

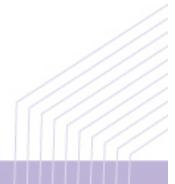
**If God had meant us to understand
file formats, He would have told us
AONS ago**

Matthew Walker
AONS Technical Lead
National Library of Australia
mwalker@nla.gov.au





AONS headquarters ☺





AONS II: Automatic Obsolescence Notification System - version 2

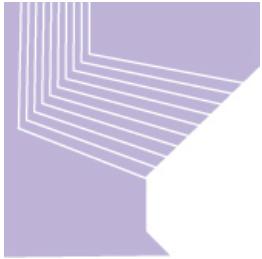
Project Manager: David Pearson

Technical Lead: Matthew Walker

Software Developer: David Levy

AONS II – latest beta release (2007-10-23) 23 Oct 2007.
Represents seven months' worth of software development.





Project Objective (APSR Nov 2006)

Objective:

To refine the Automatic Obsolescence Notification System (AONS) developed in an earlier stage of APSR, to a platform-independent downloadable tool that automatically provides information from authoritative international registries to support decisions on preservation action required to retain access to information resources stored in repositories.





Why AONS II?

Format obsolescence is potentially a major problem for every repository manager. This is particularly true given:

- ▶ Ever-increasing volume of digital material
- ▶ Plethora of file formats
- ▶ Dynamic nature of computing environments
- ▶ Rapid and unpredictable drivers that cause formats to become obsolete
- ▶ High business value of the specific content of some digital materials or collections can result in policies that mandate that access be maintained to this data for extended periods of time



Repository managers need help to manage the quantity and diversity of file formats and their obsolescence risks.





High-level Software Design Principles

- Support three different business environments
 - ▶ a national federated infrastructure
 - ▶ enterprise business models
 - ▶ individual standalone repository sites
- Open source & use Java code
- Modular & Reusable/Adaptable
- Platform independent
- Interoperable, using common interfaces, protocols and standards
- RESTful Service-Oriented Architecture
- Loosely-coupled configurable design
- Demonstrable





Deployment Modes

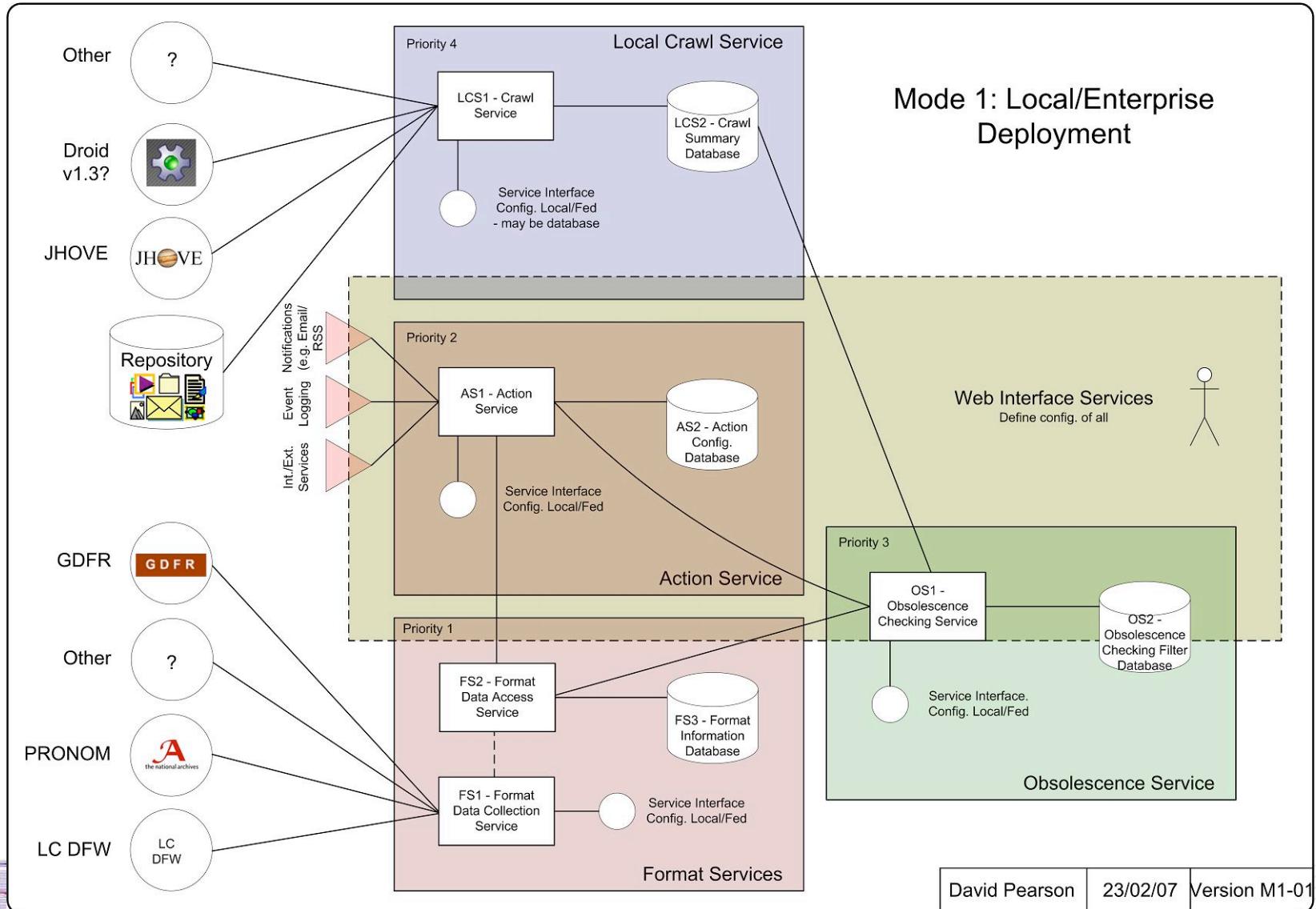
AONS II Deployment Mode and Common/ Uncommon Modules

