The Cutting Edge of SWORD
18th May 2009
OR09, Atlanta, GA

Adrian Stevenson and Julie Allinson
SWORD Project Managers

UKOLN is supported by: JISC MLA
SWORD Quick Introduction

• Vision: “lowering barriers to deposit”
• Simple Web service Offering Repository Deposit
• Aims to provide a standard mechanism for ‘doing deposit’ into repositories
• JISC funded project started 2007, continuation funding for SWORD 2 from June 2008
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, MS Office plugin, Facebook, widgets
Motivations – why?

- no standard interface 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
The Project Partners

• SWORD partners:
  – UKOLN, University of Bath and University of York (Project Management) – Adrian Stevenson & Julie Allinson
  – University of Aberystwyth (DSpace, Fedora, & clients) – Stuart Lewis, Neil Taylor, Glen Robson, Richard Jones
  – University of Southampton (EPrints) – Les Carr
  – Intrallect (IntraLibrary) – Sarah Currier

• Plus some friendly advisors
  – Jim Downing, Richard Green
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
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 of our parameters and requirements, in particular file deposit
  – it already exists and has growing support
  – it is well-used in popular applications
  – it has an extension mechanism
  – good fit with the Web architecture

• **drawbacks / risks**
  – too much of a retrofit?
  – it is designed for a single package/file OR an atom document – this means that we need to package up metadata and files
SWORD AtomPub Profile

- SWORD profile builds on AtomPub
- Provides set of extensions, constraint relaxations and enforcements when:
  - Clients post compound resources (zip,tar)
  - Mediated deposit required
  - Workflows involved
- 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, SCORM, 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 specification highlighting where SWORD profile diverges

• Part B 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
    • new elements: version, verbose, noOp, maxUploadSize
How it Works

• APP works by issuing HTTP requests (GET, POST)
  – GET Service Document (explain/discover)
  – POST ATOM document or file to collection URI
• HTTP response and ATOM document is returned
• HTTP basic authentication is required
SWORD 2 Profile Updates

• SWORD Profile Version 1.3 includes:
  • Revised deviations from AtomPub and Atom
    – increasing requirement for persistent Atom Entry Documents
  • Includes description of SWORD specific extensions
  • Removed notion of levels of compliance
  • Added sword:userAgent, sword:error, sword:service, sword:version and sword:maxUploadSize elements
SWORD In Use
Implementations

- Repository implementations
  - DSpace
  - EPrints
  - IntraLibrary
  - Fedora

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

SWORD Servlet Client - select a service document

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

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

Get Service Document
Fedora deposit

### Server details:
- **Status**: Code: 200, Message: 'OK'
- **URL**: [http://o2en.dnsalias.org/sword/servicedocument](http://o2en.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
    - **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/jpeg`, `image/jpg`
    - **Accepted packaging formats**: [http://purl.org/net/sword-types/METSDSpaceSIP (0.9)](http://purl.org/net/sword-types/METSDSpaceSIP (0.9)), [http://www.loc.gov/METS/ (0.9)](http://www.loc.gov/METS/ (0.9))
  - **Collection**: Geography Collection
    - **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/METSDSpaceSIP (0.9)](http://purl.org/net/sword-types/METSDSpaceSIP (0.9))
Fedora Deposit response
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

1. 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/jpg"
     - atom:content src="http://glen.dnsalias.org:8080/fedora/get/sword:1008/ServletClient-1"

   - atom:generator
     - atom:generator url="http://glen.dnsalias.org/sword"
Netvibes Widget
Deposit Created!

Title: Attempts to detect retrotansposition 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...
Deposit in Intralibrary
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 >
Welcome back to SWOR.DAPP Adrian!

Deposit an item – Step 2: Select collection

- Workspace: DSpace SWOR.D 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...
SWORDAPP – The Repository Deposit Tool

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 SWORDDAPP 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
URI: http://hdl.handle.net/123456789/58

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>
FeedForward Deposit

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 gRSSShopper but you have to create a template, which I haven't done yet. Steve Hargadon, Weblog, March 19, 2009 [Tags: Twitter, Chatrooms, RSS] [Link] [Comments]

Links
Intralibrary preview of deposited item
OfficeSWORD Add-on

This is a test document.
SWORD in use

• In addition to the case study implementations:
  – Feedforward has already implemented
  – ICE project is looking at SWORD
  – EU PEER project implementing SWORD
  – Microsoft eChemistry work
  – OAI-ORE - FOREsite work
  – EM-Loader
  – YODL-ING – University of York
  – Others coming along all the time

• Collaboration with Nature Publishing Group
More Info and Contact

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

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

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

- SWORD Website
  - http://www.swordapp.org
- Adrian Stevenson, UKOLN
  - a.stevenson@ukoln.ac.uk