corresponding author.
*E-mail addresses:* bfabian@wiwi.hu-berlin.de (B. Fabian), steffen.kunz@wiwi.hu-berlin.de (S. Kunz), sebastian.mueller@fu-berlin.de (S. Müller), oliver.guenther@uni-potsdam.de (O. Günther).

0167-9236/$ – see front matter © 2012 Elsevier B.V. All rights reserved.
doi:10.1016/j.dss.2012.05.049

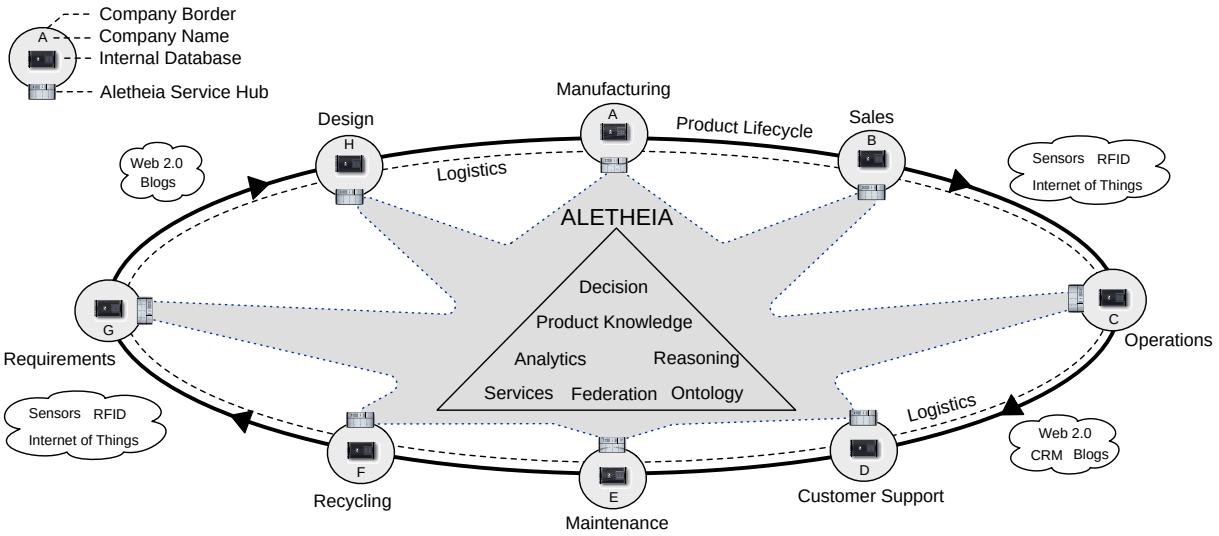



**Fig. 1.** Aletheia—business decision support for the product lifecycle.

*Information Service* as a web service that provides the capability to query semantic repositories and returns information in a semantic data format.[1]

Though on the one hand the business value of increased information exchange is expected to be high (Section 2.1), on the other hand protecting confidentiality and integrity of information as an economic asset becomes crucial and a potential barrier to participation in an information federation. Examples for confidential information include personal information of employees and customers, financial and pricing data for products, as well as research and design information in product and service development. The need for information security is supported by investigations on the business impact of confidentiality and integrity loss of corporate information. For example, [15] assessed an average loss of 2.1% of a company's market value within two days of a confidentiality breach announcement, resulting in an average loss in market capitalization of 1.65 billion dollars per breach. Goel and Shawky [26] found a statistically significant average loss of 1% of the firm market value around the event date of a security breach. Acquisti et al. [2] focused on privacy breaches, i.e., the exposure of personal information due to lack or failures of security mechanisms, indicating a cumulated average loss of 0.6% of the market value.

Already in 2006, Gartner [25] expected cyber-threats for SOA and information mashup to become more acute. Today, the maturity of SOA and semantic technologies as enabling technologies has considerably increased, and so have associated threats. Accordingly, for information federation and SOA architectures to thrive, participating companies must be able to control cross-company information flow and tailor it precisely to the amount that partners need to know for optimizing common business goals [22]. Strong emphasis must be placed on the protection of information against unauthorized access, a necessary precondition to fulfill confidentiality and integrity requirements of the various business stakeholders involved. As our literature review revealed (Section 2), there are no security architectures available that provide a comprehensive security solution for semantic, SOA-based information federations.

In order to close this research gap, our article presents a holistic approach for organizational and technical measures in order to provide a secure federation of semantic information services between organizations. Our research follows the design science paradigm [33], especially the consolidated research process activities established by [69] (Problem Identification and Motivation, Objectives of a Solution, Design and Development, Demonstration, Evaluation, and Communication by means of this article). This is also reflected in the structure of our paper and the following outline of our contributions.

First, information federation benefits and security challenges are discussed in Section 2, corresponding objectives of our work are established, and related work on security for information sharing is presented. In Section 3, we present our main contributions. First we describe a general process of designing a security architecture for federating semantic information services, and describe protection measures to meet the objectives. In particular, as the second major contribution, we present an access-control system for protecting semantic information services, capable of enforcing security policies in information federations based on semantic technologies. As our third contribution, we investigate organizational aspects of security design, and present a role-engineering process for interorganizational access-control policies. Furthermore, an implementation of our infrastructure, a demonstration in the context of an application scenario (Section 4), and an evaluation section, including several technical experiments, demonstrate the practical feasibility and, with respect to a low latency overhead, usability of our architecture (Section 5). A discussion, proposals for future work, and practical and managerial implications are presented in Section 6, before we conclude in Section 7.

## 2. Security challenges and objectives

### 2.1. Benefits of information sharing

Before turning the focus to security challenges, we first emphasize the benefits of information sharing in corporate environments by a short literature survey. In supply chain management and operations research literature, information sharing between cooperating business partners in supply networks has been shown to be beneficial for all participating companies in order to reduce demand risks [13,16,27], in particular to counter the so-called *bullwhip effect* [43]. However, many structural, organizational, and information quality factors must be considered [73]. An empirical investigation of factors positively influencing information sharing in supply chains is given in [44], including trust and a shared vision. A framework for the investigation of trust and conflict in information sharing is presented by [67], where also trust-generating strategies are analyzed. Results from [46] indicate that various information-sharing schemes improve supply chain performance, and extensive sharing of multiple types

[1] This term is different from *semantic web service*, which denotes a class of services that are described by a mark-up using formal semantics, e.g., in order to facilitate its discovery and composition [52].

of information in supply chains has advantages in volatile market conditions. In addition to demand information, the sharing of further information categories such as product histories and customer feedback on product quality can prove beneficial in product-lifecycle management in complex value chains, for example to foster efficient and advanced business analytics [78].

Earlier research on technology for information federations in product lifecycle management [3] is intrinsically related to enterprise information integration [28]. In order to exchange manufacturing data, some rudimentary standards such as STEP have been developed [30]. However, with the advent of Semantic Web technologies [9], several researchers proposed their use for intra- and inter-corporate information integration [5,36,48,83]. An overview on leveraging SOA for supply chains and manufacturing is presented in [14].

### 2.2. Security challenges for information sharing

When companies are gathering to form an information federation, several security challenges arise, as was motivated in the introduction. An established categorization of security requirements is the classical "CIA triad" (confidentiality, integrity, and availability) [57]. Since the availability problem affects every service on the Internet and can be considered as part of general research toward increasing service reliability and robustness, we focus on confidentiality and integrity in our current article. Integrity involves safeguarding the accuracy and completeness of information assets [57]. This involves the protection of information against unauthorized modification, especially during transfer. Confidentiality is the requirement that information is not made available or disclosed to unauthorized individuals, entities, or processes [57]. It refers to the protection of information against unauthorized access.

With respect to confidentiality, several questions should be answered by the management of each participating company. First, which internal information should be shared with specific partners in order to realize the anticipated benefits of the federation? Second, what information is considered confidential and must be kept internal? And not the least, can the decision to let someone access internal information of company *A* be transferred to another company *B* (implying strong trust), or should every access limitation be enforced locally? In addition, confidentiality of information must be protected during transfer and storage on local and remote servers. These general security requirements have also been empirically confirmed as relevant with the help of industry partners in a previous case study on information federation in the industrial service sector [41].

### 2.3. Security objectives

In the context of information federations, these general security requirements can be refined by establishing four major layers of security objectives and corresponding technical and organizational measures (Fig. 2, left). Intuitively, protecting confidentiality involves a decision who should have access to what information. This implies that the system has to distinguish between its users. The first objective of *federated authentication* (the lowest layer in Fig. 2) concerns the fundamental problem of providing and verifying digital identities for entities in information federations.

| | |
|---|---|
| (Semantic-Aware) Access Control | XACML, SemForce (newly developed) |
| SOA Security | WS-Security Standards, XML Encryption, ... |
| Network Security | SSL/TLS, VPN, Firewalls ... |
| Federated Authentication | Federated Identity Management, Single Sign-On Systems |

**Fig. 2.** Security objectives and measures.

Identity providers are services responsible for storing, authenticating and tracking user identities during their operations and interactions with corporate information systems. But in information federations, users usually belong to different security domains, e.g., to different companies. For a federation, one could implement a new, single, and central identity provider, thus creating a new security domain dedicated exclusively to the information federation. Though straightforward and technically viable, this solution does not follow a decentralized federation paradigm and could involve a high operational load and complexity. Furthermore, with respect to usability, this approach results in users having an extra set of credentials that are used exclusively for accessing the federation services. Finally, all participating companies must trust in the correct and secure operation and management of this central identity provider. As an alternative, one could prefer the flexibility of using individually administrated and often historically grown local identity providers, and create a federation of these existing identity stores. A federated identity system enables the portability of the existing identities between several security domains, thus reducing the number of identities a user has to keep track of.