Mode 1: Local/Enterprise Deployment

Mode 2: Federated Deployment

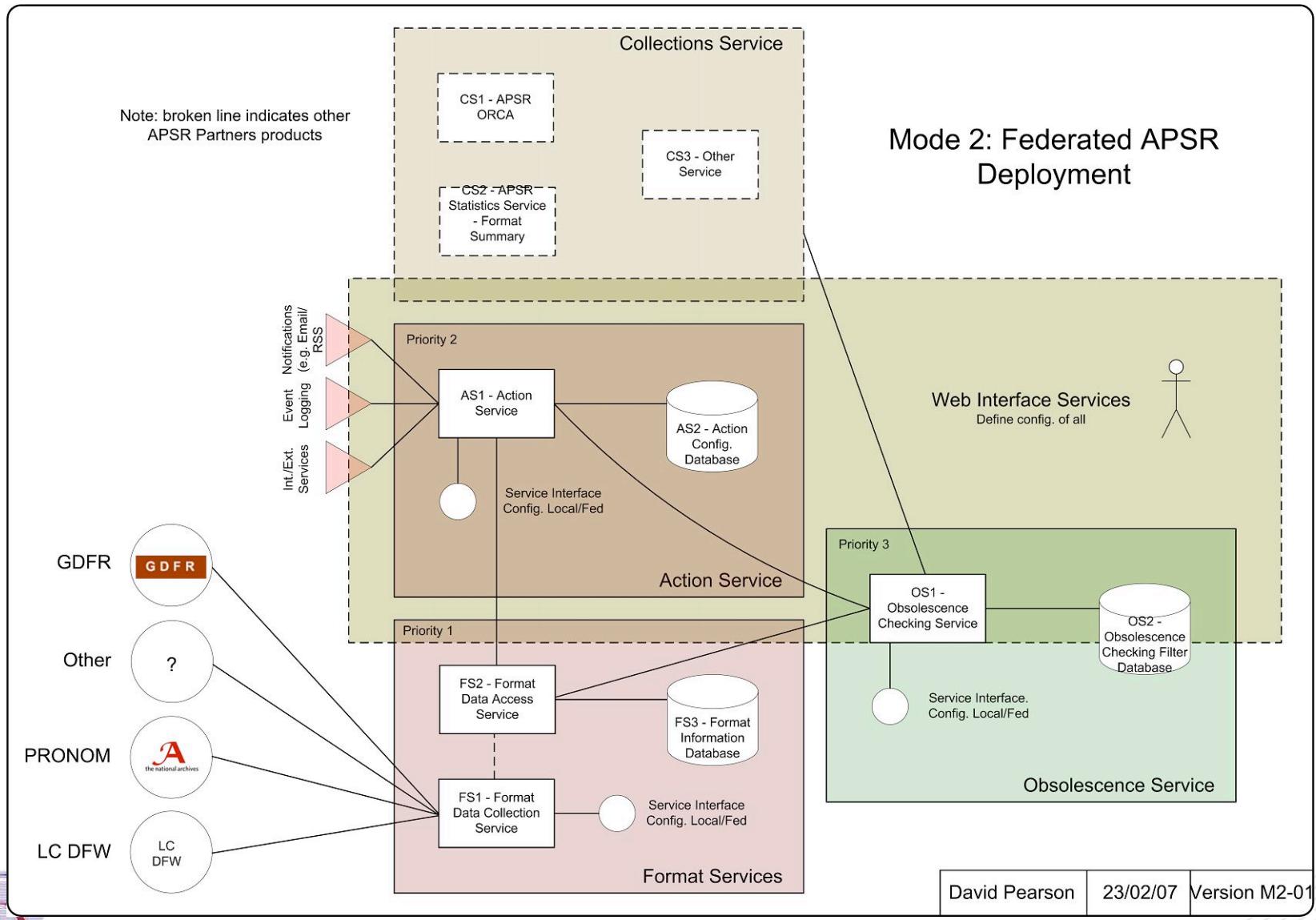
NLA Common Modules Required	Format Service	Action Service	Obsolescence Service	Format Service	Action Service	Obsolescence Service
NLA Uncommon Modules	Local Crawl Service	Web Interface Service				
APSR Modules				APSR - ORCA	APSR - Stats Service Format Summary	APSR - Web Interface Service
2007 Out of Scope Modules	Security			Security		

