NEXT GENERATION PRESERVATION MANAGEMENT

A DIGITAL REPOSITORY FOR MUSEUM COLLECTIONS

BEN FINO-RADIN | MoMA
DAN GILLEAN | ARTEFACTUAL SYSTEMS
KARA VAN MALSSSEN | AVPRESERVE

DIGITAL PRESERVATION 2014 | WASHINGTON, DC
DIGITAL ART STORAGE
DEVELOP
the
BUSINESS CASE
TOTAL GB ACQUIRED ANNUALLY

<table>
<thead>
<tr>
<th>Year</th>
<th>Total GB Acquired Annually</th>
</tr>
</thead>
<tbody>
<tr>
<td>1975</td>
<td>0</td>
</tr>
<tr>
<td>1984</td>
<td>4,000</td>
</tr>
<tr>
<td>1993</td>
<td>8,000</td>
</tr>
<tr>
<td>2002</td>
<td>12,000</td>
</tr>
<tr>
<td>2011</td>
<td>19,700</td>
</tr>
</tbody>
</table>
INVOLVE all STAKEHOLDERS
5 DEPARTMENTS
5 DEPARTMENTS
13 INTERVIEWEES
DEFINE

system

requirements

& use cases
FUNCTIONAL REQUIREMENTS
FUNCTIONAL REQUIREMENTS

USE CASES
55 FUNCTIONAL REQUIREMENTS

22 USE CASES

- SIP CREATION AND INGEST
- INTEGRITY CHECKING
- DESCRIPTION
- SEARCH & BROWSE
- ADDING VERSIONS
- REPORTING
- RISK ASSESSMENT
- SYSTEMS INTEGRATION
- ADMINISTRATION
CONSULT

the

EXPERTS
EXTERNAL ADVISORS
6 EXTERNAL ADVISORS

HOWARD BESSER, NYU MIAP
HANNAH FROST, STANFORD UNIVERSITY
STEPHEN ABRAMS, CALIFORNIA DIGITAL LIBRARY
JEROME MACDONOUGH, UNIVERSITY OF ILLINOIS
DAVID MILLMAN, NYU DLTS
BEN FINO-RADIN, RHIZOME at the NEW MUSEUM
EXTERNAL ADVISORS

HOWARD BESSER, NYU MIAP
HANNAH FROST, STANFORD UNIVERSITY
STEPHEN ABRAMS, CALIFORNIA DIGITAL LIBRARY
JEROME MACDONOUGH, UNIVERSITY OF ILLINOIS
DAVID MILLMAN, NYU DLTS
BEN FINO-RADIN, RHIZOME at the NEW MUSEUM
TMS

DATABASE

APPLICATION

STORAGE

DAM
TMS

STORAGE

ARCHIVEMATICA

MANAGEMENT/ACCESS LAYER

DAM
REQUEST for PROPOSALS
SCOPE

USE CASES

ECOSYSTEM

TECHNICAL REQUIREMENTS

FUNCTIONAL REQUIREMENTS

USE CASES
RFP ISSUED
DISTRIBUTION of the REQUEST FOR PROPOSALS
Vetting the VENDORS
THE BIG INSTITUTION and the VERY SMALL VENDOR
KNOW

your

VENDOR
DRMC TEAM

DAVID JUHASZ
Project Manager

JESÚS GARCÍA CRESPO
System Architect

DAN GILLEAN
Domain Analysis, Design

MIKE CANTELON
Developer

MISTY DE MEO
Developer

JOSÉ RADDAOUI MARÍN
Developer

HEATHER ANDERSON
Developer
Providing access to memory since 2007

AtoM stands for Access to Memory. It is a web-based, open source application for standards-based archival description and access in a multilingual, multi-repository environment.

Download
AtoM 2.0.1

Web-based
Access your AtoM installation from anywhere you have an internet

Open source
All AtoM code is released under a GNU Affero General Public License (A-GPL)

Standards-based
AtoM was originally built with support from the International Council on

Import/export friendly
Your data will never be locked into AtoM – we implement a number of
Phase 2: Code sprint #5

Due in 16 days (05/31/2014)

13 issues (1 closed - 12 open)

