ORCA-Registry v2.0 Testing

Document History

James Blanden	26 May 2008	Version 1.0	Initial document.

1. Functionality and Display

1.1. Functionality

1.1.1. ORCA Configuration in orca_init.php

1.1.1.1. Schemata URIs
Sets the URIs for XML data validation and documentation.

1.1.1.2. Administrator E-mail Address
Sets the administrator email address for the Identify response as required by the OAI-PMH specification.
If set to an empty string, the address oai@example.com is used.

1.1.1.3. Harvester Location
Set the location of a harvester that can be used by this instance of ORCA, and the IP address of the harvester as part of the validation of putHarvestData requests.
If set to an empty string, then harvester-based Harvest Methods will not be available when adding or editing a Data Source.

1.1.1.4. Harvest Methods
These settings control the Harvest Methods available when adding or editing a Data Source.

1.1.1.5. Data Provider Types
These settings control the Provider Types available when adding or editing a Data Source.

1.1.1.6. Supported Provider Types for Harvest Methods
These settings control the allowed relations between Harvest Methods and Provider Type when adding or editing a Data Source.

1.1.2. Search

1.1.2.1. Parameters and function.
The search works correctly as described in its associated help.

1.1.2.2. Pagination.
Results show 10 records per page.

1.1.2.3. Click through.
Links in results are correct.

1.1.3. Browse

1.1.3.1. Filter.
The browse filtering works correctly.
Results are returned by filtering on Registry Object class and starts with of simple names and complex name parts.

1.1.3.2. Pagination.
Results show 10 records per page.

1.1.3.3. Click through.
Links in results are rendered correctly.

1.1.4. Web Services

1.1.4.1. Content and links.
Content and links are correct.

1.1.5. Web Services.OAI Data Provider

1.1.5.1. Conformance.
The OAI data provider service validates at version 2.0 according to the protocol conformance tester at http://www.openarchives.org/data/registerasprovider.html.

1.1.6. Web Services.Get Registry Object Groups
1.1.6.1. Result XML.
Returns correct data as valid XML.

1.1.7. Web Services.Get Data Sources
1.1.7.1. Result XML.
Returns correct data as valid XML.

1.1.8. Web Services.Get Registry Objects
1.1.8.1. Filter and Result XML.
Returns correct and valid XML for the parameters passed.

1.1.9. Administration.List Data Sources
1.1.9.1. Filter.
Filters by case-insensitive matching on title.
If only one record is returned, then returns Administration.View Data Source for that record.
1.1.9.2. Pagination.
Shows 25 records per page.
1.1.9.3. Click through.
Clicking a title link returns Administration.View Data Source for that record.

1.1.10. Administration.Add Data Source
1.1.10.1. Form validation.
Mandatory fields are enforced correctly.
Optional fields for OAI Set and Harvest Date are hidden or displayed correctly based on the value selected for Harvest Method.
The relationships between Provider Type and Harvest Method, as configured in orca_init.php are enforced correctly.
1.1.10.2. Form function.
Form values are populated correctly.
1.1.10.3. Save.
Saving creates the record correctly then returns Administration.View Data Source for that record.

1.1.11. Administration.View Data Source
1.1.11.1. Data displays correctly.
1.1.11.2. Refresh.
Clicking the Refresh button returns Administration.View Data Source for this record via an HTTP GET request.
1.1.11.3. Edit.
Clicking the Edit button returns Administration.Edit Data Source for this record.
1.1.11.4. Delete.
Clicking the Delete button returns Administration.Delete Data Source for this record.
1.1.11.5. Clear Log.
Clicking the Clear Log button removes all of the Activity Log entries for this record then places a Log Cleared entry in the Activity Log.
1.1.11.6. Test.
Clicking the Test button either performs a direct retrieval and test of the data from the Data Source URI, or submits a test request to the configured harvester as per the current Harvest Method setting for this Data Source.
Testing validates the data returned, but does not import it.

1.1.11.7. Import.
Clicking the Import button either performs a direct retrieval from the Data Source URI, or submits a request to the configured harvester as per the current Harvest Method, Harvest Date, Harvest Frequency, and OAI-PMH Set settings for this Data Source.