Mode 1: Local/Enterprise deployment



David Pearson | 23/02/07 | Version M1-01

Mode 2: Federated APSR deployment





AONS II in Action

- Check files for format obsolescence
 - ▶ On-Ingest
 - ▶ On-Demand
 - adhoc or scheduled monitoring





Recognising File Formats and Building Collection Profiles

- Collection profile construction
 - ▶ Compliant XML metadata summary
 - ▶ Repository crawl using adapters purpose-built for a type of repository
- Crawl result processing
 - ▶ Repository-supplied metadata (e.g. from DSpace)
 - ▶ Automated format recognition tools (such as DROID, JHOVE or other plug-in tool).



Repository List

The screenshot shows a Mozilla Firefox browser window displaying the APSR Collection Services interface at <https://pilot.apsr.edu.au/cosif/aons/repository/repositorieslist.php>. The page title is "Collection Services".

The left sidebar contains a navigation menu:

- Logout
- Change Built-in Passphrase
- Format Obsolescence
 - Repositories
 - Registries
 - Formats Search
 - Risk Summary
- Collections Registry
 - Search
 - Browse
 - Web Services

The main content area displays a "Repositories List" table:

Repository Name	Type	Last Run	Action	Action	Action	Action
ANU DSpace (dev)	ExternalUrl	2007-09-03 12:07	details	update	delete	schedule
ANU DSpace (prod)	ExternalUrl	2007-10-01 12:00	details	update	delete	schedule
UNSW Fedora	ExternalUrl	2007-10-03 13:40	details	update	delete	schedule
USyd eScholarship (prod)	ExternalUrl	2007-09-04 11:00	details	update	delete	schedule
USyd eScholarship (dev)	ExternalUrl	2007-09-04 10:59	details	update	delete	schedule

Below the table is a "Register New Repository" form:

* Repository Type:

At the bottom of the page, there is copyright information: "Copyright © Australian Partnership for Sustainable Repositories (APSR) Partners 2007 | About The Applications | Feedback".



Registry List

APS Collection Services - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
<https://pilot.apsr.edu.au/cos/ions/registry/registrylist.php>

Latest Headlines OurWeb National Library of A...

OurWeb APSR Collection Services

Expand All Menus Collapse All Menus Skip to Menus Skip to Content About APSR Collection Services

APSR AUSTRALIAN PARTNERSHIP FOR SUSTAINABLE REPOSITORIES

Collection Services

Logout Change Built-in Passphrase

Format Obsolescence

Repositories Registries Formats Search Risk Summary

Collections Registry

Search Browse Web Services

Registries List

Registry Name	Status	Last Run	Format Count	Action
Library of Congress	Idle	2007-10-01 02:03	194	details
National Archives (UK)	Idle	2007-09-08 02:04	587	details

Logged in as: David Pearson (dpearson)

Current Registry Utilisation

National Archives (UK) [17] Library of Congress [9]

Registry Utilisation Over Time

Formats Utilised

Change Date

National Archives (UK) — Library of Congress

Copyright © Australian Partnership for Sustainable Repositories (APSR) Partners 2007 | [About The Applications](#) | [Feedback](#)

Done

Risk Summary

APSR Collection Services - Mozilla Firefox
File Edit View Go Bookmarks Tools Help
<https://plot.apsr.edu.au/cos/aons/risk/risksummary.php>
Latest Headlines OurWeb National Library of A...
OurWeb APSR Collection Services

Expand All Menus Collapse All Menus Skip to Menus Skip to Content About APSR Collection Services

Logged in as: David Pearson (dpearson)