24%

Related issues

Feature #6283: AIP record
Feature #6374: Context browser (relationships)
Feature #6377: Navigation
Feature #6428: Add multiple URI access points (slug, TMS access)
Feature #6429: Add breadcrumbs
Feature #6463: Archivematica user authentication
Feature #6588: Allow upload of DIPs to NetX
Feature #6589: Store AIPs uncompressed
Feature #6595: Integrity checker reporting
Feature #6596: Integrity checker integration
Feature #6597: Integrity checker app
Feature #6598: Reports
Feature #6622: Context browser (usability enhancements)
EXPECT to be INVOLVED
1. "Artwork storage by codec" is referencing MIME type for all files, not codec of AV files. We do want this for format, but we also need the codec widget with the same "count" and "size" tabs.
2. "Total artworks in repository" as seen on slide 1 of the PPT - dashboard is missing the line graphs for year ingested and year collected.
3. Monthly ingest totals looks unfinished – reminder, we need this for codec and format - and we want to see the monthly totals, as well as running totals (slide 6)
4. Artwork sizes – median is missing
5. the widgets on slide 9 – the departmentally focused widgets – of the ppt are missing

**Updated by David Juhasz 5 days ago**

- Status changed from QA/Review to Feedback
- Assignee changed from Ben Fino-Radin to Mike Cantelon

**Updated by Mike Cantelon 5 days ago**

Ben Fino-Radin wrote:

1. "Artwork storage by codec" is referencing MIME type for all files, not codec of AV files. We do want this for format, but we also need the codec widget with the same "count" and "size" tabs.
2. "Total artworks in repository" as seen on slide 1 of the PPT - dashboard is missing the line graphs for year ingested and year collected.
3. Heather and I worked on this yesterday, fixing the underlying issues. For the year ingested, maybe year/month would be more suitable initially (until you ran the DRMC for a few years)?
4. Artwork sizes – median is missing
5. I'm not sure if we can do median because of a limitation with ElasticSearch. The limitation is being worked on, though, so a future version of ES might make it easy to add this on.

**Updated by David Juhasz 5 days ago**

- Assignee changed from Mike Cantelon to Ben Fino-Radin

**Updated by Ben Fino-Radin 5 days ago**

- Assignee changed from Ben Fino-Radin to Mike Cantelon

The widgets have nothing to do with DIP processing. For all of these widgets we want to be reporting the AIP not the DIR. This information is all available in the METS.

Yes – year-month is best for such cases.

If you can facet by department in search/browse, then why can't you leverage this metadata for the widget?

Not doing median is fine. All the other problems you listed seem most given the above points.

**Updated by David Juhasz 5 days ago**

Mike, it looks like "Department" is pulled from TMS here:


**Updated by David Juhasz 5 days ago**

I sent an email to Radda asking him to assist Mike with getting the AV Codec data.

**Updated by Ben Fino-Radin 5 days ago**

Great - thanks David!

**Updated by José Raddaoui Marin 2 days ago**

Hi Mike,

The Mediainfo data from the METS file is now included in the ES index. I did the import in dev/issue-6519-mediainfo-import, it’s in one commit if you want to cherry-pick it to your branch, commit: 14769ed. You'll need to clear the cache and populate the search index.

I've grouped the tracks by type, so you can query over "generalTracks", "videoTracks" or "audioTracks". You can reach this information in:

Artworks/Components/Technology Records -> aips.digitalObjects.metaData.mediainfo
FROM USER STORIES...

<table>
<thead>
<tr>
<th>Step</th>
<th>User Action</th>
<th>System Response</th>
<th>UI elements needed</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Conservator navigates to artwork record</td>
<td>System displays artwork’s main image, core metadata from TMS, listing of Artwork AIPs and their relationship to the artwork (instantiation, version of instantiation, supplementary materials for instantiation), and documentation AIPs</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>Conservator selects documentation AIP</td>
<td>System lists contents of selected documentation AIP</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>Conservator clicks an individual file in the documentation AIP</td>
<td>System loads and renders the selected file in the web browser, eliminating the need to download and view locally.</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>Conservator has gained knowledge to inform conservation treatment.</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

MoMA DRMC-MA Workflow 9

MoMA DRMC-MA Workflow 9 Depicted in 4 wireframes

START
- Conserver logs into DRMC-MA Dashboard
- Conserver searches for artwork
- Conserver selects documentation AIP
- Conserver selects individual AIP file

System loads and renders the selected file in the web browser,

Wireframe 9-4: render file

END
FROM UX FLOWCHARTS...

