



## Project Document Cover Sheet

| Project Information                          |  |                 |                  |
|--|--|-----------------|------------------|
| <b>Project Acronym</b>                       | SOFA   |                 |                  |
| <b>Project Title</b>                         | Service-Oriented Federated Authorization   |                 |                  |
| <b>Start Date</b>                            | 1 January 2010   | <b>End Date</b> | 31 December 2010 |
| <b>Lead Institution</b>                      | University of Oxford   |                 |                  |
| <b>Project Director</b>                      | Andrew Simpson   |                 |                  |
| <b>Project Manager &amp; contact details</b> | Andrew Simpson<br>Oxford University Computing Laboratory,<br>Wolfson Building, Parks Road, Oxford OX1 3QD<br>Andrew.Simpson@comlab.ox.ac.uk, |                 |                  |
| <b>Partner Institutions</b>                  | None   |                 |                  |
| <b>Project Web URL</b>                       | www.comlab.ox.ac.uk/projects/SOFA/   |                 |                  |
| <b>Programme Name (and number)</b>           | Access and Identity Management   |                 |                  |
| <b>Programme Manager</b>                     | Christopher Brown  |                 |                  |

| Document Name                       |  |   |                          |
|-------------------------------------|--|---|--------------------------|
| <b>Document Title</b>               | Final report                                       |   |                          |
| <b>Reporting Period</b>             | January—December 2010                              |   |                          |
| <b>Author(s) &amp; project role</b> | Andrew Simpson, Project Director                   |   |                          |
| <b>Date</b>                         | January 20, 2011                                   | <b>Filename</b>   | final_report_updated.doc |
| <b>URL</b>                          | NA   |   |                          |
| <b>Access</b>                       | <input type="checkbox"/> Project and JISC internal | <input checked="" type="checkbox"/> General dissemination |                          |

| Document History |                   |   |
|------------------|-------------------|---|
| Version          | Date              | Comments  |
| 1.0              | December 16, 2010 | First draft   |
| 1.1              | January 12, 2011  | Second draft  |
| 1.2              | January 20, 2011  | Final version—following input from the JISC Programme Manager |



## **JISC Final Report**

### **SOFA: Service-Oriented Federated Authorization**

**Andrew Simpson,\* David Power, and Mark Slaymaker**

**January 12, 2011**

\*Andrew.Simpson@comlab.ox.ac.uk  
Oxford University Computing Laboratory,  
Wolfson Building, Parks Road, Oxford OX1 3QD

## Table of Contents

|                       |         |
|-----------------------|---------|
| Acknowledgements      | Page 4  |
| Executive Summary     | Page 5  |
| 1 Background          | Page 6  |
| 2 Aims and Objectives | Page 8  |
| 3 Methodology         | Page 9  |
| 4 Implementation      | Page 10 |
| 5 Outputs and Results | Page 13 |
| 6 Outcomes            | Page 15 |
| 7 Conclusions         | Page 17 |
| 8 Implications        | Page 18 |
| References            | Page 19 |

## **Acknowledgements**

The SOFA (Service-Oriented Interoperability Framework) project was funded by JISC, through the Access Identity and Management Programme. The project took an application-led approach—without the input of collaborators at various stages, it would have been impossible for the project team to succeed to the degree that they did. As such, we are grateful to each of the following for their enthusiastic engagement: Professor Paul Jeffreys, Richard Dunnaway, Keith Zimmerman, Professor David Gavaghan, Dr James Osborne, Professor Gordon Wilcock, and Grzegorz Agacinski.

## Executive Summary

The SOFA (Service-Oriented Federated Authorization) project was funded via JISC's Access Identity and Management project, and ran from January—December 2010. The project was concerned with the development of a service-oriented framework to support the secure sharing and aggregation of data within distributed, heterogeneous contexts. The project built upon previous work undertaken within the TSB-funded Generic Infrastructure for Medical Informatics (GIMI) project.

The primary aim of SOFA was to “deliver tools and technologies to address issues of authorization interoperability within virtual organizations.” That is to say, we wished to develop tools and technologies to allow collaborators to share data sets without being prescriptive as to how they might express access control policies.

At its simplest, the aim of SOFA was to “lift” the interoperability afforded by the middleware of GIMI. The solution delivered through GIMI allowed collaborators to integrate diverse data sets, independent of underlying data models or technologies. So, for example, cancer researchers could share data sets, even though one collection might utilize a high-end data warehouse product and another might be based on an open source XML database solution. One drawback of the GIMI solution was the requirement for data owners to prescribe access control policies in terms of the policy language, XACML (eXtensible Access Control Markup Language). XACML has many benefits—it is, for example, very expressive. But these benefits come at a cost: taking advantage of XACML's flexibility can give rise to complex policies that verge on the incomprehensible. Thus, at the heart of this project, was the drive to allow collaborators to bring together their data sources in a straightforward way, while allowing them to leverage their authorization approach of choice: whether it be a policy language such as XACML, a more accessible role-based approach, or through simple access control lists.

