Making Repository Deposit Easier With SWORD

15th October 2009
DSpace User Group Meeting 2009
Göteborgs Universitet, Sweden

Adrian Stevenson
SWORD Project Manager
SWORD Quick Introduction

• Vision: “lowering barriers to deposit”
• Simple Web service Offering Repository Deposit (at least for now!)
• Aims to provide a standard mechanism for ‘doing deposit’ into repositories
• JISC funded project started 2007, SWORD 2 from June 2008
  SWORD3 starting now
What is it?

• A lightweight protocol for deposit
• A profile of the Atom Publishing Protocol
• Implementations of SWORD in IntraLibrary, Fedora, DSpace and Eprints repositories
• SWORD clients – web-based, desktop, Facebook client, MS Office add-on, widgets
Motivations – why?

• no standard interface available for tagging, packaging or authoring tools to upload objects into a repository
• no standard interface for transferring digital objects between repositories
• no way to deposit into more than one repository with one ‘click’
• no way of initiating a deposit workflow from outside a repository system
Use Cases

• Deposit from a Desktop/Online tool
• Multiple deposit - e.g. deposit to institutional and (mandated) funders’ repository with one action
• Machine deposit - e.g. automated deposit from a laboratory machine
• Migration/transfer - e.g. to a preservation service
• Mediated deposit - e.g. deposit by a nominated representative, to additional repositories
The Project Partners

• SWORD partners:
  – UKOLN - Adrian Stevenson (project management)
  – University of Cambridge – Jim Downing (profile)
  – University of Aberystwyth (DSpace, Fedora, & clients) – Stuart Lewis, Neil Taylor, Glen Robson, Richard Jones
  – University of Southampton (EPrints) – Les Carr, Seb Francois
  – Intrallect (IntraLibrary) – Andrew Robson
  – University of York - Julie Allinson
SWORD AtomPub Profile
Standards

- WebDAV (http://www.webdav.org/)
- JSR 170 (http://www.jcp.org/en/jsr/detail?id=170)
- SRW Update (http://www.loc.gov/standards/sru/)
- Flickr Deposit API (http://www.flickr.com/services/api/)
- Fedora Deposit API (http://www.fedora.info/definitions/1/0/api/)
- OKI OSID (http://www.okiproject.org/)
- ECL (http://ecl.iat.sfu.ca/)
“The Atom Publishing Protocol is an application-level protocol for publishing and editing Web resources”

• benefits
  – supports many parameters and requirements - file deposit
  – already exists and has growing support - blogs
  – has an extension mechanism
  – good fit with Web Architecture

• drawbacks / risks
  – retrofit?
  – designed for a single package/file or an atom document – means that we need to package metadata and files
SWORD AtomPub Profile

- SWORD profile builds on AtomPub
- Provides set of extensions, constraint relaxations and enforcements for:
  - Clients posting compound resources (zip,tar)
  - When mediated deposit required
  - Where workflows involved
- Part A adds to AtomPub, Part B highlights how SWORD diverges
- SWORD compliance does not preclude AtomPub compliance
SWORD APP Package Support

- AtomPub uses MIME to describe resources
- Inadequate for compound types e.g.
  - Zip, tar
  - METS, IMS-CP, MPEG21, DIDL packages
- SWORD extends AtomPub:
  - sword:acceptPackaging element
  - Value taken from SWORD package types
SWORD APP Mediated Deposit

- SWORD deposit client user may not be owner of resource
- SWORD allows clients to set a HTTP header:
  - X-On-Behalf-Of
- Assumes trust between owner and mediating user
SWORD APP Developer Features

- No-Op (Dry Run)
- Verbose Output
- Client and Server Identity
- Auto-Discovery
- Error Documents
- Nested Service Description
SWORD APP Error Documents

- SWORD adds new class of doc to AtomPub to allow better error description
  - ErrorContent
  - ErrorChecksumMismatch
  - ErrorBadRequest
  - TargetOwnerUnknown
  - MediationNotAllowed
SWORD Profile of AtomPub

- Part B follows AtomPub spec highlighting where SWORD profile diverges
- Covers:
  - Protocol Operations
    - Retrieving Service Document
    - Listing Collections
    - Creating a Resource
    - Editing a Resource - Not currently implemented
  - Category Documents – MUST NOT be required
  - Service Documents
    - SWORD requires support for service documents
    - new elements: version, verbose, noOp, maxUploadSize
How it Works

- APP and SWORD work by issuing HTTP requests (GET, POST)
  - GET Service Document (explain/discover)
  - POST a file or package to collection URI
- HTTP response and ATOM document is returned
- HTTP basic authentication is required
SWORD In Use
Implementations