MoMA DRMC-MA Workflow 9
Depicted in 4 wireframes

START
Conservator logs into DRMC-MA Dashboard
Conservator searches for artwork
DRMC displays artwork main image, metadata, list of AIPs by class (artwork, software, documentation etc) and AIP relationships
Conservator selects individual AIP file
System lists contents of selected documentation AIP
Wireframe 9-3: list AIP contents
Conservator selects documentation AIP
System loads and renders the selected file in the web browser,
Wireframe 9-4: render file

END

TO WIREFRAMES

MoMA DRMC-MA Workflow 9
Depicted in 4 wireframes

START
Conservator logs into DRMC-MA Dashboard
Conservator searches for artwork
DRMC displays artwork main image, metadata, list of AIPs by class (artwork, software, documentation etc) and AIP relationships
Conservator selects individual AIP file
System lists contents of selected documentation AIP
Conservator selects documentation AIP
System loads and renders the selected file in the web browser,
Wireframe 9-4: render file

END
FROM WIREFRAMES... TO PROTOTYPES
Phase 1A
Analysis, workflow, mockups

Phase 1B
Analysis and prototyping

Phase 2
Agile development and more user testing

Phase 3
Deployment, optimization

2013

2014

Seb Chan
Mike Giarlo
Martina Haidvogl
Anna Henry
Mark Hellar
Hanna Hölling
Agathe Jarczyk
Pip Laurenson
Jeff Martin
Mark Matienzo
Joanna Phillips
Dave Rice
Gwynne Ryan
Mike Smorul
Kara Van Malssen
Farris Wahbeh

ADVISORY MEETINGS
PREPARE for CHALLENGES
Bumps Ahead
Self-Playing Violin
Archival Information Package #2

object:
_1922007ax1SAPM.wav-c3c5fcbe-d522-4ef0-bd72-...

manifestation:
Original

UUID: c3c5fcbe-d522-4ef0-bd72-0ae3928758ff
sha256: d763aa8e017998a1e39f2785c2ee212a03dde43392bea2de08467df79a9ef8c
size: 191.90 MB
format: Waveform Audio
puid: fmt/6

numSamples: 268291200
sampleRate: 9600
audioDataEncoding: PCM
channels: 2
bitDepth: 24
offset: 12288

status: Well-Formed and valid
mimeType: audio/x-wave
Boomerang (1974), Richard Serra

TMS metadata

- **Accession number**: F2289
- **Object ID**: 143808
- **Title**: Boomerang
- **Year**: 1974
- **Artist**: Richard Serra
- **Classification**: Video
- **Department**: Film
- **Medium**: Video (color, sound)
- **Dimensions**: 10 min.

Context browser

Dublin Core metadata

Files (24 items)

- Pass2_legaltitlecard-7a32904...: 9.47 GB
- tagmanifest-md5.txt: 142 bytes
- tagmanifest-md5.txt: 142 bytes
- manifest-sha256.txt: 108 bytes
Demo time!
FY14 ASSESSMENT RESULTS

it’s not just about technology!
DRMC SPECIFICATIONS

ISO 16363:2012
audit & certification of trustworthy digital repositories

NDSA LEVELS OF PRESERVATION
DRMC SPECIFICATIONS
<table>
<thead>
<tr>
<th>NDSA LEVELS OF PRESERVATION</th>
<th>Level 1: Protect your data</th>
<th>Level 2: Know your data</th>
<th>Level 3: Monitor your data</th>
<th>Level 4: Repair your data</th>
</tr>
</thead>
<tbody>
<tr>
<td>STORAGE &amp; GEOGRAPHIC LOCATION</td>
<td>Green</td>
<td>Yellow</td>
<td>Red</td>
<td>Red</td>
</tr>
<tr>
<td>FILE FIXITY &amp; DATA INTEGRITY</td>
<td>Green</td>
<td>Yellow</td>
<td>Green</td>
<td>Yellow</td>
</tr>
<tr>
<td>INFORMATION SECURITY</td>
<td>Green</td>
<td>Red</td>
<td>Yellow</td>
<td>Red</td>
</tr>
<tr>
<td>METADATA</td>
<td>Green</td>
<td>Yellow</td>
<td>Yellow</td>
<td>Green</td>
</tr>
<tr>
<td>FILE FORMATS</td>
<td>Green</td>
<td>Yellow</td>
<td>Yellow</td>
<td>Green</td>
</tr>
</tbody>
</table>
ISO 16363
In conclusion