1.1.11.8. Clear.
Clicking the Clear button removes all data gathered from this Data Source from the registry.

1.1.11.9. Plain text Activity Log.
Clicking the text file icon below the Activity Log label returns a plain text file of the Activity Log.

1.1.12. Administration.Edit Data Source

1.1.12.1. Form validation.
Mandatory fields are enforced correctly.
Optional fields for OAI Set and Harvest Date are hidden or displayed correctly based on the value selected for Harvest Method.
The relationships between Provider Type and Harvest Method, as configured in orca_init.php are enforced correctly.

1.1.12.2. Form function.
Form values are populated correctly.

1.1.12.3. Cancel.
The Cancel button returns Administration.View Data Source for this record.

1.1.12.4. Save.
Saving updates the data for this record correctly and returns Administration.View Data Source for this record.

1.1.13. Administration.Delete Data Source

1.1.13.1. Data displays correctly.

1.1.13.2. Cancel.
The Cancel button returns Administration.View Data Source for this record.

1.1.13.3. Delete.
The Delete button removes the record and its associated records correctly and returns Administration.List Data Sources.

1.2. Cross Platform Display and Errors

1.2.1. Firefox 2

1.2.1.1. Search
Renders and functions correctly with no errors.

1.2.1.2. Browse
Renders and functions correctly with no errors.

1.2.1.3. Web Services
Renders correctly with no errors.

1.2.1.4. Administration.List Data Sources
Renders and functions correctly with no errors.

1.2.1.5. Administration.Add Data Source
Renders and functions correctly with no errors.

1.2.1.6. Administration.View Data Source
Renders and functions correctly with no errors.

1.2.1.7. Administration.Edit Data Source
Renders and functions correctly with no errors.

1.2.1.8. Administration.Delete Data Source
Renders and functions correctly with no errors.

1.2.2. Safari 3

1.2.2.1. Search
Renders and functions correctly with no errors.
1.2.2.2. Browse
Renders and functions correctly with no errors.
1.2.2.3. Web Services
Renders correctly with no errors.
1.2.2.4. Administration.List Data Sources
Renders and functions correctly with no errors.
1.2.2.5. Administration.Add Data Source
Renders and functions correctly with no errors.
1.2.2.6. Administration.View Data Source
Renders and functions correctly with no errors.
1.2.2.7. Administration.Edit Data Source
Renders and functions correctly with no errors.
1.2.2.8. Administration.Delete Data Source
Renders and functions correctly with no errors.

1.2.3. IE 7

1.2.3.1. Search
Renders and functions correctly with no errors.
1.2.3.2. Browse
Renders and functions correctly with no errors.
1.2.3.3. Web Services
Renders correctly with no errors.
1.2.3.4. Administration.List Data Sources
Renders and functions correctly with no errors.
1.2.3.5. Administration.Add Data Source
Renders and functions correctly with no errors.
1.2.3.6. Administration.View Data Source
Renders and functions correctly with no errors.
1.2.3.7. Administration.Edit Data Source
Renders and functions correctly with no errors.
1.2.3.8. Administration.Delete Data Source
Renders and functions correctly with no errors.

1.3. XHTML Validation
http://validator.w3.org/

1.3.1. Search
This page is valid XHTML 1.0 strict.
1.3.2. Browse
This page is valid XHTML 1.0 strict.
1.3.3. Web Services
This page is valid XHTML 1.0 strict.
1.3.4. Administration.List Data Sources
This page is valid XHTML 1.0 strict.
1.3.5. Administration.Add Data Source
This page is valid XHTML 1.0 strict.
1.3.6. Administration.View Data Source
This page is valid XHTML 1.0 strict.
1.3.7. Administration.Edit Data Source
This page is valid XHTML 1.0 strict.
1.3.8. Administration.Delete Data Source
This page is valid XHTML 1.0 strict.

1.4. CSS Validation

http://jigsaw.w3.org/css-validator/

1.4.1. Orca.css

This document validates as CSS level 2.1.