- Repository implementations
  - DSpace
  - EPrints
  - IntraLibrary
  - Fedora

- Client implementations
  - command-line, desktop and web clients
  - Facebook Client
  - Java, PHP and .NET libraries
  - Deposit from within MS Word
  - Feedforward / FOREsite and others:
    http://www.swordapp.org/sword/implementation
Web Interface

SWORD Servlet Client - select a service document

Select a service document, or enter another URL:
URL: http://dspace.swordapp.org/sword/servicedocument Other:

Username: a.stevenson@ukoln.ac.uk
Password: ********
On behalf of: 

Get Service Document
Fedora deposit

SWORD Web Client: Service Document

Server details:
- Status: Code: 200, Message: 'Ok'
- URL: http://gen.dnsalias.org/sword/servicedocument
- Version: 1.3
- Supports verbose output: true
- Supports noOp: true
- Maximum file upload size: ~2147483648 KB

Service document:
- Workspace: Fedora SWORD Workspace
  - Collection: Open Collection Deposit
    - Abstract: This is a collection of objects which can be freely deposited to. This is available for the SWORD test project
    - Treatment: Preservation actions may occur on submitted deposits
    - Collection policy: This collection accepts any deposit from anyone
    - Mediation: true
    - Accepts: text/xml, application/zip, application/x-zip-compressed, application/atom+xml, image/gif, image/jpg, image/jpeg, image/png
    - Accepted packaging formats: http://purl.org/net/sword-types/METS/SpaceSIP (0.9), http://www.loc.gov/METS/ (0.9)
  - Collection: Geography Collection Deposit
    - Abstract: This is a nested collection of geography objects
    - Treatment: Preservation actions may occur on submitted deposits
    - Collection policy: This collection accepts any deposit
    - Mediation: true
    - Accepts: application/zip
    - Accepted packaging formats: http://purl.org/net/sword-types/METS/SpaceSIP (0.9)
Fedora Deposit response

SWORD Web Client: Deposit

ID: sword:1008
- Title: Uploaded by the JISC funded SWORD project
- Author: sword
- Summary: null
- Updated: 2009-03-22T16:53:00Z
- Link: (edit) http://glen.dnsalias.org/sword/collection/open/sword:1008
- Treatment: Preservation actions may occur on submitted deposits
- Generator: null
- User Agent: CASIS Test Client
- Packaging: null

Full response

<?xml version="1.0" encoding="UTF-8"?>
<atom:id>sword:1008</atom:id>
<atom:author>
  <atom:name>sword</atom:name>
</atom:author>
<atom:content type="image/jpg" src="http://glen.dnsalias.org:8080/fedora/get/sword:1008/ServletClient-1"/>
<atom:generator url="http://glen.dnsalias.org/sword" version="1.3"/>
<atom:link href="http://glen.dnsalias.org/sword/collection/open/sword:1008" rel="edit" hrelflag="en"/>
<atom:published>2009-03-22T16:53:00Z</atom:published>
<atom:rights type="text"/>
<atom:summary type="text">
  Uploaded by the JISC funded SWORD project
</atom:summary>
<atom:title type="text">Uploaded by the JISC funded SWORD project</atom:title>
<atom:updated>2009-03-22T16:53:00Z</atom:updated>
<sword:treatment>Preservation actions may occur on submitted deposits</sword:treatment>
<sword:noOp>false</sword:noOp>
<sword:userAgent>CASIS Test Client</sword:userAgent>
</atom:entry>
Validation

SWORD Validation Results

These results show a list of any errors, warnings and information messages that are relevant to the specified element. The full element is shown below.

The results represent the hierarchy of elements that are present in the submitted element.

Context

The following context was used when validating the element:

- No Op: false
- Verbose: false
- User Agent: null
- Contributor: null

```
<atom:entry
  sword:packaging
  This element is not present, but it SHOULD be included. If the POST request results in the creation of packaged resource, the server MAY use this element to declare the packaging type. If used it SHOULD take a value from [SWORD-TYPES].

  - atom:id Value: sword:1008
  - atom:author
    - atom:name Value: sword
  - atom:content
    - atom:content type="image/jpeg"
    - atom:content src="http://glen.dnsalias.org:6080/fedora/get/sword:1008/ServletClient-1"
  - atom:generator
    - atom:generator url="http://glen.dnsalias.org/sword"
```
Deposit via Facebook

SWORDAPP – The Repository Deposit Tool

Welcome back to SWORDAPP Adrian!

Deposit an item – Step 1: Select a repository

Select an existing repository: DSpace test server

Or Enter a new one:

Username: a.stevenson@ukoln.ac.uk

Password: ********

Deposit on behalf of:

Next >
SWORDAPP – The Repository Deposit Tool

Welcome back to SWORDAPP Adrian!