APSR Collection Services

Risk Summary

Format Name	Type	Status	Quantity	Local Risk	Community Risk	Final Risk	Multiplicative Risk
Binary Interchange File Format (BIFF) Workbook 8X	AONS	Risk Assessment Performed	6	0.0	0.0	0.0	0.0
Exchangeable Image File Format (Compressed) 2.2	AONS	Risk Assessment Performed	17361	10.0	9.0	9.5	164929.5
Exchangeable Image File Format (Uncompressed) 2.2	AONS	Risk Assessment Performed	143	10.0	10.0	10.0	1430.0
Extensible Markup Language 1.0	AONS	Risk Assessment Performed	5002	0.0	0.0	0.0	0.0
Graphics Interchange Format 1987a	AONS	Risk Assessment Performed	89	0.0	3.0	1.5	133.5
Graphics Interchange Format 1989a	AONS	Risk Assessment Performed	761	0.0	0.0	0.0	0.0
JPEG File Interchange Format 1.00	AONS	Risk Assessment Performed	107	4.0	2.0	3.0	321.0
JPEG File Interchange Format 1.01	AONS	Risk Assessment Performed	16734	4.0	6.0	5.0	83670.0
JPEG File Interchange Format 1.02	AONS	Risk Assessment Performed	1704	1.0	0.0	0.5	852.0
Microsoft Word for Windows Document 97-2003	AONS	Risk Assessment Performed	25	5.0	8.0	6.5	162.5
Portable Document Format 1.1	AONS	Risk Assessment Performed	126	6.0	7.0	6.5	819.0
Portable Document Format 1.3	AONS	Risk Assessment Performed	4172	4.0	6.0	5.0	20860.0
Portable Document Format 1.4	AONS	Risk Assessment Performed	2786	3.0	5.0	4.0	11144.0
Portable Document Format 1.5	AONS	Risk Assessment Performed	573	1.0	3.0	2.0	1146.0
Portable Document Format 1.6	AONS	Risk Assessment Performed	744	0.0	1.0	0.5	372.0
Tagged Image File Format 6	AONS	Risk Assessment Performed	11369	0.0	0.0	0.0	0.0
ZIP Format	AONS	Risk Assessment Performed	30	4.0	7.0	5.5	165.0
name: [Exchangeable Image File Format (Compressed)], version: [2.1], puid: [x-fmt/390]	Unidentified	Unidentified	6153	10.0	10.0	10.0	61530.0
name: [Exchangeable Image File Format (Uncompressed)], version: [2.1], puid: [x-fmt/388]	Unidentified	Unidentified	19	10.0	10.0	10.0	190.0
name: [Extensible Hypertext Markup Language], version:	Unidentified	Unidentified	46	10.0	10.0	10.0	460.0



Format Identification Issues

APSR Collection Services - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

https://pilot.apsr.edu.au/cos/aons/risk/risksummary.php

Latest Headlines OurWeb National Library of A...

Portable Document Format 1.1	AONS	RISK ASSESSMENT PERFORMED	126	5.0	7.0	6.5	819.0
Portable Document Format 1.3	AONS	Risk Assessment Performed	4172	4.0	6.0	5.0	20860.0
Portable Document Format 1.4	AONS	Risk Assessment Performed	2786	3.0	5.0	4.0	11144.0
Portable Document Format 1.5	AONS	Risk Assessment Performed	573	1.0	3.0	2.0	1146.0
Portable Document Format 1.6	AONS	Risk Assessment Performed	744	0.0	1.0	0.5	372.0
Tagged Image File Format 6	AONS	Risk Assessment Performed	11369	0.0	0.0	0.0	0.0
ZIP Format	AONS	Risk Assessment Performed	30	4.0	7.0	5.5	165.0
name: [Exchangeable Image File Format (Compressed)], version: [2.1], puid: [x-fmt/390]	Unidentified	UnIdentified	6153	10.0	10.0	10.0	61530.0
name: [Exchangeable Image File Format (Uncompressed)], version: [2.1], puid: [x-fmt/388]	Unidentified	UnIdentified	19	10.0	10.0	10.0	190.0
name: [Extensible Hypertext Markup Language], version: [1.0], puid: [fmt/102]	Unidentified	UnIdentified	46	10.0	10.0	10.0	460.0
name: [Hypertext Markup Language], version: [3.2], puid: [fmt/98]	Unidentified	UnIdentified	24	10.0	10.0	10.0	240.0
name: [Hypertext Markup Language], version: [4.01], puid: [fmt/100]	Unidentified	UnIdentified	1369	10.0	10.0	10.0	13690.0
name: [Hypertext Markup Language], version: [4.0], puid: [fmt/99]	Unidentified	UnIdentified	136	10.0	10.0	10.0	1360.0
name: [Hypertext Markup Language], version: [], puid: [fmt/96]	Unidentified	UnIdentified	1772	10.0	10.0	10.0	17720.0
name: [MPEG-2 Video Format], version: [], puid: [x-fmt/386]	Unidentified	UnIdentified	1	10.0	10.0	10.0	10.0
name: [MS-DOS Executable], version: [], puid: [x-fmt/409]	Unidentified	UnIdentified	3	10.0	10.0	10.0	30.0
name: [Microsoft Powerpoint Presentation], version: [97-2002], puid: [fmt/126]	Unidentified	UnIdentified	17	10.0	10.0	10.0	170.0
name: [Microsoft Word for Windows Document], version: [97-2002], puid: [fmt/40]	Unidentified	UnIdentified	1	10.0	10.0	10.0	10.0
name: [OLE2 Compound Document Format], version: [], puid: [fmt/111]	Unidentified	UnIdentified	31	10.0	10.0	10.0	310.0
name: [OpenDocument Text Format], version: [1.0], puid: [fmt/136]	Unidentified	UnIdentified	2	10.0	10.0	10.0	20.0
name: [OpenOffice Impress], version: [1.0], puid: [fmt/130]	Unidentified	UnIdentified	1	10.0	10.0	10.0	10.0
name: [OpenOffice Writer], version: [1.0], puid: [fmt/128]	Unidentified	UnIdentified	2	10.0	10.0	10.0	20.0
name: [Portable Document Format], version: [1.0], puid: [fmt/14]	Unidentified	UnIdentified	3	10.0	10.0	10.0	30.0
name: [Portable Document Format], version: [1.2], puid: [fmt/16]	Unidentified	UnIdentified	1965	10.0	10.0	10.0	19650.0
name: [Portable Network Graphics], version: [1.0], puid: [fmt/11]	Unidentified	UnIdentified	73	10.0	10.0	10.0	730.0

