

Project Acronym: Presto4U
Grant Agreement no: 600845
**Project Title: European Technology for Digital
Audiovisual Media Preservation**

Report accompanying D4.6: Software Tools Catalogue

Project funded by the European Community in the 7th Framework Programme



Table of contents

Scope	3
Executive summary	4
1. Audiovisual Communities and the use of software tools	5
1.1 Pre-existing software tools initiatives	5
2 Emerging Research Outputs	6
2.1 Incorporating PrestoKAT into the Tools Catalogue	6
3 Interdependency of Standards and Software	8
4 Functional Specifications	9
4.1 Merging PrestoKAT and PrestoCentre Tools	9
4.2 Creating a persistent link between the Tools Catalogue Standards Register	10
4.3 Tools Catalogue Landing Page	13
4.4 Tools Catalogue Record Page	14
4.4.1 Description Fields	14
4.4.2 Sustainability Factors	15
4.4.3 Research outputs	18
4.5 Discovery Mechanisms for the Tools Catalogue	19
4.5.1 Facets as an integration point with Research Outcomes	19
4.5.2 Facets as an integration point with the Standards Register	20
4.5.3 Implementation of Facets in Tools Catalogue	21
4.6 Technical Implementation	23
5. Conclusion	25
References	26
Document information	28

Scope

The report accompanies the launch of D4.6 Software Tools Catalogue. The report details the methodology, technical build, and rollout of the Tools Catalogue for the Presto4U project, part of a suite of interlocking services geared towards understanding the needs of distributed and varied communities within audiovisual preservation. The Tools Catalogue is largely based on the draft methodology represented in D4.5 “Interim Report on Software Tools Catalogue.” In addition the close linkage between the outcomes of D3.2 and D3.3 on Research Outcomes Assessments represent a means of identifying tools most relevant to the needs and skillsets of those involved in audiovisual preservation. The work takes into account pre-existing and parallel initiatives around the dissemination of information relating to audiovisual preservation, highlighting the new avenues to the discovery of tools, and the essential need to centre activities relating to software within a suite of standards and services.

The Tools Catalogue can be found at <https://www.prestocentre.org/library/tools-catalogue>.

Executive summary

This Report represents the methodology and build of a Software Tools Catalogue, as part of a suite of project outcomes and services within the Presto4U project. This Catalogue, in conjunction with the Standards Register, forms the pillars of the Community Marketplace. The idea is that the Marketplace will become something like a forum where users in the audiovisual archiving and preservation community can investigate and interact with tools and services to fulfil their preservation needs and requirements. It could become a place where vendors or technology providers can assess current challenges in digital preservation, and respond by offering services and solutions.

It is based around the idea of interlocking types of organisations who are seeking to find tools and services to solve their digital preservation needs. During the first year Presto4U has identified these communities as:

- Film Collections and Filmmakers
- Footage Sales Libraries
- Learning and Teaching Repositories
- Music and Sound Archives
- Personal Audiovisual Collections
- Research and Scientific Collections
- TV, Radio and New Media Broadcasting
- Video Art, Art Museums and Galleries
- Video Production and Post-Production

By focusing the Software Tools Catalogue around these Communities of Practice, we are acknowledging that while there are common challenges across all communities in digital audiovisual preservation, differing communities have different needs.

Development of this Catalogue has involved research into both previous and current initiatives that have included software tools as a component of their project or institutional outcomes. By acknowledging past initiatives, we can learn lessons on structure, impact and sustainability. By working with current initiatives in the field of digital preservation we can acknowledge our points of commonality and our different focus in order to be interoperable across initiatives and to foster an open exchange of ideas.

The Tools Catalogue can be found at <https://www.prestocentre.org/library/tools-catalogue>.

