

Project Identifier: RAPTOR

Version: 1.0

Contact: Rhys Smith ([smith@cardiff.ac.uk](mailto:smith@cardiff.ac.uk)) & Graham Mason ([gmason@kidderminster.ac.uk](mailto:gmason@kidderminster.ac.uk))

Date: 31/3/2011



## JISC Final Report

Project Information			
<b>Project Identifier</b>	RAPTOR		
<b>Project Title</b>	Retrieval, Analysis, and Presentation Toolkit for usage of Online Resources (RAPTOR)		
<b>Project Hashtag</b>	#raptor		
<b>Start Date</b>	01/01/2010	<b>End Date</b>	31/03/2011
<b>Lead Institution</b>	Cardiff University		
<b>Project Director</b>	Rhys Smith		
<b>Project Manager</b>	Graham Mason		
<b>Contact email</b>	<a href="mailto:smith@cardiff.ac.uk">smith@cardiff.ac.uk</a> & <a href="mailto:gmason@kidderminster.ac.uk">gmason@kidderminster.ac.uk</a>		
<b>Partner Institutions</b>	Kidderminster College		
<b>Project Web URL</b>	<a href="http://iam.cf.ac.uk/RAPTOR">http://iam.cf.ac.uk/RAPTOR</a>		
<b>Programme Name</b>	Identity and Access Management		
<b>Programme Manager</b>	Christopher Brown		

Document Information			
<b>Author(s)</b>	Rhys Smith, Graham Mason		
<b>Project Role(s)</b>	Project Director, Project Manager		
<b>Date</b>	31/03/2011	<b>Filename</b>	RAPTOR-FinalReport-v1.0
<b>URL</b>	<a href="http://iam.cf.ac.uk/RAPTOR">http://iam.cf.ac.uk/RAPTOR</a>		
<b>Access</b>	This report is for general dissemination		

Document History		
Version	Date	Comments
0.1	31/3/2011	First Draft of report, Circulated to team and CB
0.2	13/6/2011	Second Draft: More conclusions, etc
1.0	6/7/2011	Final Version

## Table of Contents

<b>ACKNOWLEDGEMENTS .....</b>	<b>3</b>
<b>1 PROJECT SUMMARY .....</b>	<b>3</b>
<b>2 MAIN BODY OF REPORT .....</b>	<b>3</b>
2.1 PROJECT OUTPUTS AND OUTCOMES .....	3
2.2 HOW DID YOU GO ABOUT ACHIEVING YOUR OUTPUTS / OUTCOMES? .....	4
2.2.1 <i>Methodology and Work Plan</i> .....	4
2.2.2 <i>Aims and Objectives</i> .....	5
2.3 WHAT DID YOU LEARN? .....	5
2.4 IMMEDIATE IMPACT .....	6
2.5 FUTURE IMPACT .....	6
<b>3 CONCLUSIONS .....</b>	<b>6</b>
<b>4 RECOMMENDATIONS .....</b>	<b>6</b>
<b>5 IMPLICATIONS FOR THE FUTURE .....</b>	<b>7</b>
<b>6 REFERENCES .....</b>	<b>7</b>
<b>7 APPENDICES .....</b>	<b>7</b>
<b>APPENDIX 1 – RAPTOR ARCHITECTURE .....</b>	<b>8</b>

## Acknowledgements

JISC funded the RAPTOR project as a part of its Access and Identity Management (AIM) programme. The authors of this document would like to thank Chris Brown, RAPTOR's JISC programme manager, for his input. They would also like to thank the members of the RAPTOR Governance Group – Nicole Harris, David Harrison, Alec McSkimming, and Sue Stevens; particularly David Harrison both for chairing the group and for chairing its last meeting while officially retired!

## 1 Project Summary

Given the current economic climate and likelihood of tightening funding, understanding the usage of e-resources is becoming increasingly important as it allows an institution to understand which resources they need to keep subscribing to, and those which they may wish to unsubscribe from - potentially resulting in real-world cost savings.

This project therefore built a software toolkit for reporting e-resource usage statistics (from the Shibboleth Identity Provider, EZProxy, and OpenAthens LA software) in a user-friendly manner suitable for non-technical staff.