1.4.2. Search

This document validates as CSS level 2.1.

1.4.3. Browse

This document validates as CSS level 2.1.

1.4.4. Web Services

This document validates as CSS level 2.1.

1.4.5. Administration.List Data Sources

This document validates as CSS level 2.1.

1.4.6. Administration.Add Data Source

This document validates as CSS level 2.1.

1.4.7. Administration.View Data Source

This document validates as CSS level 2.1.

1.4.8. Administration.Edit Data Source

This document validates as CSS level 2.1.

1.4.9. Administration.Delete Data Source

This document validates as CSS level 2.1.

2. Security

2.1. Authorisations

2.1.1. Functional Roles.

2.1.1.1. Administration.

There are two administrative functional roles for ORCA: ORCA Administrator and ORCA Data Source Administrator.

These roles provide access to the List Data Sources, Add Data Source, View Data Source, and Edit Data Source activities.

2.1.1.2. Add Data Source.

Only members (direct or inherited) of ORCA Administrator have access to the Add Data Source Activity.

2.1.1.3. Edit Data Source Record Owner.

Members (direct or inherited) of ORCA Data Source Administrator cannot change the Record Owner of a Data Source.

2.1.2. Organisational Roles.

2.1.2.1. List Data Sources.

Members (direct or inherited) of ORCA Data Source Administrator only see Data Sources for which the Record Owner is an organisational role of which they are a member (direct or inherited).

2.1.2.2. View Data Source.

Members (direct or inherited) of ORCA Data Source Administrator can only view Data Sources for which the Record Owner is an organisational role of which they are a member (direct or inherited).

2.1.2.3. Edit Data Source.

Members (direct or inherited) of ORCA Data Source Administrator can only edit Data Sources for which the Record Owner is an organisational role of which they are a member (direct or inherited).

2.2. Form data, characters, and query strings

    2.2.1.   Search.
           Validated and handled correctly.

    2.2.2.   Browse.
           Validated and handled correctly.
           Note that addition classes of registry object can be added to the browse results by setting them in the query string. This is then reflected correctly in the user interface, and is not a problem.

    2.2.3.   Administration.List Data Sources
           Validated and handled correctly.

    2.2.4.   Administration.Add Data Source
           Validated and handled correctly.

    2.2.5.   Administration.View Data Source
           Validated and handled correctly.

    2.2.6.   Administration.Edit Data Source
           Validated and handled correctly.

    2.2.7.   Administration.Delete Data Source
           Validated and handled correctly.

2.3. OWASP Top Ten
http://www.owasp.org/index.php/Top_10_2007

    2.3.1.   Cross-site scripting (XSS).
           Output is escaped.
           The output encoding is specified in the HTTP headers, and in the XML declaration, as UTF-8.

    2.3.2.   Injection flaws.
           Parameterised queries to user-defined functions.
           Least privilege for web user operating on a schema other than the default 'public' schema and which is owned by another user.
           Non-verbose error messages.

    2.3.3.   Malicious file execution.
           No file access is determined by user input.

    2.3.4.   Insecure direct object reference.
           All object references are checked.

    2.3.5.   Cross-site request forgery.
           XSS mitigation.
           POST for all data modification.

    2.3.6.   Information leakage and improper error handling.
           Error and debug framework and settings.

    2.3.7.   Broken authentication and session management.
           SSL.
           Hashed storage of built-in credentials.
           Uses the inbuilt session management mechanism.
           Does not accept new, preset or invalid session identifiers.
           No custom cookies for session management.
           Login/logout page ends any existing session.
           Session timeout.
           No exposure of session identifiers or credentials.
           Checks the old password for change of built-in authentication.
           Binds IP address and user agent along with session identifier.

    2.3.8.   Insecure cryptographic storage.
           No requirement for strong cryptographic storage.

    2.3.9.   Insecure communications.
           This is installation/environment specific.

Best practices for SSL communication between application components is advised in the documentation.

2.3.10. Failure to restrict URL access.
Handled by the COSI framework.

3. Source Code
3.1. Style
3.1.1. Change logs present and correct.
3.1.2. License blocks present and correct.
3.1.3. Directory and file naming.
3.1.4. Style standards are followed.
3.1.5. Commenting is extensive and descriptive.