Format Identification/Linkage

The screenshot shows a Mozilla Firefox browser window with the title bar "AONS II - Mozilla Firefox". The address bar contains the URL http://localhost:8080/aons/gui/aons2Workflow.html?_flowExecutionKey=_c2ACDBE6B-2077-BDCC-1BE1-2065C22CA. The page itself has a dark header with the "AONS" logo on the left and a search bar on the right. Below the header, the main content area is titled "Choose Identification Method". It displays the extension "[js]" and a list of three items under the heading "Extension: [js]":

- [Search Internal](#): Search and link to an existing internal format
- [Search External](#): Search for one or more external formats to use as a basis for a new format
- [Create New](#): Utilise found identification metadata to create a new format

To the right of the main content area, there is a sidebar with a vertical list of links:

- [Tasks](#)
- [Risk Summary](#)
- [Reports](#)
- [Registries](#)
- [Repositories](#)
- [Tools](#)
- [Configuration](#)
- [Notification](#)
- [Logging](#)
- [About](#)

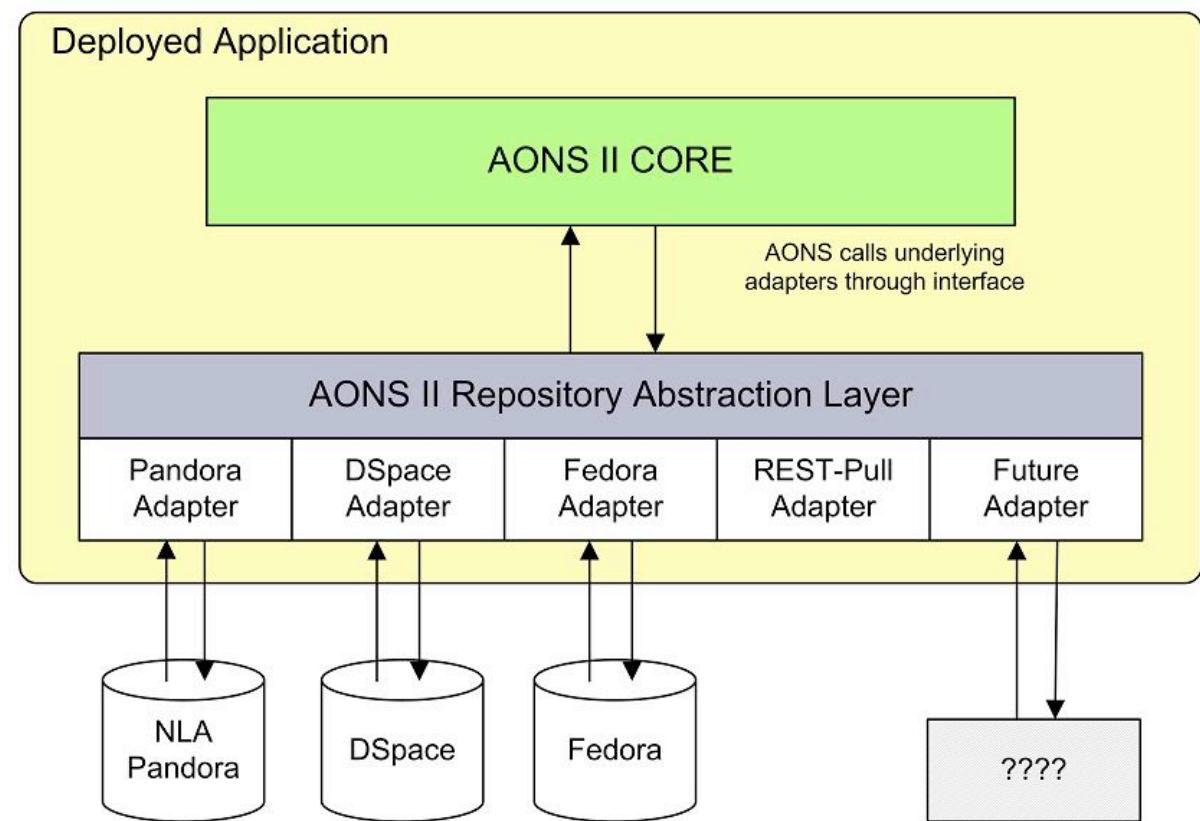
AONS 2 Beta Release [20071003], built on 2007/10/03 07:18

© Copyright 2004-2007, <http://www.nla.gov.au>, under the terms of the Apache 2.0 software license.