The approach taken was application-oriented: three case studies were used to drive and validate the work. The first was concerned with academic administration (specifically linking student admissions and progression data); the second pertained to Dementia research; the application area of the third was Systems Biology modelling.

As well as stabilizing and extending the middleware of GIMI, tool support for the construction, analysis, transformation and deployment of access control policies was developed through this project. Several publications resulted from the work,<sup>1</sup> some of which were concerned with particular applications and some of which were concerned with the underlying theory. In addition, an open source release of the SOFA extension to the GIMI middleware framework has been made available.<sup>2</sup>

---

<sup>1</sup> See <http://www.comlab.ox.ac.uk/projects/publications/date/SOFA.html>

<sup>2</sup> See <http://www.comlab.ox.ac.uk/projects/SOFA/>

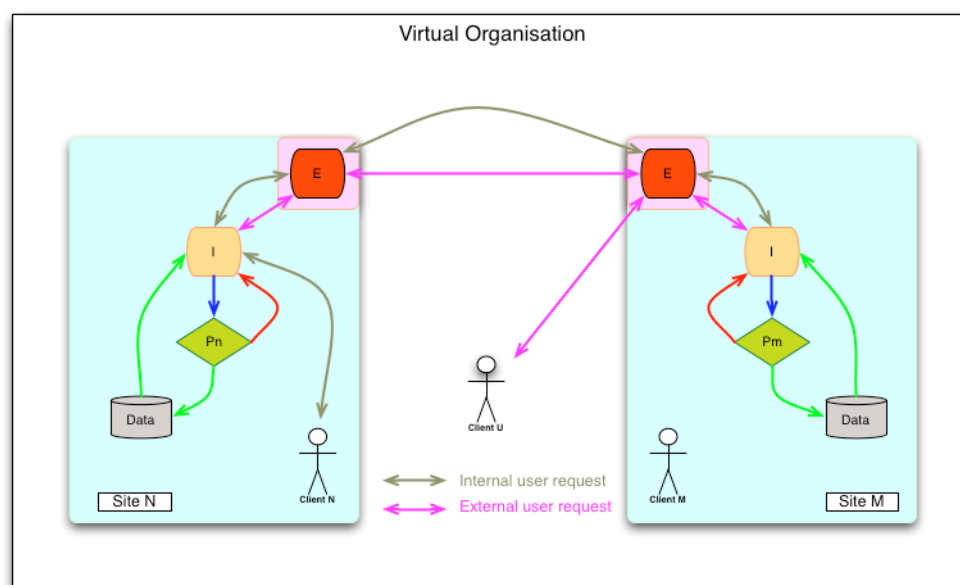
## 1. Background

It is not unusual for large organizations to have disparate, diverse sources of data. In an academic environment, for example, there might be a central student administration system, a separate admissions system, and departments running their own stand-alone databases. Then there might be separate financial systems and personnel systems. In such contexts, there will be tens—possibly hundreds—of administrators who regularly need to combine data to conduct everyday tasks. A very fortunate few will work in an environment in which technological solutions are in place to allow them to navigate these diverse systems; some will rely on friendly IT support staff to write low-level scripts; the majority will resort to asking for spreadsheets to be emailed or DVDs to be posted—with all of the inherent security implications associated with such an approach; some will simply give up.

There are other examples, of course, in which organizations might wish to integrate disparate data sources: researchers that wish to pool their data sets; newly merged companies that wish to combine their data; and organizations that see potential in aggregating information from their local database with publically available data sets.

The work of SOFA is driven by the requirement to facilitate the integration of existing data sources, but in a way that is relatively lightweight: there should be no need for organizations to change or throw away existing systems, processes or data models. In addition, data owners should be in a position to prescribe exactly which users and applications can see their data and under what circumstances.

The origins of this work can be found in [1]. There, use cases and requirements—as well as a simple architecture—for the secure sharing and aggregation of data from diverse data sources are described. Consider, as an example, the following diagram.



Here, a virtual organisation—spread across two or more geographically or physically distinct units—is characterised in terms of external interfaces, internal interfaces and policies, with data nodes communicating via their external interfaces and data being accessed via internal interfaces. Access to data is regulated by policies; importantly, each data source 'owner'

has control over their data—meaning that a data owner shares only the data that they wish to. As such, the approach acts as both a secure gateway and data integration framework.