3.2. Function
3.2.1. Input validation.
3.2.2. Output escaping.
3.2.3. Logic and error handling.

4. Load & Performance
4.1. Benchmarking
Benchmarking tests were performed with all software running on a single virtual machine with the following specifications:

- o Linux RHEL 4 x64
- o 1 vCPU (share level - normal)
- o 2048 vRAM (share level - normal)
- o 36GB vHDD (share level - normal)
- o 1 vNIC

This virtual machine was living on a virtual cluster (with around 40 other virtual machines running unknown operating systems and software).

The cluster had a single resource pool shared across 3 physical servers. The cluster also utilised HA (high availability) and DRS (dynamic resource scheduling). If an underlying host failed then HA would automatically bring all VMs online on the remaining hosts in the cluster. DRS automatically moves VMs between hosts depending on the current resource demands of a given VM and available resources on each host.

Each server in the cluster had the following specs:

- o 4x dual-core AMD Opteron64 875 2.2GHz processors (8 cores per server).
- o 16GB 400MHz ECC DDR memory.
- o Dual 2Gb FC SAN paths (redundant/load balanced links).
- o 4x 1gigabit ethernet links.
- o VM data stores live on FC 15K rpm SCSI RAID5 LUNs on an HDS USP-VM storage array.

Optimally, the web server, database, and java application server components would be running on separate hardware (or separate VMs spread across multiple physical machines). Also note that in a deployment situation, the data provider components would be running elsewhere from the harvester.

Times are averages of the times taken to process and dispatch a response to a single request. The OAI-PMH provider code was tested heavily for its capacity to service multiple concurrent requests.

4.1.1. Search execution.
No parameters, 213 from 213 records: 0.824 seconds.
No parameters, 6460 from 6460 records: 0.884 seconds.
No parameters, 12992 from 12992 records: 2.489 seconds.
Single keyword, 4 from 213 records: 0.071 seconds.
Single keyword, 127 from 6460 records: 0.446 seconds.
Single keyword, 256 from 12992 records: 1.370 seconds.

4.1.2. Browse execution.
Collections 'A', 31 from 213 records: 0.033 seconds.
Collections 'A', 992 from 6460 records: 0.322 seconds.
Collections 'A', 1984 from 12992 records: 1.189 seconds.
Services 'D', 6 from 213 records: 0.024 seconds.
Services 'D', 86 from 6460 records: 0.197 seconds.
Services 'D', 192 from 12992 records: 0.686 seconds.

4.1.3. OAI-PMH Data Provider.
The OAI-PMH data provider uses the same data retrieval code as the search, and so the time to service an OAI-PMH ListRecords request should scale in a similar way to the search times. The quantity of the data returned will be greater though, and so the time to receive the complete response will be greater.
The OAI-PMH ListRecords and ListIdentifiers utilise resumption tokens, and will only return data for up to 100 records.

4.1.4. Get Registry Object Groups.
Returned in less than a second with over 6000 records.

4.1.5. Get Data Sources.
Returned in less than a second with over 6000 records.

4.1.6. Get Registry Objects.
This service uses the same data retrieval code as the search, and so the time to service requests should scale in a similar way to the search times.
Note that this service will not limit the amount of data returned, and so the time taken to deliver the results over the network could be substantial.

4.1.7. Put Harvest Data.
Put 100 records into the registry: 6.418 seconds.

4.1.8. Clear records from a Data Source.
Clear 203 records from a Data Source: 0.955 seconds.
Clear 3045 records from a Data Source: 11.72 seconds.
Clear 6293 records from a Data Source: 25.10 seconds.