The second security objective in Fig. 2 aggregates classical protection measures from *network security* for protecting information flow from eavesdropping or modification by an external adversary. Extending classical network security, the third objective of *SOA security* is to provide capabilities for protecting a service-oriented architecture. Here, web services messages should be protected from end to end, in particular if they are forwarded by multiple services and cross multiple (application layer) hops in complex SOA infrastructures [35].

Turning to the fourth objective in Fig. 2, an important strategy for satisfying confidentiality and integrity requirements in information storage systems is to provide mechanisms for *access control*: detailed policies should determine who has access to certain information, possibly taking also different contexts into account. The currently established de facto standard for access control is based on roles: *Role-Based Access Control* (RBAC). Users are mapped onto roles, which in turn are mapped onto permission sets [23,65,74], for a comprehensive introduction see [24]. RBAC enables the creation of role hierarchies and permission inheritance, and facilitates policy-change management in typical business environments where user population changes frequently.

Moreover, access control for semantic information services must be *semantic-aware*. When information is processed and exchanged by semantic information services, it is transformed into specific machine-understandable data formats and annotations, which describe the semantics and also allow performing automatic reasoning in context of a formal ontology. This ontology defines objects, their properties, and relations, and thus provides the vocabulary and corresponding semantics describing the domain, such as a product lifecycle. The semantic repositories of the stakeholders will contain statements about entities of this domain. Confidentiality and integrity requirements on the information stored in semantic repositories demand security mechanisms that are semantic-aware, i.e., they are not only able to protect SOA communication, but are also capable of analyzing and protecting semantic technologies involved with semantic information services, such as the *Resource Description Framework* (RDF) data format [71] and the corresponding query language *SPARQL Protocol And RDF Query Language* (SPARQL) [62].

As will be discussed in Section 3, the design process for a comprehensive security architecture involves at every objective layer decisions on adopting centralized or decentralized security measures. This decision involves a careful consideration of inter-organizational trust, power structure, and control—similar to the decision of actually participating in the federation at all.

### 2.4. Related work on security for information sharing

In the following, we discuss further related work on providing security for corporate information sharing. We organize the presentation

according to the layered security objectives identified in Fig. 2, where example technologies are given on the right-hand side. For an overview, [75] addresses confidentiality and privacy concerns and presents a general research framework for secure and useful data sharing. A general overview on technical standards for data sharing and their security is presented by [31].

As a basis for *authentication*, identity management provides digital identities and corresponding roles as part of the so-called user lifecycle [80]. Those entities encompass human users, services, and sometimes even network nodes. Usually, corporations manage identities by deploying an identity provider. Classical examples are identity stores based on *Lightweight Directory Access Protocol* (LDAP) or Microsoft Active Directory. Federation of identities is supported by the OASIS standard *Security Assertion Markup Language* (SAML) [59]. Federated identity management and single sign-on systems are solutions that are already used in practice, but are also still the subject of further research, in particular with respect to privacy [29].

For *network security*, common measures include *Secure Sockets Layer* (SSL) or its newer variant *Transport Layer Security* (TLS) [21], which are capable of protecting application data that is transmitted using the *Transmission Control Protocol* (TCP). An example of such a protocol is the *Hypertext Transfer Protocol* (HTTP), which is commonly used for web services. Similarly, Virtual Private Networks (VPN) protect information flows from external adversaries [17]. It is important to note that SSL, TLS, and VPN are only protecting a direct communication process between authenticated communication partners. Once one partner sends the received information to a third partner, a second secure connection must be established independently of the first.

Concerning *SOA security*, [38] strongly recommends integrating security into SOA from the start. A comprehensive guide to many practical aspects of SOA security is [35]. For protecting message confidentiality and integrity, a suite of protocols called *Web Services Security* (WS-Security or WSS) has been developed by OASIS for ensuring end-to-end security in complex web-service architectures based on the *Simple Object Access Protocol* (SOAP). In addition to other associated mechanisms, it protects SOAP messages by protocols such as XML Signature and XML Encryption [60] for the fundamental standard *Extensible Markup Language* (XML). However, in less complex infrastructures, classical mechanisms such as TLS could be sufficient and are in general less complex to administer.

*Access control* and disclosure control in general have a long history in research, which we cannot extensively discuss in this paper. An influential standard for defining access-control policies and roles has been developed by the OASIS consortium, the *Extensible Access Control Markup Language* (XACML) [54]. Security policies in distributed environments using XACML are discussed by [47] and [51]. [50] describes policy integration algorithms that could become relevant if a central policy administration point for the whole information federation is acceptable for all partners. In contrast, in this paper we focus on decentralized control and policy administration.

Only a few research articles address the specific issues of providing access control for semantic repositories. An introduction and high-level survey on many aspects of semantic web security is presented in [77]. Furthermore, some theoretical designs for this specific challenge have been proposed [1,34], some also aim to support access-control decision by semantic technologies [37]. Though [70] provides an interesting design for policy enforcement and management on semantic data, its implementation status and further availability are unclear. An approach for access control of distributed workflows using XACML and RDF can be found in [20], but RDF is only used as a format for communicating context during access-control decisions. [18] describes an XACML modification supporting Semantic Web extensions, but the focus is on authentication and not authorization, lacking the implementation of policy enforcement and decision points.

At present, we are not aware of any other available security architecture that provides an appropriate semantic-aware access control. In contrast, our semantic-aware access control mechanism *SemForce*, described in Section 3.6, uses a fine-grained level of policy formulation for access on semantic resources, and is based on the established standard XACML. Furthermore, SemForce is not restricted to RDF and SPARQL, but is designed to be easily extendible by providing additional query-processing engines.

## 3. Design and implementation — security architecture for information federations

### 3.1. Design process for a security architecture

In the following, we present the step-by-step design process of a security architecture for federations of semantic information services, in order to achieve high generalizability and to illustrate the process

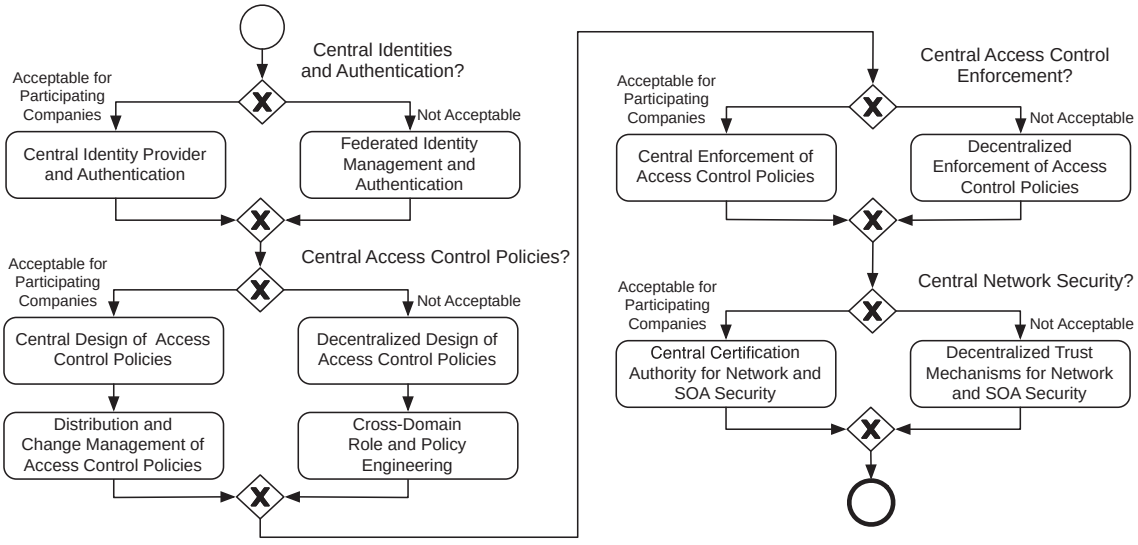



**Fig. 3.** Decision process for designing the security architecture.

of integrating security components. This process involves several decisions concerning the centralization of trust and scalability (Fig. 3).

First, concerning the provisioning of identities and authentication procedures, it must be decided if a central identity provider is acceptable for all participating companies, and if the central management overhead of this approach is scalable for the planned federation. Similar decisions have to be made with respect to access control, as well as for policy design, administration, and enforcement. In the case of a decentralized policy design, a common business process for cross-domain role and policy engineering must be applied, such as the one we propose in Section 3.7. Furthermore, the question of centrality resurfaces in the domain of network and SOA security where truly decentralized trust mechanisms are subject to active research, but have not been widely applied and tested in practice, yet. In the following, we describe the architecture-design process using Aletheia as a reference federation.

### 3.2. Aletheia information federation

Aletheia [55] is an information federation for product-lifecycle management (Fig. 1). Relevant lifecycle phases [3] include product-requirement analysis, product design and development, manufacturing, sales, operation, customer management, maintenance, recycling, and interconnecting logistics processes.

Collaborating business partners will form a small-scale and closed membership peer-to-peer network based on web services. However, even though in theory all peers should be equivalent, certain organizations may in practice still adopt specific functions in the federation, such as membership control or updating of software components. Interorganizational communication in Aletheia, such as semantic queries for product information, is conducted via special proxy nodes deployed in each company, the *Aletheia Service Hubs* (ASH) [79]. These service-oriented federation hubs are responsible for information requests between companies, distributed indexing of services and information, and also act as gateways for internal information requests. They store the global and company-specific parts of ontologies, and conduct the extraction, semantic analysis, and categorization of the information available in the domain of the local company.

In the basic federation, every participant *A* would access all data directly without any security measures (Fig. 4a). This scenario serves as the initial state of a federation of semantic information services and as a baseline in later performance experiments that investigate the possible overhead of security mechanisms. Company *A* accesses its own store, and the stores of companies *B* and *C* remotely via its own and the respective remote federation hubs as service gateways. Access is symmetrical for each company (so also *B* accesses *A* and *C* remotely, and similar for *C*).