Repository Adapters

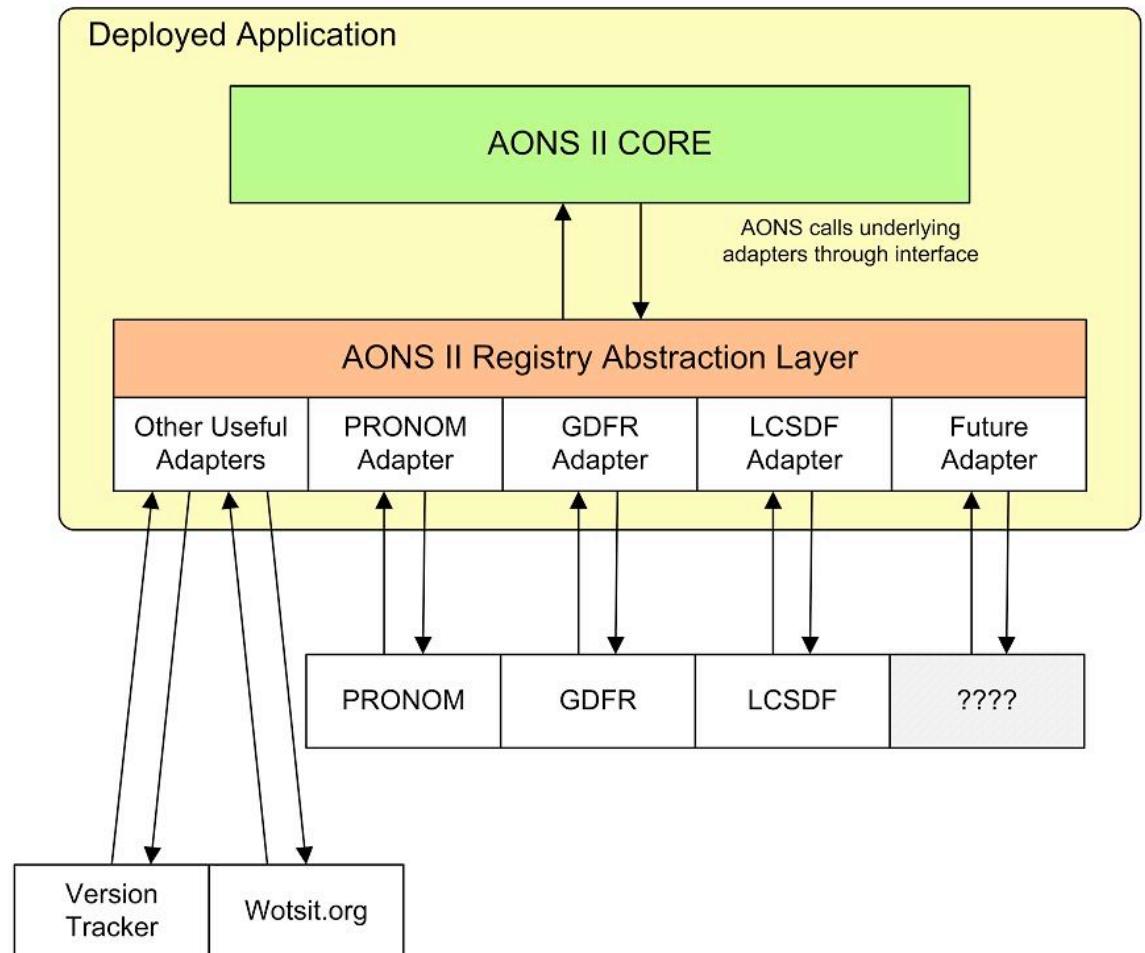
- Completed:
 - ▶ DSpace (1.4)
 - ▶ Fedora (2.2)
 - ▶ NLA Pandora
 - ▶ REST/Pull
 - ▶ Generic File System
- Not completed 2007
 - ▶ EPrints
 - ▶ SRB (UniQLD)
 - ▶ TRIM





Registry Adapters

- Completed
 - ▶ LoC SDF
 - ▶ PRONOM
- Not completed
 - ▶ GDFR



Task List

The screenshot shows a Mozilla Firefox browser window with the title bar "AONS II - Mozilla Firefox". The address bar contains the URL "http://localhost:8080/aons/gui/aons2Workflow.html?_flowExecutionKey=_cD9AEE154-3646-48D4-FAF0-805D075C2". The main content area displays the AONS II interface. At the top left is the AONS logo (a red stylized flower or sunburst). The top right features a large orange "XO" logo. Below the logo is a search bar with the placeholder "Search". On the left, a sidebar titled "Tasks" includes a link "Show Completed Tasks". The main content area is titled "Identify Format Metadata ([View All](#))". It lists three tasks, all marked as "Outstanding" and created on "2007/10/04 09:08". Each task has a yellow star icon. The details for each task are as follows:

Read	Date Created	Status	Cause	Details	Available Actions
*	2007/10/04 09:08	Outstanding	Format identification metadata [extension: [gif]] is not linked to an AONS format	Details	Identify Format Metadata
*	2007/10/04 09:08	Outstanding	Format identification metadata [extension: [js]] is not linked to an AONS format	Details	Identify Format Metadata
*	2007/10/04 09:08	Outstanding	Format identification metadata [extension: [msg]] is not linked to an AONS format	Details	Identify Format Metadata

A vertical sidebar on the right lists navigation links: "Tasks", "Risk Summary", "Reports", "Registries", "Repositories", "Tools", "Configuration", "Notification", "Logging", and "About".

RDSS - Risk Questions

Step 1 – Risk assessment based on Community Information	Step 2 - Risk assessment based on Collection / Repository Information
Community Information Questions	Collection / Repository Information Questions
Q1. Is this a base format? A ubiquitous format which is likely to be rendered by most applications (e.g. plain text). If unknown, state "Unknown".	Q1. The original primary rendering software has been identified as ... (see Step 1 - Q6.) Is this primary rendering software available to you? Q2. How many new versions have been released since then? Q6. Is the original primary rendering software for this format version identified? What is it? If unknown, state "Unknown".

Q1. The original primary rendering software has been identified as ... (see Step 1 - Q6.)

Is this primary rendering software available to you?

Q1.
by n
If ye

If unknown, state "Unknown".

Q2. How many new versions have been released since then?

Q6. Do you have any other alternative means of providing safe and effective access? (i.e. custom designed applications, scripts, emulators). What are they?

Q6. Is the original primary rendering software for this format version identified?
What is it?
If unknown, state "Unknown".

Q6. Overall, how many access options are effectively available to you (i.e. how many can you make work), including the original rendering software?

Q6. Is the original primary rendering software for this format version identified?

What is it?

If unknown, state "Unknown".

Q9. For each alternative, are there critical hardware and software dependencies for effective use of the alternative rendering software?
What are they?
If unknown, state "Unknown".

Format Risk Assessment

AONS II - Mozilla Firefox

File Edit View Go Bookmarks Tools Help

http://localhost:8080/aons/gui/aons2Workflow.html?_flowExecutionKey=_c968A6150-385C-362B-714C-B0ED825EC/

[Help] Q2. Is this file format and version referenced in any searched information resources?

[Help] Q3. Is there a known support end date for this format version?

If yes, how many years to that support end date?

[Help] Q4. How many years since this version was released?

[Help] Q5. How many new versions have been released since then?

[Help] Q6. Is the original primary rendering software for this format version identified?

Done

Risk Summary

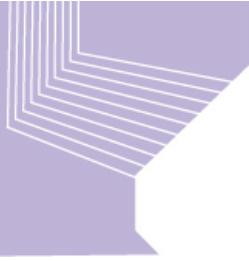
The screenshot shows a Firefox browser window displaying the 'Global Risk Summary' page from the AONS II application. The URL in the address bar is http://localhost:8080/aons/gui/aons2Workflow.html?_flowExecutionKey=_c3A46F959-7696-0140-7261-1D797211F1.

The main content area displays the 'Global Risk Summary as of 2007/10/04 11:08'. It lists various file formats with their risk assessment status, quantity, and calculated risks.

Name	Type	Status	Quantity	Community Risk	Local Risk	Final Risk	Multiplicative Risk	Review Date
AutoCAD Drawing 2004-2005	Format	Risk Assessment Performed	1	10.0	5.0	7.5	0.0011316484	2007/10/04 11:00
Binary Interchange File Format SX	Format	No Risk Assessment	14	10.0	0.0	10.0	0.021124104	No Record
Broadcast WAVE 1	Format	Risk Assessment Performed	1	5.0	-5.0	0.0	0.0	2007/10/04 11:07
Cascading Style Sheet	Format	No Risk Assessment	2	10.0	0.0	10.0	0.0030177291	No Record
Extensible Markup Language 1.0	Format	Risk Assessment Performed	2	-5.0	-5.0	-5.0	-0.0015088646	2007/10/04 11:08
extension: [bak]	Format	UnIdentified	3	10.0	10.0	10.0	0.0045265937	Not ...