The TSB-funded project, *GIMI* (Generic Infrastructure for Medical Informatics) [2], gave the first opportunity for the implementation of these ideas. There, a middleware framework, called *sif* (service-oriented interoperability framework) [3,4,5], was developed to support a wide variety of healthcare-related applications, including: support of, for example, Colorectal Cancer and Breast Cancer research; a system to facilitate the secure transfer of data between desktop systems and hand-held devices to help Asthma and Diabetes patients manage their conditions; and a tool to support the secure round-trip of images from remote devices, to desktops, to servers (for image analysis), and finally back to remote devices.

While motivated, originally, by the needs of healthcare applications, the approach is data-agnostic—meaning that the framework has the potential to be of benefit in any context in which there is a requirement either to integrate data from diverse data sources or to provide some assurance that potential data access is *appropriate*.

At its simplest, the aim of SOFA was to “lift” the interoperability afforded by *sif*. Prior to SOFA, *sif* required data owners to prescribe access control policies in terms of the policy language, XACML (eXtensible Access Control Markup Language).<sup>3</sup> XACML has many benefits—it is, for example, very expressive. But these benefits come at a cost: taking advantage of XACML’s flexibility can give rise to complex policies that verge on the incomprehensible. Thus, SOFA was concerned with extending GIMI to allow data owners and policy writers to deploy access control policies in their paradigm of choice: whether it be a policy language such as XACML, a more accessible role-based approach, or through simple access control lists. Accessible tool support allows for the construction, analysis, transformation and deployment of policies. Thus, the SOFA solution offers the potential to integrate data from heterogeneous data sources, which are supported by different underlying technologies and protected by different authorization mechanisms.

The work is, of course, timely in that increasing amounts of data are being collected and combined—at precisely the time that the public’s consciousness as to issues of data security and privacy has been pricked. Furthermore, the Information Commissioner’s Office’s recently increased powers means that effective data management strategies must be of concern to all organizations that deal with personal data—including all higher education institutions.

---

<sup>3</sup> See [www.oasis-open.org/committees/xacml/](http://www.oasis-open.org/committees/xacml/)

## 2. Aims and Objectives

Recalling the project proposal, the aim of the project was stated quite simply:

“The aim of this project is to deliver tools and technologies to address issues of authorization interoperability within virtual organizations; specifically, we are concerned with ensuring fine-grained access to, and federation of, data within virtual organizations.”

The anticipation was that we would measure success against the following criteria.

1. Is it possible to support federation across at least three disparate data sources, one of which uses an XACML implementation for authorization, one of which uses an RBAC implementation, and one of which uses an ACL?
2. Does the policy construction tool allow the construction of policies for the RBAC and ACL paradigms?
3. Is it possible to support a definitive use case, linking three or more heterogeneous data sources from within the University of Oxford, to enable administrators to undertake analysis that would otherwise be impossible, thereby delivering tangible benefits to the organization?
4. Has a second definitive use case been identified and supported?
5. Currently, sif supports authentication via standard X.509 certificates. Has the feasibility of supporting authentication via, for example, Shibboleth been determined?

Early in the project, there were two changes to these criteria. Criterion 3 was adapted: we built a more complex use case than had originally been anticipated—but at the expense of using fictional data.

Criterion 5 proved to be met very quickly: the answer was “yes, we have established the feasibility—it can’t be supported without extensive, additional work.”

### 3. Methodology

In previous projects, including the aforementioned GIMI, as well as e-DiaMoND [6] and NeuroGrid [7], the SOFA team had taken an iterative, test-led approach to development—with close engagement with end-users being at the heart of this activity. A similar approach was taken here. Further, in each of these projects, a collection of diverse use cases was used to drive and validate the work; hence the application-led approach to SOFA.

The first use case built upon ongoing collaboration with UAS (University Administration and Services) at the University of Oxford. The driver here was to aggregate student admission data (held on a stand-alone) database, with student progression data (held in an Oracle data warehouse) so as to allow administrators to determine how effective admissions tests were as a predictor of future performance.

The second use case emerged from engagement with the Oxford Biomedical Research Centre.<sup>4</sup> The OPTIMA (Oxford Project To Investigate Memory and Ageing) project<sup>5</sup> has been collecting data on Dementia patients for over 20 years, with that data being stored in an SQL database. In parallel, a group based at the Neuropathology group based at the Department of Clinical Neurology has been collecting post-mortem data (stored in Microsoft Excel spreadsheets) on patients—many of whom appear in the OPTIMA data sets. Clearly, to be able to link these collections would be of value to both groups.

The third use case emerged from engagement with the Oxford Centre for Integrative Systems Biology (OCISB).<sup>6</sup> The objective of the Cell Based Chaste (Cancer, Heart and Soft Tissue Environment) (or *Cancer Chaste*) initiative [8,9] is to develop a mathematical and computational model that bridges across these spatial and temporal scales within a single, generic modelling framework. In conjunction with collaborators from Chaste, the third use case was concerned with facilitating the sharing of data and models to support the mathematical modelling process.