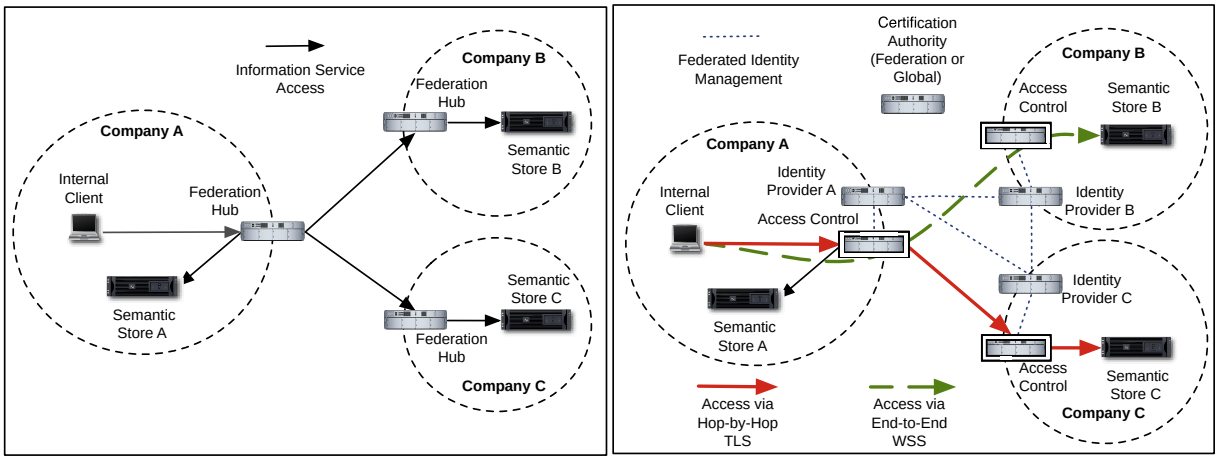
### 3.3. Overview of the Aletheia security architecture

Since this basic federation architecture is insecure, it must be extended by security mechanisms where possible to satisfy the security objectives (see Section 2). Primarily, there is a need for access control to information and services, involving authentication and authorization. This corresponds to both the lowest (3) and highest (8) objective layers of Fig. 2. The extended architecture depicted in Fig. 4b provides a solution for meeting these objectives. First, every company leverages an identity provider for internal user and service identities as well as for corresponding roles. These identity providers should be federated in order to enable a user at company *A* to authenticate using his or her local credentials at the federation hub of company *A*, and use this authentication also for the information access at companies *B* and *C* if those companies authorize the authenticated identity for access.

Every access attempt to internal information services is controlled locally by an access-control component of each company involved. For our architecture, we developed SemForce, a semantic-aware access-control decision and enforcement point, which we discuss in Section 3.6. This component can be deployed as an independent set of security services, or can be directly integrated into the federation hub.

### 3.4. Aletheia-SSO—federated authentication

In order to protect access to resources in Aletheia, individuals and roles have to be identified. In our project, no central identity provider for the whole federation should be installed since participating companies are very independent of each other and leverage existing identity providers. Instead, local control over identities and user credentials should be maintained. The authentication of entities, as well as their assignment to roles is the task of an authentication component (the lowest layer of Fig. 2). Within each company, individual identity providers that also include authentication components are already established (e.g., Active Directory, LDAP) and should be reused for Aletheia. Each company registers new users, clients, or services to these local identity providers, which need to be federated in order to provide a capability for shared identity and role information and for single-sign on throughout Aletheia.




(a) Basic Federated Information Access (b) Protected Federated Information Access

**Fig. 4.** Addition of security components in order to protect information access.

For Aletheia, we systematically evaluated existing identity management frameworks in terms of suitability for our security architecture. Selection criteria included the exact set of functionality provided, the existence of an active user community, and the frequency of releases and updates. As basis for our federated authentication system *Aletheia Single-Sign On* (Aletheia-SSO), we selected and adopted components from the open-source project OpenSSO [61]. With OpenSSO, it is possible to use existing or define new subjects (i.e., users and roles). The OpenSSO implementation can be integrated into any web application or web service. Every access to this application and service will then first be intercepted and handled by OpenSSO. Once an entity, such as a user, has been authenticated, OpenSSO assigns her a token. The user can then confirm her identity and role to any web application and service by sending this token.

### 3.5. SOA and network security

In general, all communications between services and clients within the information federation need to be protected. This involves the middle layers of Fig. 2. As discussed in Section 2, the choice of protection measures depends on the complexity of the SOA message routing and the web service protocols involved, and cannot be made final at the abstract level of discussion in this section. The WSS protocol suite should be employed if messages are often routed across several service hops in multiple domains, for example if an information request involves multi-hop information transfer from an information source in company *B* to a client in company *A*, as shown in Fig. 4b.

Moreover, one has to consider if authenticity of the information should also later be provable for non-repudiation. In this case, storage of the original message—in addition to the semantic data extracted from the message—including its original XML signature issued by company *B* could be useful. In simpler scenarios, or if the web service protocols are not based on SOAP, the classical SSL or TLS protocols should be preferred. If the membership set of the federation is not large and very stable, a VPN solution and the formation of a protected Extranet ([17], p. 247), tunneling all federation communication over secure virtual connections, could be an alternatively viable approach.

All SOA and network security measures must provide means to identify and authenticate services and the nodes that are communicating. An important scalable measure in network security engineering is to use public-key cryptography and certificates, which are binding identities to public keys. In order to issue and manage certificates, a *Certification Authority* (CA) is used. In larger infrastructures, hierarchies of CAs would form a *Public-Key Infrastructure* (PKI). Classical CAs and PKI involve a central entity as the root of trust, which may be considered as a violation of the paradigm of forming a federation of equal peers in some application scenarios. In research, several approaches for distributing trust have been proposed, see [82] for distributed CAs, and [45] for a general overview on trust in distributed systems. Adoption of such decentralized trust systems is still a challenge for information systems research and practice. Because some Aletheia components are using *Representational State Transfer* (REST), we decided to use TLS as the main communication-security mechanism for our implementation, which can be used to secure both approaches. In order to avoid a single, central CA for the federation, we recommend using official TLS certificates signed by well-known, Internet-wide CAs.

### 3.6. SemForce—authorization and semantic-aware access control

The actual authorization and enforcement of access control are conducted by an interplay of several security components of our architecture. Since we use XACML as language to define and transmit RBAC policies (Section 2), we also adopted the XACML security-architecture model. This includes a *Policy-Administration Point* (PAP) for administering policies, a *Policy-Decision Point* (PDP) for matching incoming requests and their roles to the policy base and issuing permissions or denials of access. Furthermore, we provide a *Policy-Enforcement Point* (PEP) that is responsible for enforcing the decision in the particular service environment. Since no PEP or PDP for semantic repositories existed, we designed and developed a new solution for enforcing access-control policies for semantic information services, the *SemForce PEP* and *SemForce PDP*. These SemForce components can be all deployed on the same server, or as distributed services over the network.

The default query language used within SemForce is SPARQL [62], but also further query languages and protocols can be supported by a plug-in mechanism that loads corresponding query-processing engines during runtime. The store protected by SemForce may be any semantic data store that is compatible with RDF [71], for example the open source Jena framework [58] that we used for our own implementation and experiments.

SemForce follows a query-rewrite approach: The access-control policy is enforced by rewriting the SPARQL query and extending it by filtering statements in order to prevent access to non-permitted resources. This includes an analysis of resources explicitly mentioned in the query, but also of implicit resources that would be returned by inference of the reasoning engine within the repository. An alternative approach for semantic-aware access control could be the creation of views in the semantic repository in order to enforce the policies. The view generation, however, is resource-intensive and would have to be repeated for each change in the access-control policies.

A further advantage of the rewriting approach is that performance optimization and materialization strategies can be considered separately. Moreover, SemForce works with different repository implementations and serves as an independent service for audit and control. In order to support the service-oriented architecture of the Aletheia federation, we implemented SemForce as a web service based on SOAP.

### 3.7. Decentralized access-control management

Accompanying the development of technical components for a security architecture, an important organizational challenge is to develop and manage access-control policies for information federations. This is due to the full organizational complexity of the federated information landscape. Concerning role modeling within a single organization, in [32], a goal-driven role engineering process is introduced, which does not reflect business scenarios. Organizational structures are used by [49] in order to model access-control policies by applying the SI* modeling notation, which was originally developed for socio-technical systems. [7] developed their own modeling notation SecureUML, though policy development for roles and access-control was not a focus of their work. A mapping between XACML and the *Business Process Modeling Notation* (BPMN) is developed and applied to a banking industry scenario by [81].