This document is structured as follows: Chapter 1 is a general overview of the current use of tools for digital audiovisual preservation and the rationale for the Software Tools Catalogue. Chapter 2 focuses on the need to assess emerging research outcomes, to remain flexible and forward thinking in terms of the type of tools that are being developed to solve future preservation issues. Chapter 3 details the symbiosis between standards and software to create reliable and durable tools. Chapter 4 presents the functional requirements and build of the Tools Catalogue, with a particular focus on discoverability and flexibility as a means of creating multiple paths to the understanding of tools. Finally the Conclusion looks to the future of the Tool Catalogue and its place within the Community Marketplace.

1. Audiovisual Communities and the use of software tools

The Tools Catalogue responds to the perceived need to create a registry of tools focused on the workflows within the different types of communities engaged in digital audiovisual preservation. It acknowledges the wide range of technical skillsets within organisations and necessitates the development of different points of entry to satisfy the needs of those different audiences. Communities of Practice are at the heart of the Tools Catalogue in terms of being the audience for certain tools whether they have identified this or not, and as an indication of actual uptake and adoption of as a means to identifying trends of use.

1.1 *Pre-existing software tools initiatives*

The Tools Catalogue takes into account previous work within this sector and their relevance to audiovisual preservation needs, as referenced in D4.5 “Interim Report of Software Tools Catalogue.”[1] Information has been scattered across several initiatives, often without audiovisual preservation at its core interests. By building the software Tools Catalogue within PrestoCentre we are drawing upon many years of excellence in audiovisual preservation knowledge transfer. Central to this is acknowledgement of previous PrestoCentre investigations into, and representation of, tools. By assessing and incorporating that work into the new Tools Catalogue, we were able to build and expand upon the strong foundation that previous initiatives have created.

The Tools Catalogue, as with the Standards Register [2], is based around community adoption of resources related to audiovisual preservation as a driver to uptake of resources. Being open to collaboration with related projects is a central philosophy. The project has already illustrated the strength of collaborative output through the relationship between Presto4U Standards and the FP7 APARSEN initiative on standards. [3] By creating an agreement to share ideas on schemas and vocabularies we have created interoperability opportunities for data and a common approach to field labelling within user interfaces. To extend the reach of the register, the project will continue to be open to similar collaborations with other tools initiatives such as COPTR [4] and the Library of Congress Tools Showcase [5], as well as being open to other collaborations that will best create robust interchange of information between resources. At the most basic this may manifest itself as web links between different initiatives, through to more fully interoperable collaborations on schemas and vocabularies.

2 Emerging Research Outputs

To highlight the rapidly evolving environment of tools seen as relevant to audiovisual preservation workflows, the parallel Presto4U project track “Identification and Analysis of Research Outcomes” is identifying and analysing research outputs that can potentially match the Community of Practice audiovisual preservation needs. This is the opportunity for communities to become aware of emerging research that may solve their issues at a later date and to potentially leverage the outcomes of those research outcomes by making their needs known at the very beginning of the development process.

The Presto4U Research Outputs Assessment v1 report [6], published in December 2013, presented the methodologies by which tools were chosen for testing, and the outcomes of those tests performed during the first year. The project team continues to work with the Communities of Practice in year two to further identify emerging tools. The Tools Catalogue will include these emerging tools and monitor community adoption in the same way as it will monitor the adoption of more mature tools that are currently available.

2.1 Incorporating PrestoKAT into the Tools Catalogue

The Emerging Research Outcomes project team created an internal-only web application called PrestoKAT for the WP3 team members to track and assess emerging tools and software related to audiovisual preservation. With a focus on the metrics needed to investigate and analyse the basic use and maturity of emerging tools, this interface is tracking the research lifecycle using a Technology Readiness Level (TRL) scale [7]. As an assessment methodology it presents a clear indication of the development lifecycle. To users it shows a transparent indication of the readiness of a tool, and an opportunity to decide at what part of the development lifecycle they would be comfortable investing resources into. A picture of TRL where level 3 is selected is shown in Figure 1.