The first three months (January—March 2010) of the project were concerned with engaging with the three user groups so as to capture their intentions, in terms of data sharing. The focus of the second quarter (April—June 2010) was on delivering initial, working applications to the groups.

The focus of the second half of the project was on refining the two core technical deliverables: the policy construction tool and the open source release of the SOFA extension to *sif*.<sup>7</sup>

A fourth—unexpected—use case provided an additional means of collaboration. The Software Engineering Programme at the University of Oxford delivers one-week intensive courses in a variety of software engineering and security subjects to professional software engineers. One of these courses—Data Security—was run for the first time in mid-2010. The SOFA extension to *sif*, as well as an early version of the policy construction tool, was used to support the course's running case study exercises. This meant that the systems were exposed to 18 users simultaneously exposing data sources, constructing and deploying access control policies, and executing federated queries.

---

<sup>4</sup> See <http://www.oxfordbrc.org/>

<sup>5</sup> See <http://www.medsci.ox.ac.uk/optima>

<sup>6</sup> See [www.sysbio.ox.ac.uk/](http://www.sysbio.ox.ac.uk/)

<sup>7</sup> See <http://www.comlab.ox.ac.uk/projects/SOFA/>

## 4. Implementation

The project's work packages were as follows:

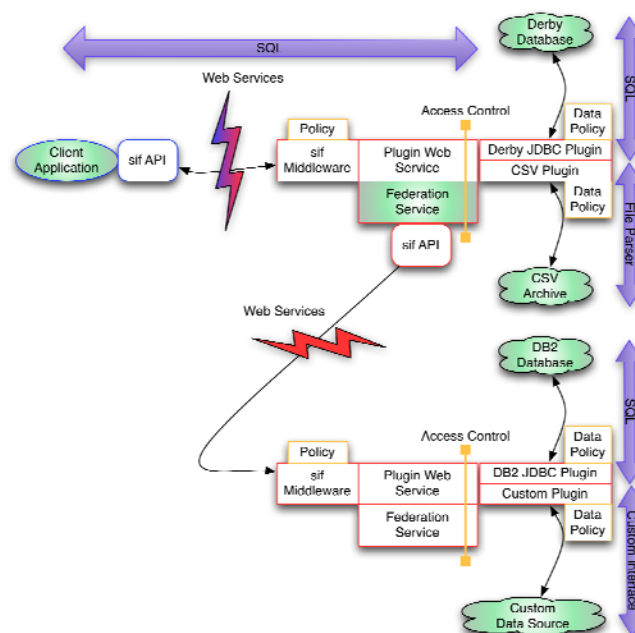
1. Project management
2. Use cases and requirements
3. Middleware extension and refinement
4. Tool development
5. Application development and support
6. Testing and validation
7. Community engagement and dissemination

Our original intention with respect to the capture of requirements, the validation of systems and the dissemination of results, was to hold a series of workshops. Due to variety of factors—not least the relatively short time-scales associated with the project—this proved not to be possible. Rather, a series of meetings with the individual user groups was carried out. Arguably, this approach proved to be more useful than a series of workshops would have been.

Given the nature of the project, there is little of interest to say with respect to implementation: requirements were captured; code was written and tested; applications were deployed. Perhaps, though, it is worth considering the key technical aspects of the project: the middleware and the policy construction toolset.

### The middleware

sif makes data accessible via “plug-ins”—standard interfaces. There are three kinds of plug-in: data, file and algorithm; it is data plug-ins that facilitate access to databases (our concern in the following). If, in a distributed context, a user runs a query across several data nodes, then the middleware will distribute that query to the nodes and aggregate the results. The reason that sif can expose any relational database is that it makes no assumptions about structure or semantics; this, of course, makes the task of federation much easier. The architecture of a typical deployment is given below.