5. Accessibility Guidelines

5.1. Web Content Accessibility Guidelines 1.0

5.1.1. Guideline 1. Provide equivalent alternatives to auditory and visual content.
5.1.2. Guideline 2. Don't rely on colour alone.
5.1.3. Guideline 3. Use markup and style sheets and do so properly.
5.1.4. Guideline 4. Clarify natural language usage.
5.1.5. Guideline 5. Create tables that transform gracefully.
5.1.6. Guideline 6. Ensure that pages featuring new technologies transform gracefully.
5.1.7. Guideline 7. Ensure user control of time-sensitive content changes.
5.1.8. Guideline 8. Ensure direct accessibility of embedded user interfaces.
5.1.9. Guideline 9. Design for device-independence.
5.1.10. Guideline 10. Use interim solutions.
5.1.11. Guideline 11. Use W3C technologies and guidelines.

5.1.12. Guideline 12. Provide context and orientation information.
5.1.13. Guideline 13. Provide clear navigation mechanisms.
5.1.14. Guideline 14. Ensure that documents are clear and simple.

5.2. Checklist of Checkpoints for Web Content Accessibility Guidelines 1.0
   http://www.w3.org/TR/WCAG/full-checklist.html
   5.2.1. Priority 1 Checkpoints (MUST)
   5.2.1.1. | YES | ~~NO | N/A~~ | Provide a text equivalent for every non-text element.
   5.2.1.2. | YES | ~~NO | N/A~~ | Ensure that all information conveyed with colour is also available without colour.
   5.2.1.3. | ~~YES | NO~~ | N/A | Clearly identify changes in the natural language of a document's text and any text equivalents.
   5.2.1.4. | YES | ~~NO | N/A~~ | Organise documents so they may be read without style sheets.
   5.2.1.5. | ~~YES | NO~~ | N/A | Ensure that equivalents for dynamic content are updated when the dynamic content changes.
   5.2.1.6. | YES | ~~NO | N/A~~ | Avoid causing the screen to flicker.
   5.2.1.7. | ~~YES | NO~~ | N/A | Use the clearest and simplest language appropriate for a site's content.
   5.2.1.8. | YES | ~~NO | N/A~~ | For data tables, identify row and column headers.
   5.2.1.9. | ~~YES | NO~~ | N/A | For data tables that have two or more logical levels of row or column headers, use markup to associate data cells and header cells.
   5.2.1.10. | YES | ~~NO | N/A~~ | Ensure that pages are usable when scripts are turned off or not supported.

   5.2.2. Priority 2 Checkpoints (SHOULD)
   5.2.2.1. | YES | ~~NO | N/A~~ | Ensure that foreground and background colour combinations provide sufficient contrast when viewed by someone having colour deficits or when viewed on a black and white screen.
   5.2.2.2. | YES | ~~NO | N/A~~ | When appropriate markup language exists, use markup rather than images to convey information.
   5.2.2.3. | YES | ~~NO | N/A~~ | Create documents that validate to published formal grammars.
   5.2.2.4. | YES | ~~NO | N/A~~ | Use style sheets to control layout and presentation.
   5.2.2.5. | YES | ~~NO | N/A~~ | Use relative rather than absolute units in markup language attribute values and style sheet property values.
   5.2.2.6. | ~~YES | NO~~ | N/A | Use header elements to convey document structure and use them according to specification.
   5.2.2.7. | YES | ~~NO | N/A~~ | Mark up lists and list items properly.
   5.2.2.8. | YES | ~~NO | N/A~~ | Mark up quotations. Do not use quotation markup for formatting effects such as indentation.
   5.2.2.9. | YES | ~~NO | N/A~~ | Ensure that dynamic content is accessible or provide an alternative presentation or page.
   5.2.2.10. | YES | ~~NO | N/A~~ | Avoid causing content to blink.
   5.2.2.11. | YES | ~~NO | N/A~~ | Do not create periodically auto-refreshing pages.
   5.2.2.12. | YES | ~~NO | N/A~~ | Do not use markup to redirect pages automatically (use the server to perform redirects).
   5.2.2.13. | YES | ~~NO | N/A~~ | Do not cause pop-ups or other windows to appear and do not change the current window without informing the user.
   5.2.2.14. | YES | ~~NO | N/A~~ | Use W3C technologies when they are available and appropriate for a task and use the latest versions when supported.
   5.2.2.15. | YES | ~~NO | N/A~~ | Avoid deprecated features of W3C technologies.
   5.2.2.16. | YES | ~~NO | N/A~~ | Divide large blocks of information into more manageable groups where natural and appropriate.
   5.2.2.17. | YES | ~~NO | N/A~~ | Clearly identify the target of each link (i.e. not "click here").