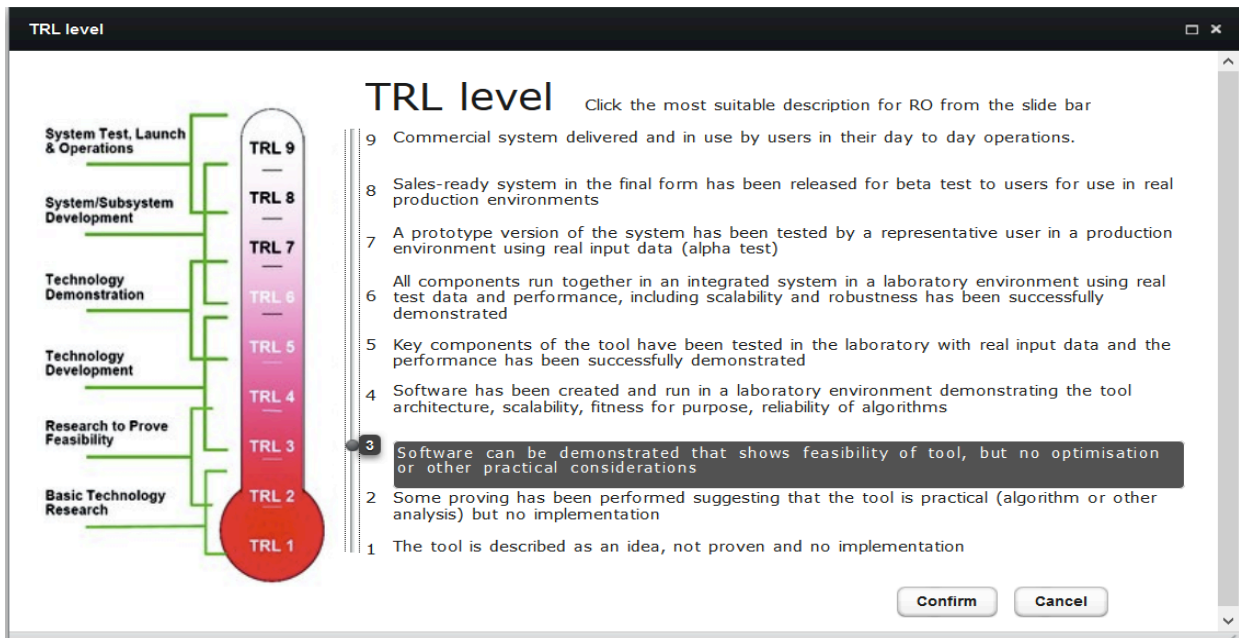


Figure 1. TRL level in PrestoKAT user interface [8].

Many of the other fields follow existing schemas and standards, such as OAIS functional specification model[9], as well as the Knowledge Schema developed as part of Presto4U [10]. Figure 2 below shows the PrestoKAT user interface.

The screenshot shows the 'Add a new RO' form in the PrestoKAT user interface. The form is organized into several sections:

- Details:** This section contains the primary input fields. It starts with a 'Name' field followed by a 'Check Name' button. Below this are 'Start Year' and '~ Finish Year' fields, both with date pickers (yyyy-MM-dd). The 'Early Research Output' field is a dropdown menu. The 'Commercial type' field is also a dropdown menu. The 'TRL level' field has a text input and a 'Get my RO's TRL level' button. The 'Software Type' field is a dropdown menu with 'Add' and 'Remove' buttons below it.
- Metadata:** A large, empty text area for entering additional information.
- Tool Type:** A dropdown menu with 'Type support Tool type' selected.
- Platform:** A dropdown menu with 'Type support platform' selected.
- Required Software:** A text input field with the placeholder text 'example: jdk'.

Figure 2. PrestoKAT user interface

By having mapped some of the PrestoKAT fields and data to the publicly available Tools Catalogue, we can now present users with both emerging and mature tools, facilitating a broader range of choices available to differing types of organisations involved in audiovisual preservation.