One of the most prominent approaches was presented in [65,76]. These articles describe scenario-driven role engineering, a methodology for reducing the complexity of developing an organization-specific RBAC model within one single domain. Fig. 5 depicts the control flow of the scenario-driven role-engineering process for a single company *A* in the upper part of the diagram. First, (1) the underlying scenario has to be determined and modeled. This can be achieved through structured text or any suitable diagrammatic model. Scenarios consist of several steps and each of them can occur in multiple scenarios. A step activates an action, which is defined by an operation and a target object. Based on the defined scenarios, (2) permissions are derived, which are in turn used for defining tasks. In addition, (3) constraints are defined. After all these steps have been completed, (4) the model has to be checked and refined. If the model does not represent the scenario correctly, it has to be improved. If the model is complete, in the next step (5) work profiles have to be derived from the defined tasks. These work profiles (6) function as provisional roles, which are used for (7) defining the RBAC model. A final check

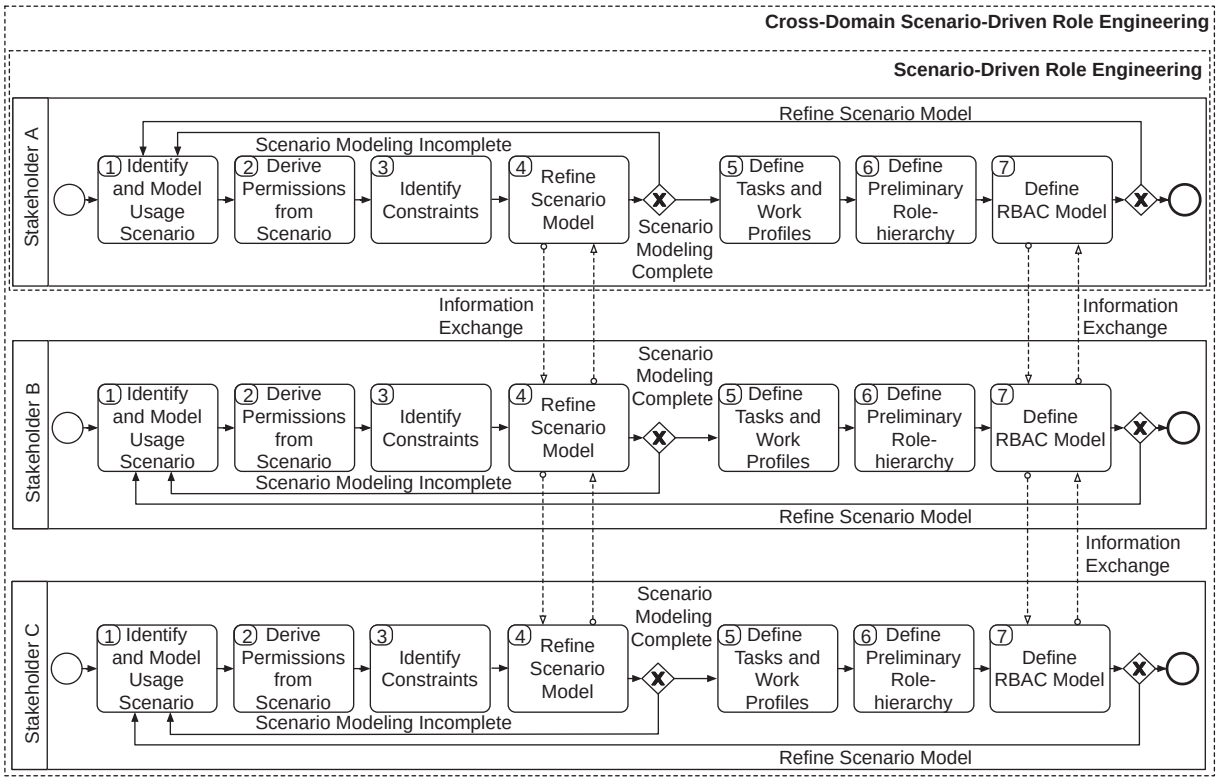



**Fig. 5.** Cross-domain scenario-driven role and policy engineering.

determines if the RBAC model is complete or if the model has to be refined; for details see [65,76].

All of the presented approaches have one shortcoming in common: they do not address the interorganizational aspect, which is key for our research on federations of semantic information services. This issue of cross-company information exchange is currently also intensively discussed in the business-process management community with respect to process choreographies. Considering that some approaches for extracting RBAC models from process models have already been developed [53], an extension of the scenario-driven role-engineering process through the coordination of inter-organizational choreography of stakeholders by using a common business process modeling language seemed promising and was developed for Aletheia.

The box entitled *Cross-Domain Scenario-Driven Role Engineering* in Fig. 5 displays an extension of the scenario-driven role engineering approach of [65,76], including cross-company information exchange at crucial points of the engineering process: in step (4), the refining of the scenario model, and step (7), the definition of the RBAC model. Based on this theoretical extension and the presented research gap identified, we developed the BPAX converter for modeling access-control policies in an interactive, systematic and collaborative way by using process-choreography models, for details see [42].

However, design of access-control policies in practice is a task that involves a deep background knowledge on business processes, and must be strongly supported by management. It must be handled with due care and sufficient time since it cannot be fully automatized. In information federations, it also involves dedicated coordination activities between process experts and policy administrators from cooperating companies.

### 3.8. Security and power structures

We conclude the description of the security architecture design process with a general remark on centrality and power structure. In an application scenario where the main motivation for a federated design is the avoidance of central points of failure or control, also all security measures must be investigated in order to prevent the reintroduction of centralized elements into the architecture. In other application scenarios, hybrid designs involving a mixture of centralized or decentralized security components may be less critical, except for possible performance or management bottlenecks. In any case, power structures in security should be analyzed carefully by experts and management during security design.

## 4. Demonstration—secure information federation in the industrial service sector

For illustration of the practical usefulness of our security architecture, we conceptually applied it to a real-world application scenario from the domain of an industry partner in the Aletheia project. The scenario is displayed in Fig. 6, and was originally discussed in our previous research [40,41]. In this application scenario from the maintenance phase of the product lifecycle, we focus on the industrial service sector, in particular on a service technician conducting a service job. The three companies involved are depicted in dashed circles: the *Industrial Service Provider* (SP), the *Partner Company* (PC), and the *Logistics Provider* (LP).

The main business processes in this application scenario are as follows. The partner technician belongs to the Partner Company, which was subcontracted by the Industrial Service Provider to conduct a service job, i.e., to repair a defective part at a customer site—here a power plant of an electric power provider. On the one hand, the technician has to access company internal resources for the preparation of the service job such as information provided by the internal enterprise resource planning system of the Partner Company. On the other hand, she needs access to the federated information from the distributed semantic repositories in order to conduct the service job, including historic information of previous service jobs and documentation from

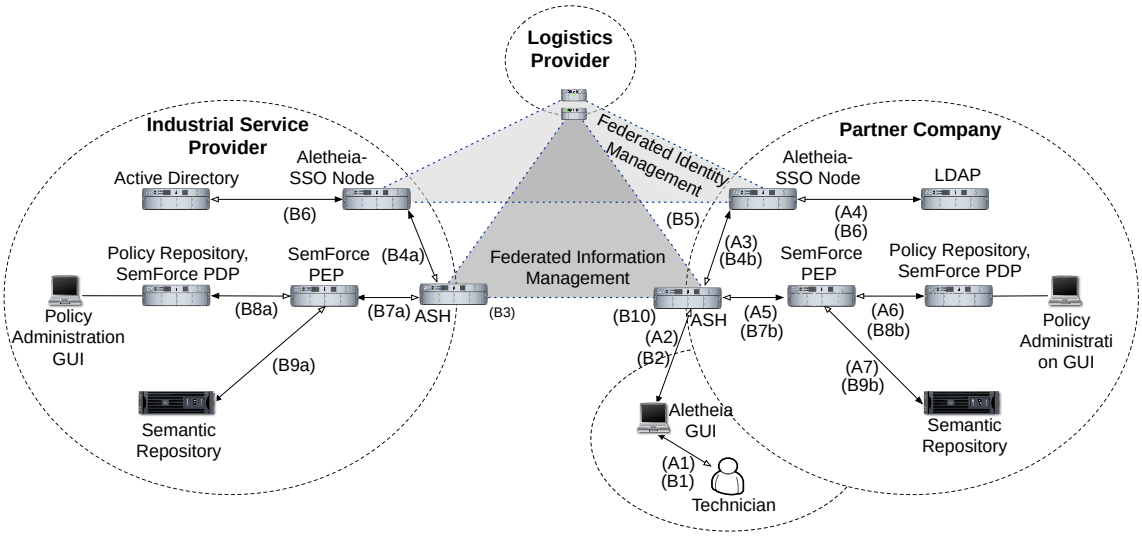



**Fig. 6.** Security architecture in an application context.

both the Industrial Service Provider and the Partner Company. The information exchange takes place via the Aletheia Service Hubs (ASH).

We will now introduce and explain the interplay of the security components by two sample information requests from the application scenario in Fig. 6. Both queries are executed by the Partner Company technician. Whereas query *A* targets the internal semantic repository of the Partner Company only, query *B* needs to contact the semantic repositories in both domains of the Industrial Service Provider and the Partner Company, thus requiring a federated management of information and identity. Access-control policies have been engineered earlier at the Policy Administration Points.

Authentication and authorization for request *A* are conducted as follows: In step (A1) in Fig. 6, the partner technician requests information from the local semantic repository via the Aletheia GUI. This request is directed to the ASH (A2), which demands authentication from the user and forwards the request to the local Aletheia-SSO node (A3). The latter authenticates the technician with the help of the identity provider of the Partner Company, here an LDAP Server. This information is transferred back to the ASH, which then sends the query to the SemForce PEP (A5). Before the query can be executed on the semantic resources, the PDP has to check the roles and associated access-control policies assigned to the partner technician, and informs the SemForce PEP which resources the user is allowed to access (A6). If necessary, the SemForce PEP rewrites the query according to the information provided by the PDP and enforces the access-control policies on the semantic repository (A7). The results are then displayed to the technician via the Aletheia GUI.

Request *B* is an information query that requires federation, e.g., a request for all information about a defective device from the domain of the Partner Company as well as the Industrial Service Provider. Again, the partner technician sends her request via the Aletheia GUI (B1) to her ASH (B2). The ASH of the Partner Company routes the request to the ASH of the Industrial Service Provider (B3). Both ASHs request their local Aletheia-SSO nodes for authentication (B4a/b). Since the Active Directory of the Industrial Service Provider is not able to authenticate the service technician, the authentication request is directed to the Aletheia-SSO Node of the Partner Company (B5). There, the service technician is authenticated by the Partner Company's LDAP Server (B6).

The corresponding authentication token is transferred to the ASH in the respective domains and is used for identifying the user and her role during the information requests to both SemForce PEPs (B7a/b). The PEPs contact the PDPs for an access decision (B8a/b). If necessary, the PEPs then rewrite the queries according to the information provided by the PDP, and enforce the access-control policies on the semantic repositories (B9a/b). The query results of the Industrial Service Provider are then transferred from its ASH to the ASH of the Partner Company (B10), where they are merged and presented to the technician via the Aletheia GUI.