## 2 Main Body of Report

### 2.1 Project Outputs and Outcomes

*Note that all Project Deliverables are available at:*

<http://iam.cf.ac.uk/trac/RAPTOR/wiki/Project/Documentation>

Output / Outcome Type	Brief Description and URLs (where applicable)
RAPTOR Software	The core RAPTOR software, consisting of information collector agents and an accounting server. Enables an institution to collect and analyse electronic resource usage data. Also enables federations to collect aggregated statistical information. Available at <a href="http://iam.cf.ac.uk/Raptor">http://iam.cf.ac.uk/Raptor</a>
Reporting Analysis Report	A report describing what reporting information relevant software makes available through various means, and what reporting software is currently available and their capabilities. Available at <a href="https://iam.cf.ac.uk/trac/RAPTOR/wiki/Project/Documentation">https://iam.cf.ac.uk/trac/RAPTOR/wiki/Project/Documentation</a>
Requirements Analysis Report	A report describing the process and results of a requirements gathering exercise that aimed to tease out requirements for usage reporting software requirements. Available at <a href="https://iam.cf.ac.uk/trac/RAPTOR/wiki/Project/Documentation">https://iam.cf.ac.uk/trac/RAPTOR/wiki/Project/Documentation</a>
SAML Accounting Attribute Specifications, SAML Accounting Bindings, SAML Metadata Accounting Extensions.	Working with the SWITCH project AMAAIS, we have identified the need for various standards in terms of accounting to allow inter-operation between the RAPTOR software and the AMAAIS software. These also improve security and usability aspects of the RAPTOR software directly. These particular outputs are the subject of on-going work with the AMAAIS project, beyond the lifetime and scope of RAPTOR.
Dissemination	The project has disseminated at various events, building knowledge in the community of reporting options and helping to establish an initial community of potential RAPTOR users.

Document title: JISC Final Report Template

Last updated : Feb 2011 – v11.0

	<p>These events included:</p> <ul style="list-style-type: none"> <li>• 2010 Gregynog Colloquium (Mid-Wales) <a href="http://whelf.ac.uk/training.shtml">http://whelf.ac.uk/training.shtml</a></li> <li>• TTP EMEA (Regensburg, Germany) <a href="http://www.mindworksuk.com/ttp-emea-10.html">http://www.mindworksuk.com/ttp-emea-10.html</a></li> <li>• FAM10 (Cardiff) <a href="https://sites.google.com/site/jiscfam10/breakout-session-details#1a">https://sites.google.com/site/jiscfam10/breakout-session-details#1a</a></li> <li>• Internet2 2010 Fall Members Meeting (Atlanta, GE) <a href="http://events.internet2.edu/2010/fall-mm/agenda.cfm?go=session&amp;id=10001481&amp;event=1159">http://events.internet2.edu/2010/fall-mm/agenda.cfm?go=session&amp;id=10001481&amp;event=1159</a></li> <li>• Internet2 2011 Spring Members Meeting (Arlington, VA). <a href="http://events.internet2.edu/2011/spring-mm/agenda.cfm?go=session&amp;id=10001676&amp;event=1035">http://events.internet2.edu/2011/spring-mm/agenda.cfm?go=session&amp;id=10001676&amp;event=1035</a></li> <li>• JISC Annual Confence 2011 <a href="http://www.jisc.ac.uk/events/2011/03/jisc11/programme/workshopsmond ay/2statisticgathering.aspx">http://www.jisc.ac.uk/events/2011/03/jisc11/programme/workshopsmond ay/2statisticgathering.aspx</a></li> <li>• TNC 2011 (Prague, Czech Republic) <a href="https://tnc2011.terena.org/core/presentation/56">https://tnc2011.terena.org/core/presentation/56</a></li> </ul>
Collaboration	<p>In the process of developing RAPTOR, the project developed ties with the SWITCH project AMAAIS, JISC project JUSP (JISC Usage Statistics Portal), Shibboleth Consortium (developers of the Shibboleth IdP), Eduserv (makers of the OpenAthens LA product), and the JISC Monitoring Unit (who will host the UK federation RAPTOR aggregator).</p>

## 2.2 How did you go about achieving your outputs / outcomes?

### 2.2.1 Methodology and Work Plan

The first stage of the project investigated both the reporting requirements that potential users of the RAPTOR software might have, and what reporting options would be available to the project.