3 Interdependency of Standards and Software

Standards provide stability and reliability in product development, and the ability for tools, and services to be interoperable. By adhering to the principles of standardisation in terms of the rules by which a tool is created, originators are ensuring a level of consistency in the products they develop. In terms of community acceptance, the evidence of a tool being standards-based generates trustworthiness in the product. Migration and future development should be easier in a standards based environment due to the stability of the original framework.

To recognise the importance of standards to tool development and sustainability, the project launched the Standards Register [2], a catalogue of standards that are relevant to audiovisual preservation. As well as focusing on those standards that are important to audiovisual preservation, the Standards Register illustrates actual adoption of particular standards by differing communities. Together the Standards Register and PrestoKAT form the basis of the Tools Catalogue schema, to create a user interface that is compliant with the structure of the Standards Register, while incorporating fields that are particular to research development. Drawing on these two forms of field structures facilitates a more uniform view and navigation experience for users.

To reinforce the relationship between standards and tools, we have created transparent and strong links between the two. By creating such persistent links between tools and standards, we are reinforcing the use of standards as a benchmark to the trustworthiness of tools. These links should be at both a general level where users can easily link to the Standards Register to browse information contained therein, and also link back to specific standards as used by the tools they are investigating.

4 Functional Specifications

The build of the Tools Catalogue is based on current and past trends in Tools Catalogues relating to digital audiovisual preservation. It also marks the convergence of current project outcomes on Research Outputs and the Standards Register.

The Tools Catalogue presents a list of processes, tools and techniques used at different stages of the audiovisual preservation lifecycle. Specifically, the Software Tools Catalogue will have the potential to:

- Provide a list of tools that are both commercial and open source
- Provide an aggregated catalogue of tools that are currently available for use in digital preservation
- Provide a list of current emerging research outputs for use in digital preservation, with special attention to multimedia contents
- As a community driven initiative, the Software Tools Catalogue will allow users to add new records for software outputs, or update existing records within the Software Tools Catalogue
- Monitor and approve all user-driven data submission for authenticity, traceability and suitability
- Create links to code repositories such as SourceForge [11], GoogleCode [12] and Github [13]
- Create links to guidelines or manuals for self-installation or deeper review of tools
- Provide evaluation/feedback mechanism for users (and Community of Practice) to rate software
- Provide links to vendor/research outputs demonstrations of tools and services
- Provide links to brokerage activities within marketplace

Reassessing the potential for the catalogue to list not only software, but also different types of research outcomes, the project has taken the step to rename the Software Tools Catalogue as the Tools Catalogue.

4.1 *Merging PrestoKAT and PrestoCentre Tools*

Acknowledging the strength of the pre-existing PrestoCentre Tools Catalogue, and emerging Research Outcome as expressed in PrestoKAT, the initial task was to map those schemas and taxonomies into a new schema for the Presto4U Tools Catalogue. Once migrated into the expanded schema, this facilitated use of the pre-existing records from the PrestoCentre website as the initial corpus of the Tool Catalogue.

Analysis of taxonomies across both interfaces revealed some common paths and deviations across the two interfaces. One such example is the representation of “Type.”

PrestoCentre tools were organised by type:

- Calculator
- Client Application
- Command Line
- Downloadable Database
- Downloadable Software

- Java Tool
- Library
- Manual
- Metadata schedule
- Model
- Toolkit
- Web Application
- Web Service

The PrestoKAT prototype database has organised “Type” in a similar way, with the exception that they also use the terms “Hardware” and “Integration” to describe “Types.” Further investigation into outside initiatives such as the Digital Curation Centre [14], COPTR [4], and LoC [5] illustrated differing approaches to the use of “Type” to describe tools. The mapping of these taxonomies allowed us to reflect on the current taxonomies used and to amend the taxonomies we will use for Presto4U outcomes. As such, the taxonomies used in the launch of the Tools Catalogue represent an initial taxonomy that will be refined throughout the remainder of the project.