## 5. Evaluation

### 5.1. Reflection of the design based on the objectives

In Section 2.3, we refined the stakeholder requirements of confidentiality and integrity in the context of information federations into four layers of security objectives (Fig. 2): Federated authentication, network and SOA security, semantic-aware access control. These objectives are fulfilled by our architecture: Authentication is provided by Aletheia-SSO that is federating existing identity providers in each company for cross-company single sign-on. In our design, network security (confidentiality, integrity) is generally provided by TLS, whereas for securing multi-hop SOAP message exchange, WSS would be adopted. Our architecture provides access control in the form of RBAC. Corresponding XACML policies are tailored by an interorganizational process and enforced by several SemForce instances, each under local control of the respective company.

A complete and holistic view on all facets of practical security for critical deployments is out of the scope of our research and this article. For example, potential security flaws in TLS implementations would also affect our infrastructure. But since we chose to adopt established security protocols where possible, such flaws would be quickly fixed by a larger community. For mission-critical applications, a verified secure implementation, installation, and operation are major practical challenges. Furthermore, the factor of human error during policy design, handling of cryptographic keys, or during parameter choice for cryptographic algorithms in TLS or WSS, needs to be reduced by policies and processes in practice.

### 5.2. Network performance as indicator for usability

Even though a security mechanism works correctly, overall security could be impeded by a lack of usability. If users are not able or not willing to use the mechanisms, they may find ways to circumvent it. Many usability criteria are not directly relevant to the design of service architectures as software artifacts without explicitly designed user interfaces. However, for architecture design in information retrieval,

*latency* is a major usability issue [64,72], and therefore also represents a challenge for security and privacy measures. In [12], we discovered that users are willing to accept a higher latency in order to gain more privacy (in this case via an anonymization network). But, this result is only valid for a personal context and small increases in latency. We hypothesize that our proposed security features will gain larger user acceptance if their latency is comparable to the latency without using them. In order to investigate this problem, we provide experimental performance evaluations in the following.

### 5.3. Performance experiments on secure federation

In order to demonstrate the practical feasibility and usability of our security architecture, performance experiments for federating information from a distributed installation of several semantic repositories were conducted. Each repository was situated in a different domain and protected by a local SemForce instance. The context of the experiments is the application scenario from Section 3.4 displayed in Fig. 6. The gathering and federation of information is the task of one special component of the ASH: the *union service*. This service is responsible for contacting remote ASHs in order to query for information, and for merging the returned answers to a federated information set. For avoiding performance influences, the union service, which works on behalf of the user as a first point of contact for a query, was deployed separately from the SemForce instances on different machines.

Pre-test measurements established that authentication latency for using Aletheia-SSO is fairly constant in each environment except for random network conditions; furthermore, since authentication tokens can be reused for multiple queries, this step was excluded from the main experiments, which focus on assessing the performance of federated SemForce and repository services. The distributed setup included the following four nodes: Industrial Service Provider (SP), Logistics Provider (LP), and Partner Company (PC), each represented by an *Amazon Elastic Cloud 2* (EC2) *small instance* server located in Ireland. For Request 1, *extra large* EC2 server instances were used due to high main memory requirements. The union service was deployed on an EC2 small instance server; for Request 1 also the extra large EC2 server instance was used. Technical details are displayed in Table 1.

In order to populate the data stores of the different domains in the experiments with small (1000 instances), medium (10,000 instances) and large data sets (100,000 instances), classes were extracted from a real-world ontology of a project partner, replicated, and slightly modified. This way we assured data consistency for our experiments while still processing real-world and realistic data. The industrial service provider maintains the class *service*, the logistics provider the class *delivery*, and the partner company the class *job*. The class for the industrial service provider is shown in *Turtle* notation for RDF with an exemplary instance in Table 2, while Table 3 displays a fragment of the corresponding policy.

Each single instance is equivalent to eight triples in the store of the industrial service provider, seven triples in the logistics provider store, and six triples in the partner company store. The experiments were conducted for every possible combination of store sizes in the three companies. Furthermore, decentralized access-control policies were used as input for each SemForce PDP. Access-control policies were defined for each company repository individually, restricting read access for non-company employees on confidential information. For example, the property *price* of the class *service* was denied for employees of the logistics provider and the partner company, but read-access to the property *date* was permitted.

In the first series of experiments, it was evaluated how store size, its distribution, and the total return-set size of the stores influence response time. The total return rate is the percentage of data in all queried stores that is returned by a single request. The response time of four different requests with different total return rates was measured: *Request 1* returns every allowed triple for the role of an industrial service provider dispatcher (extremely high total return rate of ≈90%). *Request 2* filters triples via the *date* to a certain year for the role of a logistics provider *call center agent* (medium total return rate ≈0.4%). *Request 3* filters triples by a certain *price* for the role of a partner company *technician* (low total return rate of ≈0.007%). *Request 4* filters triples by partner for the role of an industrial service provider *dispatcher* (no returns). *Request 2* can be regarded as a realistic average filtering request.

In order to reduce random effects, each experiment was conducted five times, and the average of those measurements was

**Table 1**
Configuration of Amazon cloud server instances.

| Parameter | Small instance | Extra large instance |
|---|---|---|
| CPU | 1 EC2 Compute Unit (≈1.0–1.2 GHz 2007 Opteron [56]) | 4 EC2 Compute Units |
| Main memory/hard drive | 1.7 GB/160 GB | 15 GB/1690 GB |
| Platform/operating system | 32 bit/Ubuntu 9.10 Server | 32 bit/Ubuntu 9.10 server |
| Web service container | Axis2 1.4.1 | Axis2 1.4.1 |
| Web application server | Tomcat 6.0.26 | Tomcat 6.0.26 |
| Java runtime environment | 6.20dlj-0ubuntu1.9.10 | 6.20dlj-0ubuntu1.9.10 |

**Table 2**
Ontology fragment (SP).

```
@prefix a: <http://sp/ontology#> . &
a:Service rdf:type rdfs:Class .
a:ServiceNr rdf:type rdf:Property .
a:Date rdf:type rdf:Property .
a:ServiceDescription
    rdf:type rdf:Property .
a:Customer rdf:type rdf:Property .
a:Partner rdf:type rdf:Property .
a:Price rdf:type rdf:Property .
a:ServiceCosts rdf:type rdf:Property .
a:LogisticsCosts rdf:type rdf:Property .
a:Service1 rdf:type a:Service .
a:Service1 a:ServiceNr "1" .
a:Service1 a:Date "20100803" .
a:Service1 a:ServiceDescription
     "Heating Control Element" .
a:Service1 a:Customer "Firm 1" .
a:Service1 a:Partner "Partner 1" .
a:Service1 a:Price "1000 Euro" .
a:Service1 a:ServiceCosts "400 Euro" .
.....
```

**Table 3**
XACML policy fragment (SP).

```
<Rule RuleId="lp:employee:ruleid:1" Effect="Permit">
<Description>
Partner Company Technician
</Description>
<Target>
...
<SubjectMatch MatchId="... string-equal">
<AttributeValue DataType=" ... string">
pctechnician </AttributeValue> ...
</SubjectMatch>
...
<ResourceMatch MatchId="... string-equal">
<AttributeValue DataType="... string">
http://sp/ontology#Date </AttributeValue> ...
</ResourceMatch>
...
<ActionMatch MatchId="... string-equal">
<AttributeValue DataType=" ... string">
write </AttributeValue> ... </ActionMatch>
...
```

calculated. The metrics are: First, the overall *request duration* for the union service, i.e., the time it takes for the union service to send a request to all remote semantic information services protected by SemForce, to receive their answers (in parallel processing), and to merge them. This duration includes delays due to network traffic. The other three metrics measure the reaction time of each of the SemForce-protected information services in the three different domains—Industrial Service Provider, Logistics Provider and Partner Company—while they respond to the union service requests (excluding network traffic).

Request 1 queried for almost all data in the small and medium stores and served mainly as a stress test for the architecture. Each individual information service processed the request on small stores within 1 s. For medium stores, the return time was between 17 s (PC) and 41 s (SP). In addition, it was observed that the union service produced considerable overhead when merging the return sets. The largest total return set for Request 1 consisted of approximately 190,000 triples.

For Requests 2–4, Fig. 7a, b, and c, respectively, shows the average request durations of the union service and SemForce reaction times. Fig. 7a indicates that Request 2 was executed by the information services of small and medium stores within 1 s; on large stores, such as the 100,000 instances store of the Logistics Provider, the individual information-service processing took about 10 s with a federated total return-set size of 9000 triples. Similar observations can be made for Request 3 displayed in Fig. 7b where only the information service of the partner company provides a return set. For the large store (100,000 instances) and total return-set sizes of 210 triples, the performance of the Partner Company information service was around 1 s. Requests to smaller store sizes were processed without notable delay. Fig. 7c describes the outcome for the most restrictive Request 4, where no return set is given back. The individual information services executed the request without notable delay. The federation by the union service caused an delay of about 2 s.

Fig. 8 shows the results of the comparison of a protected store (protected by SemForce) with an unprotected store (without SemForce). In this experiment, additional waiting time caused by SemForce becomes notable. It ranges from 0.06 s, for small stores, to 6.6 s for large stores. While this latency overhead is notable, it appears acceptable for practice, though further user studies will be necessary to confirm this evaluation.

Another experiment investigated the overhead of using TLS to secure all information service connections. The outcome displayed in Fig. 9a and b is based on the same experimental setup with store distributions of SP:1 LP:1 PC:1 for the Requests 1, 2, 3, and 4. The experiments indicate an increase in network latency that varies from 21% to 55%, but can be also regarded as acceptable.

Turning to concurrency, Fig. 9c depicts the outcome of the concurrency tests, using the store distribution SP:1 LP:1 PC:1 and Request 1, while the number of concurrent request is increased from 1 to 20 in steps of 5. The x-axis shows the durations of each step. The query duration is calculated as the difference of the latest stop (response calculated, response merged) and the earliest start (request arrived) of the concurrent requests. Here, an important criterion is overall response time, which is determined by the response time of the union service. The results indicate that the response time is linear (or even below linear) in comparison to the number of requests. Hence, the architecture is able to handle and to some extent even profit from parallel requests.