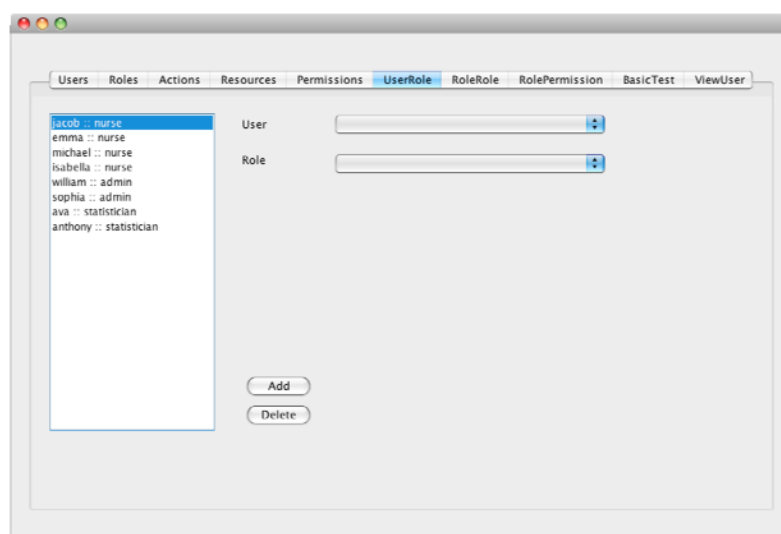


sif can be thought of as being comprised of three parts: the core middleware, the plug-ins, and the client-side API. The core middleware manages the installed plug-ins, giving them a standard interface to be written against. It also provides a federation service to facilitate the construction of queries against multiple data sources. As all data is represented as if it were a standard SQL database, these queries take the form of SQL queries across distinct data sources each exposed via a separate plug-in. The access control framework enforces policies created by the owners of the data and the owners of the machine on which sif is being hosted. This allows data owners to restrict the data they expose to users, and server owners to control who the permitted users of services are. The middleware has built-in capabilities for transferring files and data: installing, removing and updating plug-ins; advertising and defining resources exposed by plug-ins; and providing system status information. The core middleware exposes this functionality through a number of web services, all of which utilise strong cryptography to ensure privacy. The client-side API is a wrapper around web service calls to create the simplest possible interface for a new application developer to implement against; it also provides a number of helper functions to assist in common tasks.

### The policy construction toolset

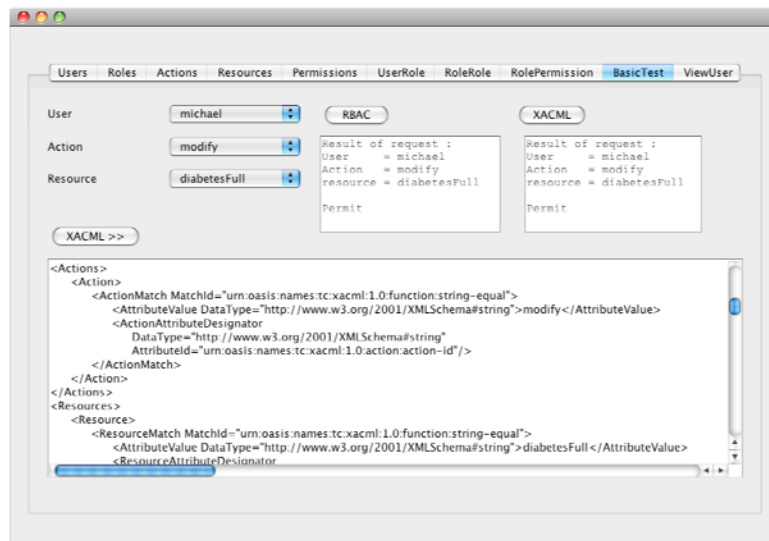
The policy construction toolset was developed in Java and allows for the capture of access control policies in terms of, for example, Role-Based Access Control (RBAC), access control lists, or XACML. Should a role-based policy be constructed, it might subsequently be deployed to facilitate access to a plug-in; alternatively, it might first be converted to XACML prior to deployment. (The potential to map from RBAC to XACML will be our focus in the following discussion.) An example of the use of the policy tool is given in [10]; we reprise that description in the following.

The user interface is divided into a number of panes—either capturing a single aspect of RBAC (in the case of that approach) or providing the facility to translate the RBAC model into XACML. The first few panes support the collection of basic information about users, roles, actions, and resources; the next pane allows the definition of permissions. The next three panes define the main RBAC relationships: the relationship between users and roles, the relationship between roles and permissions, and the role “hierarchy” (if there is one). The user—role pane is shown below.



The final pane (in this case) provides access to the mechanisms necessary to perform the translation from RBAC to XACML, along with a method for testing that the XACML generated is consistent with the original RBAC. This is achieved by checking the result of applying a given request to each access control representation produces the same result.

The following screenshot illustrates the case where a user, Michael, is requesting to modify the resource; the second case is where Ava is attempting to modify the resource. It can be seen that both the RBAC and XACML evaluations of the request result in the action requested being permitted; testing for denial results in a denial from both RBAC and XACML policies.



## 5. Outputs and Results

There have been four major outputs as a result of this work:

1. The extension to the sif middleware and its open source release
2. The policy construction tool
3. The development and deployment of three applications
4. A series of publications

We have considered the first two of these in previous sections (and we return to our success criteria in the next section); in this section we consider the applications and publications.

### 5.1 Applications

As outlined in Section 3, three applications were delivered as part of this project. One application—pertaining to support for Systems Biology research is described in [11]; we consider the other two in the following.