Deposit an item – Step 2: Select collection

Server details: Connection OK! (show details)

Please choose from one of the following collections to make your deposit:

- Workspace: DSpace SWORD Demo
  - Collection: Data sets (A collection for depositing data sets)
    - Deposit into this collection
    - See full collection details...
  - Collection: Research materials (A collection for depositing research materials)
    - Deposit into this collection
    - See full collection details...
  - Collection: Teaching materials (A collection for depositing teaching materials)
    - Deposit into this collection
    - See full collection details...
Welcome back to SWORDAPP Adrian!

Deposit an item – Step 3: Item details

- Deposit to: http://dspace.swordapp.org/sword/deposit/123456789/4
- Collection title: Research materials
- Abstract: A collection for depositing research materials

Type of item: Journal article

Has the item been peer reviewed: Yes

Title: Ade in Amsterdam

Abstract: Ade sitting in Amsterdam cafe

1st author first name: Adrian

1st author surname: Stevenson
Welcome back to SWORDAPP Adrian!

Deposit an item – item deposited!

You have successfully deposited 'Ade in Amsterdam' whose URL is now http://dspace.swordapp.org/jspui/handle/123456789/58
Title: Ade in Amsterdam
Authors: Stevenson, Adrian
Issue Date: 22-Mar-2009
Abstract: Ade sitting in Amsterdam cafe

Appears in Collections: Research materials

Files in This Item:

<table>
<thead>
<tr>
<th>File</th>
<th>Description</th>
<th>Size</th>
<th>Format</th>
</tr>
</thead>
<tbody>
<tr>
<td>ade-amsterdam.jpg</td>
<td></td>
<td>18.21 kb</td>
<td>JPEG</td>
</tr>
</tbody>
</table>

Please use this identifier to cite or link to this item: http://hdl.handle.net/123456789/58
Deposit Created!

Title: Attempts to detect retrotransposition and de novo deletion of Alus and other dispersed repeats at specific loci in the human genome

ID: http://dspace.swordapp.org/jspui/handle/123456789/56

Author: Hollies, C.R.

Summary: Dispersed repeat elements contribute to genome instability by de novo insertion and unequal recombination between repeats. To study the dynamics of these processes, we have developed single DNA molecule approaches to detect de novo insertions at a single locus and Alu-mediated deletions at two different loci in human genomic DNA. Validation experiments showed these approaches could detect insertions and deletions at frequencies below 10^-6 per cell. However, bulk analysis of germline
I’m just keeping an eye on this sort of work because I want to ensure that the same functionality is available for my ‘conference chat’ (cchat) system. CChat is RSS-enabled by default in gRSShopper but you have to create a template, which I haven’t done yet. Steve Hargadon, Weblog, March 19, 2009 [Tags: Twitter, Chatrooms, RSS] [Link] [Comment]

Links
OfficeSWORD Add-on

This is a test document
SWORD in use

- More implementations …
  - ICE project is using SWORD
  - EU PEER project implementing SWORD
  - Microsoft Zentity Research-Outputs Repository
  - OAI-ORE - FOREsite work
  - EM-Loader
  - YODL-ING – University of York
  - Others coming along all the time

- Collaboration with publishers

- Any more? Let us know.
SWORD Phase 3

- 11 months, starting now
- Support interest and activities around SWORD
- SWORD package types list/registry
- SWORD enabled repositories registry?
- Formal standardisation?
- SWORDCamp?
- Renewed and increased advocacy efforts
SWORD3 - Development work

• Maintenance and development of SWORD application profile
• Update SWORD demonstrator repositories and clients
• Synergies with project activity in the area
• Tie in with repository handshake strand of international repositories workshop
SWORD 3 - Community Support & Advocacy

• Reflective piece on why SWORD has been a success
• Devise a support model for SWORD
• Increase uptake by marketing and promotion
• Additional use and implementation case studies
SWORD3 - Prototyping Registries

- Prototype SWORD package types registry.
- Prototype ‘SWORD enabled’ repositories registry?
- Explore adding ‘SWORD enabled’ info to existing registries e.g. OpenDOAR
SWORD3 - Standardising SWORD

• Investigate standardising SWORD profile with e.g. NISO, CEN, and others. Suggestions?
• Evaluate cost/benefit and make recommendations.
• Consider alternatives to formal standardisation
SWORD3 - SWORDCamp

• Idea from Monday’s deposit show-and-tell in London
• A week of hands on development
  – agree package standard in advance
    • Based on OAI-ORE
  – update demonstrators
  – update their clients/code
• Timed to coincide with Open Repositories 2010
More Info and Contact

• SWORD Website:
  • http://www.swordapp.org
  • http://twitter.com/swordapp

• General queries:
  – Adrian Stevenson
    a.stevenson@ukoln.ac.uk

• Technical queries:
  – Sword sourceforge list
    sword-app-tech@lists.sourceforge.net