The first of these – requirements analysis – involved a requirements gathering exercise, conducted through in-depth interviews with library and technical staff at both project partners, and in-breadth surveys of the community conducted through an online survey tool (LimeSurvey). This information gathered formed the basis of a report outlining the conclusions drawn [REQ].

The second – reporting options [OPT] – involved investigating the mechanisms available to retrieve usage data from the relevant authentication systems (Shibboleth, EZproxy, OpenAthens LA). This included looking at existing tools for reporting usage information, along with an investigation into the form of usage information that each of the main authentication systems makes available. Again, the information gathered formed the basis of a report outlining the conclusions drawn.

In the second stage of the project, the project team took these two reports and designed an architecture that would enable software to fulfil the identified requirements; this architecture was used as a basis for the RAPTOR software's construction.

Early on in the project, a similar concurrently running project was identified – the SWITCH funded AMAAIS project (a collaboration of Universität Zürich, ETH Zürich, and SWITCH). The project team established links to AMAAIS; the initial outcome of these links lead to the development of a similar architecture being designed in both projects in order to enable future interoperability between the two software products produced. Later on, this collaboration went one step further and identified a common communications protocol and statistical usage information model as being desirable; work has begun on designing this. The initial release of the RAPTOR software will use its own custom

versions of these items, but the software's design enables the addition of support of this common model and protocol, once designed, in a relatively easy manner.

Throughout all of this work, the project team conducted various amounts of dissemination about the project, aiming to build up community awareness of the project and its software, and to build the beginnings of a community of interested parties. At the beginning of the project this dissemination centred on discussing the project and its aims and objectives, towards the end it centred on demonstrating the prototype software in action to build up community engagement and provoke feedback.

## 2.2.2 Aims and Objectives

The aims and objectives of the project were:

- To understand institutional accounting and reporting requirements around statistics of e-resource access via current common access management technologies (most notably Shibboleth) through a process of requirements gathering within the partner institutions, pilot institutions, and all interested parties.
- To engage with the national bodies such as the JISC and JANET(UK) in order to understand accounting and reporting requirements they face on a federation level, and how institutional-based reporting technologies could help meet these needs.
- To use these dual requirements to inform the functionality of a software toolkit called RAPTOR whose primary aim is to present statistical information about e-resource usage within an institution via Shibboleth (and also potentially EZProxy) to non-technical users of that institution, and whose secondary aim is to make basic aggregated usage information available to suitably authorised external organisations (e.g. a federation operator).
- To implement the RAPTOR software toolkit through a process of a quickly developed initial release and subsequent regular updates during the lifetime of the project.
- To install the RAPTOR software toolkit at a group of around 8-10 pilot institutions, using their feedback regarding both bugs and desired features to inform the regular updates.
- To make the RAPTOR software toolkit available on a free-to-use, open source licensed basis for the UK academic community, and beyond, with full documentation and support structures.

Of these aims and objectives, the project team feels that all were met, with only a few modifications and one exception:

- The exception is that the original plan to develop the software in an agile style and make use of pilot institutions for testing the software was changed with the agreement of the JISC programme manager and RAPTOR governance group; this was because the creation of a secure architecture was more of a prerequisite to building the software than originally envisaged, and because the creation of secure communication between client components came (arguably necessarily) late in the development process – and the project team did not want to release beta, insecure, versions of the software to pilot institutions that would be working with real live, personally identifiable, information.
- The minor modifications from the original plan are that the software's initial release will not just function with the Shibboleth v2 IdP and potentially EZproxy, but definitely functions with the Shibboleth v2 IdP, the Shibboleth v1.3 IdP, EZproxy, and the OpenAthens LA product.
- Additionally, the software supports batch loading of historical log files from the supported systems in order to quickly build a repository of such information (hence the support for the Shibboleth v1.3 IdP).