#### Student administration

The focus of the system was to demonstrate the feasibility of securely aggregating data from a central data warehouse with other, localized systems. The particular focus here was linking student progression data with admissions data.

Two entry cohorts were considered. The plug-in for the data warehouse exposed basic student information, alternative personal identifiers (UCAS identifiers)—to link on; course information; and outcomes of each year.

The sif API offers the opportunity for a wide variety of functionality; in this case, two applications were developed. The first offered a portal-style interface, supporting canned queries; the second took a query-builder approach, which was more flexible—but required a degree of query-writing ability. Both applications allowed users to execute federated queries across the data sources, and, subsequently, export the results for further statistical analysis.

The particular task that was undertaken was to compare Thinking Skills Assessment (TSA)<sup>8</sup> scores with first year outcomes so as to determine the effectiveness of the TSA tests as a predictor of future performance. The system proved to be effective in this respect.

#### Dementia research

The OPTIMA (Oxford Project To Investigate Memory and Ageing) project has been investigating the causes of dementia for over 20 years. The OPTIMA database contains all of the data captured over that period in one system—as opposed, for example, to having one database per trial; thus, the database not only includes data pertaining to studies that concluded many years ago (often associated with now deceased patients), but is also a live system, with new data associated with ongoing trials being inserted continually. As such, it is a resource looked upon with envy by other researchers—within both Oxford and the wider community, with access being requested on a regular basis. Taken together, these facts made OPTIMA an ideal case study against which the potential value of the deliverables SOFA might be evaluated.

---

<sup>8</sup> See <http://www.admissionstests.cambridgeassessment.org.uk/adt/tsaoxford>

The focus here was linking the main OPTIMA database with related data sources: the aforementioned post-mortem data, as well as a file system storing scans associated with OPTIMA patients.

What has been delivered is a system that allows users to query the live database without having to go via the data manager (other than with respect to obtaining credentials). The portal interface generates a query builder—tailored to the user's permissions—which allows the user to choose their attributes of interest, filtering as necessary; the resulting data can be either viewed on screen or downloaded as a CSV file (for subsequent analysis via Excel or the user's statistical package of choice). The application is unaffected by any changes to the underlying data schema as the query builder is built dynamically on the basis of the current database schema and that user's access permissions.

## 5.2 Publications

The following have all been published as either a direct or indirect consequence of this project:

- A. C. Simpson, M. A. Slaymaker, and D. J. Gavaghan. On the secure sharing and aggregation of data to support systems biology research. In Proceedings of the 7th International Conference on Data Integration in the Life Sciences (DILS 2010), pages 58—73. Springer-Verlag Lecture Notes in Computer Science volume 6254, 2010.
- M. A. Slaymaker, D. J. Power, and A. C. Simpson. Formalising and validating RBAC-to-XACML translation using lightweight formal methods. In Proceedings of the Second International Conference on Abstract State Machines, Alloy, B and Z (ABZ 2010), pages 349—362. Springer-Verlag Lecture Notes in Computer Science volume 5977, 2010.
- D. J. Power, M. A. Slaymaker, and A. C. Simpson. Automatic conformance checking of role-based access control policies via Alloy. To appear in the Proceedings of the 3<sup>rd</sup> International Conference on Engineering Secure Software and Systems (ESSOS 2011), 2011.
- M. A. Slaymaker, D. J. Power, and A. C. Simpson. On the formal harmonisation of distributed, heterogeneous access control policies. Submitted to POLICY 11.

## 6. Outcomes

We assess the outcomes of the project work, first, in terms of the assessment criteria of Section 2, and, second, by considering potential beneficiaries.

### 6.1 Measuring against the success criteria

We consider each in turn.

#### **1. Is it possible to support federation across at least three disparate data sources, one of which uses an XACML implementation for authorization, one of which uses an RBAC implementation, and one of which uses an ACL?**

This was demonstrated via the Data Security case study alluded to in Section 3. There, six data sources formed a virtual organization, with access to two databases being via an ACL, access to two others being via RBAC, and access to the final two being via XACML.

The case study involved 18 students (divided into six groups of three), with each group being responsible for managing access to a single database—one of which was an admissions database; one of which was a student management database; two of which were college databases; and two of which were department databases. It should go without saying that the data was fictional.

The groups were required to write and deploy plug-ins to expose their data sources, construct and deploy access control policies to facilitate access by other groups, and run federated queries via simple applications. As one might expect, the users were demanding—they were all software engineering professional that had paid for a one-week course on Data Security. The level of testing afforded through this process was invaluable; the number of minor bugs detected was, thankfully, minimal.

#### **2. Does the policy construction tool allow the construction of policies for the RBAC and ACL paradigms?**

As discussed in Section 4, we have gone further in that the policy construction tool also allows the modeling, analysis and transformation of policies. The facility to model and analyze policies prior to deployment helps give assurance to policy writers and data owners that their policies are, in some sense, appropriate.