5.2.2.18. | ~~YES | NO~~ | N/A | Provide metadata to add semantic information to pages and sites.

5.2.2.19. | YES | ~~NO | N/A~~ | Provide information about the general layout of a site (eg a site map or table of contents).

5.2.2.20. | YES | ~~NO | N/A~~ | Use navigation mechanisms in a consistent manner.

5.2.2.21. | YES | ~~NO | N/A~~ | Do not use tables for layout unless the table makes sense when linearised. Otherwise, if the table does not make sense, provide an alternative equivalent (which may be a linearised version).

5.2.2.22. | YES | ~~NO | N/A~~ | If a table is used for layout, do not use and structural markup for the purpose of visual formatting (eg do not use TH to make something bold and centred).

5.2.2.23. | YES | ~~NO | N/A~~ | For all form controls with implicitly associated labels, ensure the label is properly positioned

5.2.2.24. | ~~YES | NO~~ | N/A | For scripts, ensure that event handlers are input device-independent.

5.2.2.25. | YES | ~~NO | N/A~~ | Avoid movement in pages, or provide a means for turning it off.

5.2.2.26. | ~~YES | NO~~ | N/A | Make scripts directly accessible or compatible with assistive technologies.

5.2.2.27. | ~~YES | NO~~ | N/A | Ensure that any element that has its own interface can be operated in a device-independent manner.

5.2.2.28. | ~~YES | NO~~ | N/A | For scripts, specify logical event handlers rather than device-dependent event handlers.

5.2.3. Priority 3 Checkpoints (MAY)

5.2.3.1. | ~~YES | NO~~ | N/A | Specify the expansion of each abbreviation or acronym in a document where it first occurs.

5.2.3.2. | YES | ~~NO | N/A~~ | Identify the primary natural language of a document.

5.2.3.3. | ~~YES | NO~~ | N/A | Specify the expansion of each abbreviation or acronym in a document where it first occurs.

5.2.3.4. | YES | ~~NO | N/A~~ | Identify the primary natural language of a document.

5.2.3.5. | ~~YES~~ | NO | ~~N/A~~ | Create a logical tab order through links, form controls, and objects.

5.2.3.6. | YES | NO | N/A | Provide keyboard shortcuts to important links.

5.2.3.7. | YES | ~~NO | N/A~~ | Include non-link printable characters (surrounded by spaces) between adjacent links.

5.2.3.8. | ~~YES | NO~~ | N/A | Provide information so that users may receive documents according to their preferences (eg language, content type, etc).

5.2.3.9. | YES | ~~NO | N/A~~ | Provide navigation bars to highlight and give access to the navigation mechanism.

5.2.3.10. | YES | ~~NO | N/A~~ | Group related links, identify the group (for user agents), and provide a way to bypass the group.

5.2.3.11. | ~~YES | NO~~ | N/A | If search functions are provided, enable different types of searches for different skill levels and preferences.

5.2.3.12. | ~~YES | NO~~ | N/A | Place distinguishing information at the beginning of headings, paragraphs, lists, etc.

5.2.3.13. | ~~YES | NO~~ | N/A | Provide a means to skip over multi-line ASCII art.

5.2.3.14. | ~~YES | NO~~ | N/A | Supplement text with graphic or auditory presentations where they will facilitate comprehension of the page.

5.2.3.15. | YES | ~~NO | N/A~~ | Create a style of presentation that is consistent across pages.

5.2.3.16. | YES | ~~NO | N/A~~ | Provide summaries for tables.

5.2.3.17. | ~~YES~~ | NO | ~~N/A~~ | Provide abbreviations for header labels.

5.2.3.18. | ~~YES | NO~~ | N/A | Provide a linear alternative to side-by-side text for all tables that lay out text in parallel, word-wrapped columns.