## 2.3 What did you learn?

Given that the project was primarily a software development project, the main sessions learned centred on the requirements of the users, as outlined in the RAPTOR Requirements Gathering Report [REQ]. A summary of this is that institutions increasingly need usage stats at a time when budgets are decreasing. The people who would make use of these stats mostly want simple statistical analysis (usage per resource, for example), but with the option of doing more complex statistical analysis.

From a project management point of view, the project team learned that when implementing security software that stores and uses real world personally identifiable information, the agile method of software development may not be the most appropriate method to use. Since security software development often involves a lot of time getting the basic architecture correct and implementing things like secure communication, you simply can't release betas of the software intending on working on real live data until these core security features are implemented and considered stable.

Another project management lesson learned was a result of working with other international projects with wildly different timelines to our own. While international collaboration is a fine goal, if the other projects that you are collaborating with are working on different timescales, this interaction has to be managed carefully in order not to impact upon your own timescales and work plans.

## **2.4 Immediate Impact**

The RAPTOR project has had immediate impact on Cardiff University, as the RAPTOR software is installed and running and able to provide statistical information to members of CU staff.

Beyond the project partners, there has not yet been much wider immediate impact, as the software is only due to be released shortly.

## **2.5 Future Impact**

The project is likely to have quite a widespread impact in the UK HE, FE, and schools sectors – and beyond, as the software will enable library and IT staff at these organisations to gain access to statistics vital for decision-making.

It will also impact federations as the RAPTOR software enables them to run a federation-level statistical aggregator that will allow them to demonstrate the value-for-money of their federations.

# **3 Conclusions**

General conclusions:

- Community requirements gathering has shown that statistical usage information is as important to academic institutions as was predicted in the project proposal.

Conclusions relevant to the wider community:

- The RAPTOR software toolkit will be able to help these institutions gain access to the statistical information they need to make value for money decisions.

Conclusions relevant to JISC:

- The JISC Monitoring Unit should set up an aggregator for the UK federation. Project staff are currently working with the JISC MU to achieve this.
- When future projects propose constructing security-related software, the JISC should work with them on the most appropriate method of software development to ensure the method they propose suits the software to be developed.

# **4 Recommendations**

Recommendations for the community

- Institutions that have a need for statistical usage information of their authentication systems can deploy RAPTOR to give them a single point of interaction for this information.

Recommendations for JISC

- The JISC Monitoring Unit should set up an aggregator for the UK federation. Project staff are currently working with the JISC MU to achieve this.

- When future projects propose constructing security-related software, the JISC should work with them on the most appropriate method of software development to ensure the method they propose suits the software to be developed.

## 5 Implications for the future

Implications for the community:

- Institutions that have a need for statistical usage information of their authentication systems can deploy RAPTOR to give them a single point of interaction for this information. This will then allow institutions to make better decisions based on value for money.

Sustainability:

- In the short term, project staff will be continuing development of RAPTOR on a small scale as part of our commitment to the software and the community.
- Longer term, project staff are working with the SWITCH project AMAAIS and members of the Shibboleth Consortium with an aim to integrating both code bases with a view to the Shibboleth Consortium potentially taking ownership of the merged software.

Long term outputs:

- The long-term outputs of the project are software code (Apache2 licensed) and corresponding documentation). Alongside this, there are mailing lists, a web site, and other supporting infrastructure. Each of these will be kept alive and active for at least 3 years, as indicated in the project proposal.

Finally, there are several areas of possible further work:

- The standardisation of accounting attributes, protocols, and entity management that has started with our sister project at SWITCH could be continued and formalised.
- The RAPTOR software itself could be further developed to add additional functionality by polling its initial user community for desired improvements and implementing the most popular, and by implementing all protocols and models standardised with SWITCH.
- RAPTOR information collectors could be more directly integrated into relevant software (such as the Shibboleth IdP), rather than purely working with their log files, to gain access to more detailed usage information and removing a level of abstraction within the architecture.