#### **3. Is it possible to support a definitive use case, linking three or more heterogeneous data sources from within the University of Oxford, to enable administrators to undertake analysis that would otherwise be impossible, thereby delivering tangible benefits to the organization?**

This success criterion was not met fully—due to a variety of non-technical, organisational issues. However, the application described in Section 5.1 did allow administrators to undertake analysis in a straightforward fashion that would have otherwise have been time-consuming.

#### **4. Has a second definitive use case been identified and supported?**

Having supported three “official” applications, as well as the Data Security case study, we feel that we have met this criterion.

**5. Currently, sif supports authentication via standard X.509 certificates. Has the feasibility of supporting authentication via, for example, Shibboleth been determined?**

The sif middleware uses X.509 certificates both as a means of authentication and a source of attributes for making access control decisions. Integrating with a single sign-on system that uses Shibboleth (or related technologies) would require a fundamental change to the “ticketing” system that the middleware uses to pass credentials between servers. If this change were to be made, it would then be possible for access control decisions to be based on the results of attribute queries. While the changes that would be required are feasible, they would require significant redevelopment work (in the order of 4—5 months), and, as such, were not undertaken within this project.

**6.2 Potential beneficiaries**

The potential beneficiaries of this work are numerous—in both the public, private and “third” sectors. Effectively, the deliverables might be utilised in any context in which there is a need to share and aggregate data in a lightweight, relatively straightforward fashion. The plug-in mechanism means that any (standard) data source can be exposed; the policy construction tool allows policies to be deployed quickly; the API and existing templates mean that straightforward, generic applications can be utilised to construct and execute federated queries.

Assuming no “gotchas”, i.e. no networking difficulties, standard data formats, no complex application requirements, etc., it is possible to deploy a brand new system within a matter of days; adding subsequent data sources to that system can be achieved within a matter of minutes.

## 7. Conclusions

As a result of JISC funding, we have been able to stabilize and extend the sif middleware framework that was developed through the TSB-funded GIMI project. We have added access control extensions to provide support for federated authorization; we have developed a policy construction toolset to support the construction, analysis, transformation and deployment of access policies; we have made the SOFA extension available as an open source release. To drive and validate this work, we have engaged with end-users from three communities; additional validation for the technologies came via their use in support of a one-week intensive course on Data Security.

A number of lessons emerged as a result of this work. We consider each in turn below.

First, an iterative approach to development—with short bursts—is essential in development projects of this length (the duration of SOFA was 12 months). We were fortunate in that risks were mitigated in two ways: we were building on existing technology, and we were leveraging existing collaborative relationships. Together, this allowed us to “hit the ground running”; it also allowed for early feedback from end-users. Had this not been the case, we would have had difficulty in terms of successful delivery.

Second, and this will come as no surprise to anyone who has ever engaged in collaborative research, the expectations of end-users are increasing—fuelled by the ubiquity of the Web. Understandably, users are coming to expect applications that have the look and feel—and stability—afforded by the likes of Amazon, Expedia, etc. No matter how clever or useful an underlying technology is, the interface still matters greatly to a significant number of users. This will prove to be an interesting challenge to computer scientists and software engineers engaged in collaborative research projects in the coming years.

Third, the ability to model—and test—policies is a wonderful thing. When we embarked upon this project, the access control extensions to sif were seen as the core activity and the policy tool endeavour was considered a secondary aspect. The quality of access control tooling is notoriously patchy; particularly so with respect to XACML. To be able to give policy writers the opportunity to model *and test* policies prior to deployment allowed them to realise the limitations of their original intentions, as well as mistaken assumptions—it also granted a level of assurance and faith in the policies that were to be deployed. Fundamentally, though, it changed the process of writing and deploying policies.

## 8. Implications

We will pursue—in collaboration with others—various avenues in taking this work forward.

First, and most obviously, we will continue to seek further potential applications. To date, the majority of applications have been associated with research—primarily clinical research; however, the very nature of the technologies developed means that they are applicable in many other contexts. Our understanding of IT infrastructures within higher education institutions leads us to suggest that the work of this project has the potential for application in that sector. We would hope to continue to work with JISC in this respect so as to identify potential beneficiaries.

Second, in conjunction with Isis Innovation (the University of Oxford's technology transfer division), we will be looking at the potential for the exploitation and commercialisation of the work undertaken within SOFA. Initial discussions in this respect have been very promising.