The sidebar on the right contains navigation links:

- Tasks
- Risk Summary
- Reports
- Registries
- Repositories
- Tools
- Configuration
- Notification
- Logging
- About

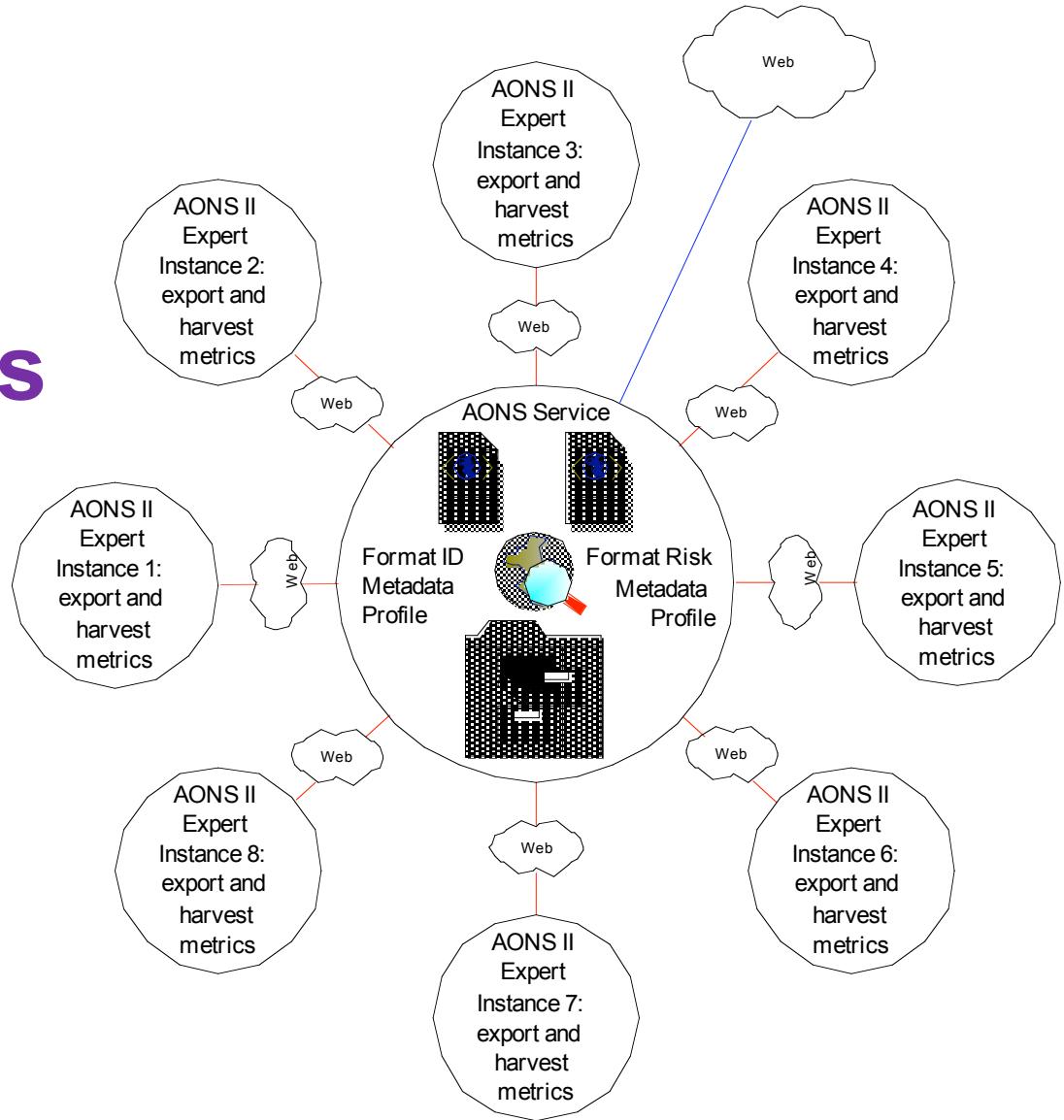


Demo

- <http://localhost:8080/aons>



Possible Futures



AONS II Service (Web)

Providing metrics for:

- Format ID Metadata Profile; and
- Format Risk Metadata Profile.

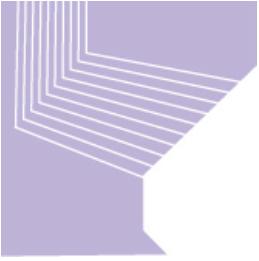
Information push/pulled from-to other AONS II expert instances. Information available to all.



AONS Ten Commandments

- I: Thou shalt not worship proprietary file formats
- II: Thou shalt not discard files of wayward format
- III: Remember not to analyse file formats manually on the weekend
- IV: Honour “base” file formats
- V: Thou shalt not kill digital file contents
- VI: Thou shalt not commit to a risky file format
- VIII: Thou shalt not falsely assess a file format’s risk
- IX: Thou shalt not covet other format rendering infrastructure
- X: Thou shalt not covet supported file formats





Acknowledgements & Links



- APSR:
 - ▶ <http://www.apsr.edu.au/>
- APSR AONS II Home Page:
 - ▶ <http://www.apsr.edu.au/aons2/index.htm>
- AONS II Wiki:
 - ▶ <http://www.apsr.edu.au/wiki/index.php/AONS>
- AONS II Development Blog:
 - ▶ <http://aons2dev.blogspot.com>
- Download AONS:
 - ▶ <http://sourceforge.net/projects/aons/>