## 6 References

[OPT] RAPTOR Options Analysis - <http://goo.gl/W2kfR>

[PROP] RAPTOR Proposal to JISC - <http://goo.gl/2XSSB>

[REQ] RAPTOR Requirements - <http://goo.gl/Yks8f>

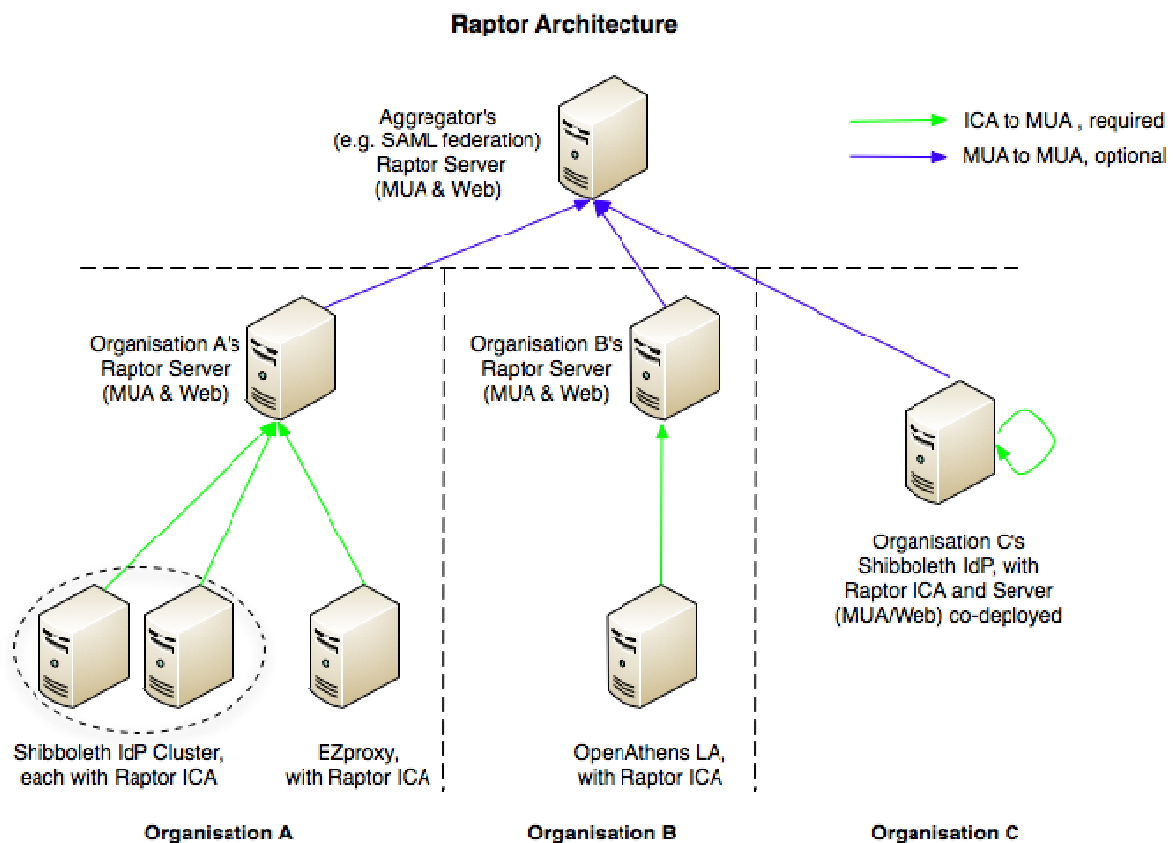
## 7 Appendices

- RAPTOR Architecture

## Appendix 1 – Raptor Architecture

Raptor is designed with a recursive hierarchical architecture:

- Raptor ICAs (Information Collector Agents) sit on the systems to be monitored for events to be accounted for. Once gathered, this event information is sent to a Raptor MUA.
- Raptor MUAs (Multi Unit Aggregators) collect event information sent to them and store it. Raptor MUAs can also send events onwards to another Raptor MUA.
- RaptorWeb is a Web interface that talks to a Raptor MUA, allowing users to view the event information in summary or queryable form.



Raptor ICAs send information to a Raptor MUA. Raptor MUA's, however, can also send information to another Raptor MUA. This allows hierarchical aggregation of event information, enabling accounting across multiple organisations to take place. A typical example of this usage might be a SAML federation operator wishing to gain usage information across the whole federation.