Third, we will continue to pursue extensions to the tool support for policy construction and analysis. The system currently has limited support for what might be termed *evolving access control* (see, for example, [12])—access control policies that can change, automatically, on the basis of observed actions. Thus, there is the potential to support policies of the nature of “Adam can access data from exactly one of A, B or C”, “Bob can download files only if there is sufficient network capacity”, or “if there has been no contact from Charlie for 30 minutes, then access from his device should be denied”. SOFA's policy toolset allows for the construction and analysis of static policies; we would hope to be able to extend that support to dynamic policies in the near future. Initial support for the modelling and analysis of such policies is described in [13].

## References

- [1] D. J. Power, E. A. Politou, M. A. Slaymaker, and A. C. Simpson. Towards secure grid-enabled healthcare. *Software: Practice and Experience*, 35(9): 857—871, 2005.
- [2] A. C. Simpson, D. J. Power, D. Russell, M. A. Slaymaker, V. Bailey, C. E. Tromans, J. M. Brady, and L. Tarassenko. GIMI: the past, the present, and the future. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, 368: 3891—3905, 2010.
- [3] A. C. Simpson, D. J. Power, D. Russell, M. A. Slaymaker, G. Kouadri-Mostefaoui, X. Ma, and G. Wilson. A healthcare-driven framework for facilitating the secure sharing of data across organisational boundaries. *Studies in Health Technology and Informatics*, 138: 3—12, 2008.
- [4] M. A. Slaymaker, D. J. Power, D. Russell, G. Wilson, and A. C. Simpson. Accessing and aggregating legacy data sources for healthcare research, delivery and training. In *Proceedings of the 2008 ACM symposium on Applied Computing (SAC 2008)*, pages 1317—1324, 2008.
- [5] M. A. Slaymaker, D. J. Power, D. Russell, and A. C. Simpson. On the facilitation of fine-grained access to distributed healthcare data. In *Proceedings of Secure Data Management 2008*, pages 169—184. Springer-Verlag Lecture Notes in Computer Science volume 5159, 2008.
- [6] J. M. Brady, D. J. Gavaghan, A. C. Simpson, M. Mulet-Parada, and R. P. Highnam. eDiaMoND: A grid-enabled federated database of annotated mammograms. In F. Berman, G. C. Fox, and A. J. G. Hey, editors, *Grid Computing: Making the Global Infrastructure a Reality*, pages 923—943. Wiley Series, 2003.
- [7] J. Geddes, S. Lloyd, A. C. Simpson, M. Rossor, N. Fox, D. Hill, J. V. Hajnal, S. Lawrie, A. McIntosh, E. Johnstone, J. Wardlaw, D. Perry, R. Procter, P. Bath, and E. Bullimore. NeuroGrid: using grid technology to advance neuroscience. In *Proceedings of the 18th IEEE Symposium on Computer-Based Medical Systems*, pages 570—573. IEEE Computer Society Press, 2005.
- [8] J. Pitt-Francis, P. Pathmanathan, M. O. Bernabeu, R. Bordas, J. Cooper, A. G. Fletcher, G. R. Mirams, P. Murray, J. M. Osborne, A. Walter, S. J. Chapman, A. Garny, I. M. M. van Leeuwen, P. K. Maini, B. Rodriguez, S. L. Waters, J. P. Whiteley, H. M. Byrne, D. J. Gavaghan. Chaste: a test-driven approach to software development for biological modelling. *Comp. Phys. Comm.* 180: 2452—2471, 2009.
- [9] I. M. M. van Leeuwen, G. R. Mirams, A. Walter, A. Fletcher, P. Murray, J. M. Osborne, S. Varma, S. J. Young, J. Cooper, J. Pitt-Francis, L. Momtahan, P. Pathmanathan, J. P. Whiteley, S. J. Chapman, D.J. Gavaghan, O. E. Jensen, J. R. King, P. K. Maini, S. L. Waters, H. M. Byrne. An integrative computational model for intestinal tissue renewal. *Cell Proliferation* 42: 617—636, 2009.
- [10] D. J. Power, M. A. Slaymaker, and A. C. Simpson. Automatic conformance checking of role-based access control policies via Alloy. To appear in the *Proceedings of the 3<sup>rd</sup> International Conference on Engineering Secure Software and Systems (ESSOS 2011)*, 2011.

- [11] A. C. Simpson, M. A. Slaymaker, and D. J. Gavaghan. On the secure sharing and aggregation of data to support systems biology research. In Proceedings of the 7<sup>th</sup> International Conference on Data Integration in the Life Sciences (DILS 2010), pages 58—73. Springer-Verlag Lecture Notes in Computer Science volume 6254, 2010.
- [12] A. C. Simpson, C. Sieunarine, D. J. Power, M. A. Slaymaker, and D. Russell. Evolving access control: models and Implementations. Submitted to CAISE 2011.
- [13] D. J. Power, M. A. Slaymaker, and A. C. Simpson. Conformance checking of dynamic access control policies. Submitted to FM 2011.