## 6. Discussion

### 6.1. Contributions

The process of introducing security to a federation of semantic information services involves several design decisions, as was discussed in Section 3 and outlined by Fig. 3. We applied this process to Aletheia, an information federation for product lifecycle management, and implemented necessary components for the application area of

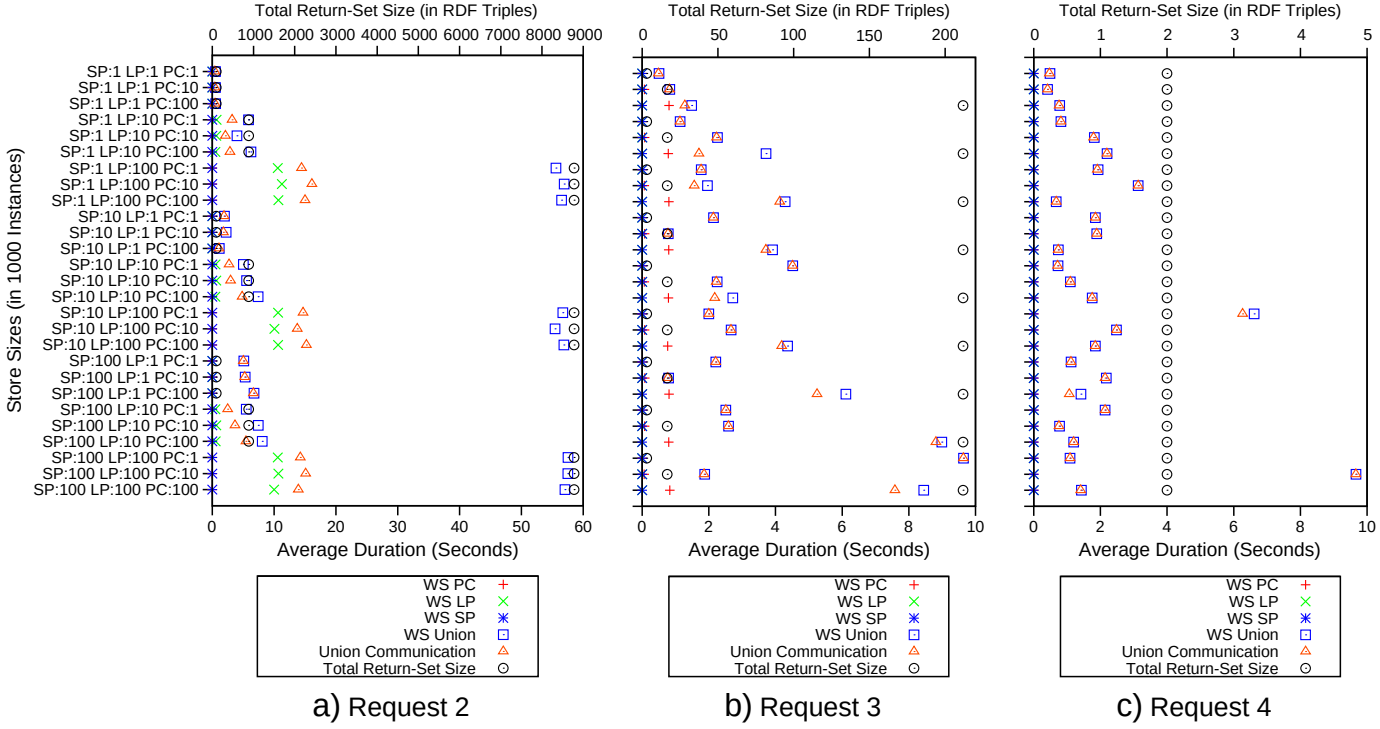



**Fig. 7.** Average duration and size of federated return sets for Requests 2–4.

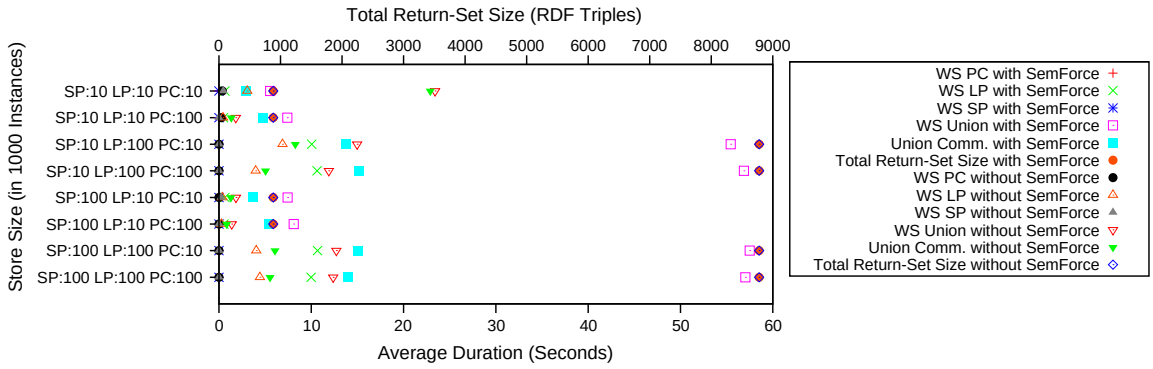



**Fig. 8.** Comparison of access with and without SemForce for Request 2.

industrial service provisioning. These components include SemForce, a semantic-aware decision and enforcement service for access-control policies, and Aletheia-SSO for federated identity management and single sign-on. Distributed experiments on part of our infrastructure indicated the technical viability of our security architecture, though the performance of our system in terms of latency could be still improved. Moreover, we provided a new methodology and organizational process choreography to design roles and policies for access control in inter-organizational information federations.

### 6.2. Limitations and future work

Future research will have to be conducted along both the technical and organizational dimensions. Concerning technical future work on our prototype, we will investigate how our components such as the federating union service and the security components SemForce and Aletheia-SSO can be made more efficient in terms of latency for enhanced usability. In order to improve response time for semantic federations in general, the use of relational databases for the triple store could be promising. Concerning communication security for SOA, the benefits and drawbacks of WSS and TLS in several application scenarios and federation architectures should be systematically investigated. Further research must also be conducted on designing decentralized security services and testing them in real-world applications. Important topics here include decentralized federated identity management and certification authorities. Decentralization will also positively impact availability of the federation and its security components. Further research will also have to be conducted on replication mechanisms for all components involved in order to provide better scalability and availability.

An evolving use case for information federations involves continuous queries, for example in order to support stream processing [11]. Continuous queries are queries over data streams that are executed when new data arrives [39]. Though the presentation of our design focuses on ad hoc queries, there are only minor modifications necessary in order to support processing of continuous queries, namely explicit handling of longer authorization intervals and the extension of SemForce for new SPARQL dialects such as C-SPARQL [6] by providing a new plug-in engine. In C-SPARQL, similar to other approaches [4,10], the requester can include options indicating the time frame when data should be requested. The main change in our design would be to extend the rule interpreter for checking request times against policy lifecycles. Furthermore, there is already an option to include time functions in XACML in order to express lifecycles of policies. Therefore, we expect no major hurdle in order to provide security for continuous queries and stream-processing in our approach.

Concerning organizational aspects, it should be investigated how formal ontologies could be applied to the mapping of business process models to access-control policies. We will continue working on tool support for security processes, and on designing our converter of process models to policies as a web service. Further usability enhancements aim to provide a comprehensive policy-administration interface to the managers of access-control policies, allowing also for supporting the necessary inter-organizational communication. The interorganizational policy-management process should be validated further by field tests and user studies.

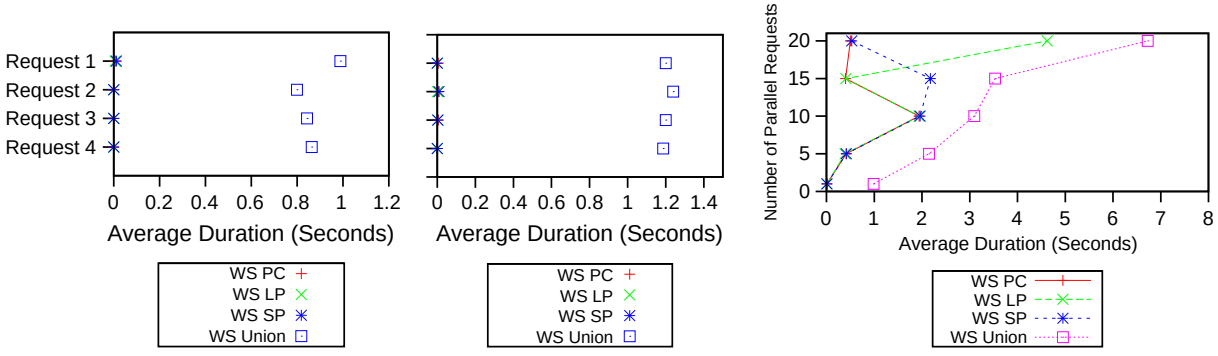



**Fig. 9.** (a/b) Impact of transport layer security; (c) concurrent requests.

Finally, it must be investigated how privacy of information subjects and users of the federation can be protected. Some of our mechanisms for access control will be applicable also in this area, but additional privacy-preserving components should be developed to achieve anonymity and confidentiality of personal information where this is demanded by individuals or public policy.

### 6.3. Practical and managerial implications

Our research has the following main managerial and practical implications:

1. Information sharing based on service orientation leads to potential benefits for companies, but only if security is addressed. Without secure information exchange, companies would refrain from participating in a federation. Security was also emphasized by Demirkan et al. [19] as a major issue for distributed and service-oriented architectures.
2. Security standards and service orientation should be adopted in order to enhance interoperability of security mechanisms. The research presented in this paper closes current research gaps in security for semantic information federations. Nevertheless the newly developed components incorporate or offer interfaces to established security standards. Furthermore, by using service orientation as design principle, our security architecture can be easily integrated into existing SOA architectures, thus fostering reusability and adoption by reducing integration overhead.
3. Design of access control should be conducted as a formal business process, supported by management, and with enough time and budget to guarantee policy quality. In the context of an information-security management process, the following questions must be answered: (1) Which internal information should be shared with specific partners in order to realize the anticipated benefits of the federation? (2) What information is considered confidential and must be kept internal? Access control should be driven by business processes, reflecting roles and a need-to-know basis. Studies such as [66] motivate that RBAC reduces administrative processing time and reduces the frequency and severity of security violations. Benefits are expected to multiply for cross-company access-control engineering. We described an approach that allows engineering RBAC policies in a systematic and collaborative way. However, design of access-control policies, in particular for federations, must be strongly supported by management. It must be handled with care and sufficient time, since it involves coordination between cooperating companies.
4. Security mechanisms could introduce new power structures between companies. Our approach allows to govern the access-control policies and to implement the security components in a centralized as well as decentralized way. Both approaches have their benefits and drawbacks. In a decentralized approach, the individual domains keep their security sovereignty, whereas in a centralized approach security could be offered by a trusted third party security service provider as a service, providing a potential for new business models. As a guideline, the impact of corresponding power structures must be investigated before committing to a final design.

## 7. Conclusion

In this article, we presented a holistic approach for introducing organizational and technical measures to an information federation. In particular, we described the design and implementation of security measures and processes for federated semantic information services. Special emphasis has been placed on reflecting the inter-organizational and decentralized character of a federation, both in the security architecture and corresponding process choreographies for designing access control. An implementation in the context of the Aletheia project and several experiments show the practical viability of our security infrastructure. In future, we will focus on further decentralizing important security services such as certification authorities, and investigate the use of semantic technologies not only as a subject of security but also for supporting security processes. Moreover, we aim to refine the business processes for security, and will extend our approach to the field of privacy for information subjects and users of the federation.

## Appendix A. List of abbreviations

| Abbreviation | Extension |
|---|---|
| Aletheia-SSO | Aletheia Single-Sign On |
| ASH | Aletheia Service Hub |
| BPMN | Business Process Modeling Notation |
| CA | Certification Authority |
| EC2 | Amazon Elastic Cloud 2 |
| HTTP | Hypertext-Transfer Protocol |
| LDAP | Lightweight Directory Access Protocol |
| LP | Logistics Provider |
| PAP | Policy-Administration Point |
| PC | Partner Company |
| PDP | Policy-Decision Point |
| PEP | Policy-Enforcement Point |
| PKI | Public-Key Infrastructure |
| RBAC | Role-Based Access Control |
| RDF | Resource Description Framework |
| REST | Representational State Transfer |
| RFID | Radio-Frequency Identification |
| SOA | Service-Oriented Architecture |
| SOAP | Simple Object Access Protocol |
| SAML | Security Assertion Markup Language |
| SP | Industrial Service Provider |
| SPARQL | SPARQL Protocol And RDF Query Language |
| SSO | Single Sign-On |
| SSL | Secure Sockets Layer |
| TCP | Transmission Control Protocol |
| TLS | Transport Layer Security |
| VPN | Virtual Private Network |
| WSS | Web Services Security |
| XACML | Extensible Access-Control Markup Language |
| XML | Extensible Markup Language |

## References

[1] F. Abel, J. De Coi, N. Henze, A. Koesling, D. Krause, D. Olmedilla, Enabling advanced and context-dependent access control in RDF stores, 6th International/ 2nd Asian Semantic Web Conference, LNCS, vol. 4825, Springer, 2007, pp. 1–14.
[2] A. Acquisti, A. Friedman, R. Telang, Is there a cost to privacy breaches? An event study, 5th Workshop on the Economics of Information Security (WEIS), 2006.
[3] F. Ameri, D. Dutta, Product lifecycle management: closing the knowledge loops, Computer-Aided Design and Applications 2 (5) (2005) 577–590.
[4] D. Anicic, P. Fodor, S. Rudolph, N. Stojanovic, EP-SPARQL — a unified language for event processing and stream reasoning, 20th International Conference on World Wide Web (WWW'11), 2011, pp. 635–644.
[5] J. Arjona, R. Corchuelo, D. Ruiz, M. Toro, From wrapping to knowledge, IEEE Transactions on Knowledge and Data Engineering (2007) 310–323.
[6] D.F. Barbieri, D. Braga, S. Ceri, E. Della Valle, M. Grossniklaus, C-SPARQL: SPARQL for continuous querying, 18th International Conference on World Wide Web (WWW'09), 2009, pp. 1061–1062.
[7] D. Basin, J. Doser, T. Lodderstedt, Model-driven security: from UML models to access control infrastructures, ACM Transaction on Software Engineering and Methodology 15 (2006) 39–91.
[8] V.R. Benjamins, J. Davies, R. Baeza-Yates, P. Mika, H. Zaragoza, M. Greaves, et al., Near-term prospects for semantic technologies, Intelligent Systems 23 (1) (2008) 76–88.
[9] T. Berners-Lee, J. Hendler, O. Lassila, The semantic web, Scientific American 284 (5) (2001) 34–43.
[10] A. Bolles, M. Grawunder, J. Jacobi, Streaming SPARQL — extending SPARQL to process data streams, 5th European Semantic Web Conference (ESWC'08), 2008, pp. 448–462.
[11] I. Botan, Y. Cho, R. Derakhshan, N. Dindar, L. Haas, K. Kim, et al., Federated stream processing support for real-time business intelligence applications, Enabling Real-time Business Intelligence, LNBIP, vol. 41, Springer, 2010.
[12] F. Brecht, B. Fabian, S. Kunz, S. Müller, Are you willing to wait longer for internet privacy? European Conference on Information Systems (ECIS), 2011.

[13] G.P. Cachon, M. Fisher, Supply chain inventory management and the value of shared information, Management Science 46 (8) (2000) 1032–1048.
[14] G. Candido, J. Barata, A.W. Colombo, F. Jammes, SOA in reconfigurable supply chains: a research roadmap, Engineering Applications of Artificial Intelligence 22 (6) (2009) 939–949.
[15] H. Cavusoglu, B. Mishra, S. Raghunathan, The effect of internet security breach announcements on market value: capital market reactions for breached firms and internet security developers, International Journal of Electronic Commerce 9 (1) (2004) 70–104.
[16] F. Chen, Information sharing and supply chain coordination, in: S. Graves, A. de Kok (Eds.), Supply Chain Management: Design, Coordination and Operation, Handbooks in Operations Research and Management Science, vol. 11, Elsevier, 2003, pp. 341–421.
[17] W.R. Cheswick, S.M. Bellovin, A.D. Rubin, Firewalls and Internet Security, 2nd ed. Addison-Wesley, 2003.
[18] E. Damiani, S.D.C. di Vimercati, C. Fugazza, P. Samarati, Extending policy languages to the semantic web, 4th International Conference on Web Engineering (ICWE'04), 2004, pp. 330–343.
[19] H. Demirkan, R.J. Kauffman, J.A. Vayghan, H.G. Fill, D. Karagiannis, P.P. Maglio, Service-oriented technology and management: perspectives on research and practice for the coming decade, Electronic Commerce Research and Applications 7 (4) (2008) 356–376.
[20] V. Dhankhar, S. Kaushik, D. Wijesekera, Securing workflows with XACML, RDF, and BPEL, Data and Applications Security XXII: 22nd Annual IFIP WG 11.3 Working Conference on Data and Applications Security London, UK, LNCS, vol. 5094, Springer, 2008, pp. 330–345.
[21] T. Dierks, E. Rescorla, RFC 4346: The Transport Layer Security (TLS) Protocol URL, http://www.ietf.org/rfc/rfc4346.txt 2006.
[22] S. Evdokimov, B. Fabian, S. Kunz, Challenges for access control in knowledge federations, International Conference on Knowledge Management and Information Sharing (KMIS 2009), 2009, pp. 224–229.
[23] D.F. Ferraiolo, J.F. Barkley, D.R. Kuhn, A role-based access control model and reference implementation within a corporate intranet, ACM Transactions of Information Systems Security 2 (1) (1999) 34–64.
[24] D.F. Ferraiolo, D.R. Kuhn, R. Chandramouli, Role-based Access Control, Artech House, 2007.
[25] Gartner, New Gartner Hype Cycle Highlights Five High Impact IT Security Risks. Tech. Rep.; Gartner URL, http://www.gartner.com/it/page.jsp?id=496247 2006.
[26] S. Goel, H. Shawky, Estimating the market impact of security breach announcements on firm values, Information & Management 46 (7) (2009) 404–410.
[27] Z. Guo, F. Fang, A.B. Whinston, Supply chain information sharing in a macro prediction market, Decision Support Systems 42 (3) (2006) 1944–1958.
[28] A. Halevy, N. Ashish, D. Bitton, M. Carey, D. Draper, J. Pollock, et al., Enterprise information integration: successes, challenges and controversies, 2005 ACM SIGMOD International Conference on Management of Data, ACM New York, NY, USA, 2005, pp. 778–787.
[29] M. Hansen, A. Schwartz, A. Cooper, Privacy and identity management, IEEE Security and Privacy 6 (2008) 38–45.
[30] M. Hardwick, D.L. Spooner, T. Rando, K.C. Morris, Sharing manufacturing information in virtual enterprises, Communications of the ACM 39 (2) (1996) 46–54.
[31] D. Harris, L. Khan, R. Paul, B. Thuraisingham, Standards for secure data sharing across organizations, Computer Standards & Interfaces 29 (1) (2007) 86–96.
[32] Q. He, A. Antón, A framework for modeling privacy requirements in role engineering, 9th International Workshop on Requirements Engineering: Foundations of Software Quality, vol. 3, 2003, pp. 137–146.
[33] A.R. Hevner, S.T. March, J. Park, S. Ram, Design science in information systems research, MIS Quarterly 28 (1) (2004) 75–106.
[34] A. Jain, C. Farkas, Secure resource description framework: an access control model, 11th ACM Symposium on Access Control Models and Technologies (SACMAT'06), 2006, pp. 121–129.
[35] R. Kanneganti, P. Chodavarapu, SOA Security, Manning, 2008.
[36] W. Kim, D. Choi, S. Park, Intelligent product information search framework based on the semantic web, 3rd International Semantic Web Conference (ISWC), Hiroshima, Japan, 2004.
[37] M. Knechtel, J. Hladik, RBAC authorization decision with DL reasoning, IADIS International Conference WWW/Internet (ICWI'08), 2008.
[38] W. Knight, Security — built-in or bolted-on to the SOA? Infosecurity Today 2 (2) (2005) 38–40.
[39] D. Kulkarni, C.V. Ravishankar, OM — a tunable framework for optimizing continuous queries over data streams, 21st Brazilian Symposium on Databases (SBBD'06), 2006, pp. 206–220.
[40] S. Kunz, F. Brecht, B. Fabian, M. Aleksy, M. Wauer, Aletheia — improving industrial service-lifecycle management by semantic data federations, IEEE International Conference on Advanced Information Networking and Applications (AINA 2010), 2010, pp. 1308–1314.
[41] S. Kunz, S. Evdokimov, B. Fabian, B. Stieger, M. Strembeck, Role-based access control for information federations in the industrial service sector, 18th European Conference on Information Systems (ECIS), 2010.
[42] S. Kunz, B. Fabian, D. Marx, S. Müller, Engineering policies for secure inter-organizational information flow, 6th IEEE International Workshop on Vocabularies, Ontologies and Rules for The Enterprise (VORTE), IEEE EDOC 2011, 2011.
[43] H.L. Lee, V. Padmanabhan, S. Whang, Information distortion in a supply chain: the bullwhip effect, Management Science 43 (4) (1997) 546–558.
[44] S. Li, B. Lin, Accessing information sharing and information quality in supply chain management, Decision Support Systems 42 (3) (2006) 1641–1656.
[45] H. Li, M. Singhal, Trust management in distributed systems, IEEE Computer 40 (2) (2007) 45–53.
[46] J. Li, R. Sikora, M.J. Shaw, G.W. Tan, A strategic analysis of inter-organizational information sharing, Decision Support Systems 42 (1) (2006) 251–266.
[47] M. Lorch, S. Proctor, R. Lepro, D. Kafura, S. Shah, First experiences using XACML for access control in distributed systems, ACM Workshop on XML Security, 2003, pp. 25–37.
[48] A. Maier, H. Schnurr, Y. Sure, Ontology-based information integration in the automotive industry, 2nd International Semantic Web Conference (ISWC 2003), LNCS, vol. 2870, Springer, 2003, pp. 897–912.
[49] F. Massacci, N. Zannone, A model-driven approach for the specification and analysis of access control policies, On the Move to Meaningful Internet Systems (OTM), Springer, 2008, pp. 1087–1103.
[50] P. Mazzoleni, B. Crispo, S. Sivasubramanian, E. Bertino, XACML policy integration algorithms, ACM Transactions of Information System Security 11 (1) (2008) 1–29.
[51] U. Mbanaso, G. Cooper, D. Chadwick, S. Proctor, Privacy preserving trust authorization framework using XACML, International Symposium on World of Wireless, Mobile and Multimedia Networks, 2006, pp. 673–678.
[52] S.A. McIlraith, T.C. Son, H. Zeng, Semantic web services, IEEE Intelligent Systems 16 (2001) 46–53.
[53] J. Mendling, M. Strembeck, G. Stermsek, G. Neumann, An approach to extract RBAC models from BPEL4WS processes, 13th IEEE International Workshop on Enabling Technologies, 2004, pp. 81–86.
[54] T. Moses, eXtensible Access Control Markup Language (XACML) Version 2.0, OASIS Standard URL, http://docs.oasis-open.org/xacml/2.0/ 2005.
[55] Aletheia Project Web Site URL, http://www.aletheia-projekt.de 2011.
[56] Amazon EC2 Instance Types URL, https://aws.amazon.com/ec2/instance-types/ 2011.
[57] ISO/IEC 13335, Information technology – security techniques – management of information and communications technology security — part 1: concepts and models (ISO/IEC 13335–1:2004). 2004.
[58] Jena Project Web Site URL, http://www.openjena.org/ 2011.
[59] OASIS Security Assertion Markup Language (SAML) V2.0 Technical Overview URL, http://www.oasis-open.org/ 2008.
[60] OASIS Web Services Security: SOAP Message Security 1.1URL, http://docs.oasis-open.org/wss/v1.1/ 2006.
[61] OpenSSO Project Web SiteURL, https://opensso.dev.java.net/ 2010.
[62] SPARQL Query Language for RDF, W3C Recommendation URL, http://www.w3.org/TR/rdf-sparql-query/2008.
[63] The EPCglobal Architecture FrameworkFinal Version 1.3 Approved 19 March. 2009. URL, http://www.epcglobalinc.org/standards/architecture/.
[64] F.F. Nah, A study on tolerable waiting time: how long are web users willing to wait? Behaviour & Information Technology 23 (3) (2004) 153–163.
[65] G. Neumann, M. Strembeck, A scenario-driven role engineering process for functional RBAC roles, 7th ACM Symposium on Access Control Models and Technologies (SACMAT'02), 2002, pp. 33–42.
[66] A.C. O'Connor, R.J. Loomis, 2010 economic analysis of role-based access control, NIST Final Report, 2010.
[67] N. Panteli, S. Sockalingam, Trust and conflict within virtual inter-organizational alliances: a framework for facilitating knowledge sharing, Decision Support Systems 39 (4) (2005) 599–617 [Collaborative Work and Knowledge Management].
[68] M.P. Papazoglou, Service-oriented computing: concepts, characteristics and directions, 4th International Conference on Web Information Systems Engineering (WISE'03), IEEE Computer Society, 2003.
[69] K. Peffers, T. Tuunanen, M. Rothenberger, S. Chatterjee, A design science research methodology for information systems research, Journal of Management Information Systems 24 (2007) 45–77.
[70] P. Reddivari, Policy-based access control for an RDF store, Policy Management for the Web Workshop, 2005, pp. 78–83.
[71] Resource Description Framework (RDF), Concepts and abstract syntax, W3C recommendation URL, http://www.w3.org/TR/rdf-concepts/ 2008.
[72] G. Ryan, M. Valverde, Waiting online: a review and research agenda, Internet Research: Electronic Networking Applications and Policy 13 (2003) 195–205.
[73] S. Samaddar, S. Nargundkar, M. Daley, Inter-organizational information sharing: the role of supply network configuration and partner goal congruence, European Journal of Operational Research 174 (2) (2006) 744–765.
[74] R. Sandhu, D. Ferraiolo, R. Kuhn, The NIST model for role-based access control: towards a unified standard, 5th ACM Workshop on Role-Based Access Control, 2000, pp. 47–63.
[75] R. Sarathy, K. Muralidhar, Secure and useful data sharing, Decision Support Systems 42 (1) (2006) 204–220.
[76] M. Strembeck, Scenario-driven role engineering, IEEE Security and Privacy 8 (2010) 28–35.
[77] B. Thuraisingham, Building Trustworthy Semantic Webs, CRC Press, 2007.
[78] P. Trkman, K. McCormack, M.P.V. de Oliveira, M.B. Ladeira, The impact of business analytics on supply chain performance, Decision Support Systems 49 (3) (2010) 318–327.
[79] M. Wauer, D. Schuster, J. Meinecke, T. Janke, A. Schill, Aletheia — towards a distributed architecture for semantic federation of comprehensive product information, IADIS International Conference WWW/Internet. Rome, Italy, 2009.
[80] P. Windley, Digital Identity, O'Reilly Media, Inc., 2005
[81] C. Wolter, A. Schaad, C. Meinel, Deriving XACML policies from business process models, Web Information Systems Engineering Workshops (WISE 2007), LNCS, Springer, 2007, pp. 142–154.

[82] L. Zhou, F.B. Schneider, R.V. Renesse, COCA: a secure distributed online certification authority, ACM Transactions on Computer Systems (TOCS) 20 (4) (2002) 329–368.
[83] J. Zhou, S. Zhang, H. Zhao, M. Wang, SGII: towards semantic grid-based enterprise information integration, Grid and Cooperative Computing (GCC 2005), vol. 3795, Springer, 2005, pp. 560–565.

**Dr. Benjamin Fabian** is a senior researcher and project manager at the Institute of Information Systems, Humboldt-Universität zu Berlin, and coordinated security and privacy research in the Aletheia project. He holds a Diplom degree in Mathematics from the Free University of Berlin and a Ph.D. in Information Systems from Humboldt-Universität zu Berlin.

**Dr. Steffen Kunz** is a researcher at the Institute of Information Systems, Humboldt-Universität zu Berlin. He holds a Diplom degree in Betriebswirtschaftslehre (equivalent to Master of Business Administration) from the University of Mannheim and a Ph.D. in Information Systems from Humboldt-Universität zu Berlin.

**Sebastian Müller, M.Sc.**, is a researcher at the Databases and Information Systems Group, Department of Computer Science, Free University of Berlin. He holds a Master of Science in Information Systems from the Universities of Hohenheim and Stuttgart.

**Prof. Oliver Günther, Ph.D.**, is currently President of the University of Potsdam. He was Dean of the School of Business and Economics, Humboldt-Universität zu Berlin, and director of Humboldt's Institute of Information Systems. He holds a Diplom degree in Industrial Engineering from the University of Karlsruhe, and M.S. and Ph.D. degrees in Computer Science from the University of California at Berkeley.