4.2 *Creating a persistent link between the Tools Catalogue Standards Register*

To begin to build the integration between the Tools Catalogue and the Standards Register, we needed to create some form of persistent link between both the Catalogue and the Register. This link should allow users to switch seamlessly between the both interfaces. As an initial point of integration, the landing page of the Tools Catalogue, as reported in figure 3, contains a “View all standards” link back to the homepage of the Standards Register.

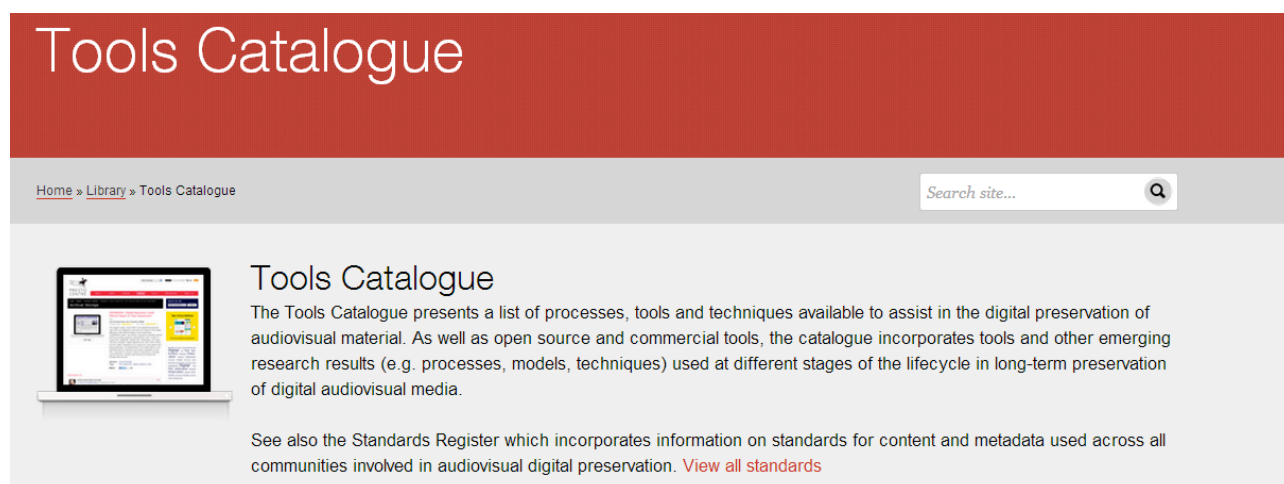


Figure 3. Tools Catalogue Landing Page - excerpted

As the Tools Catalogue develops, we will further investigate more transparent ways to link back to the Standards Register. This may for example include moving the hot-link for the Standards Register from the Tools Catalogue description into a stand-alone section of the Tools Catalogue landing page.

The link between the record page for an individual tool and the standards it employs, creates a powerful driver to trust and potential adoption of that tool. The Digital Curation Centre employs use of a field called “Standards Compliance” to list related standards [15]. The Tool Catalogue has expanded on this use case to create dynamic links between the tools record page and the record pages for related standards.

The screenshot shows a 'Documentation' section with the following details:

Available:	Yes Download java source code here: https://github.com/LibraryOfCongress/bagit-java Download python source code here: https://github.com/LibraryOfCongress/bagit-python
Fee:	
Resources:	Introductory video from the Library of Congress: http://www.digitalpreservation.gov/multimedia/videos/bagit0609.html
Standards compliance:	BagIt 0.97

Figure 4. Tools Catalogue BagIt Library record page - excerpted

In the above example reported in Figure 4, by clicking the link presented to the user on the tools record page [16], you are taken directly to the record page for that standard. [17]

Standards Register

[Home](#) » [Standards Register](#) » [BagIt](#) » BagIt 0.97

Description

BagIt 0.97

Version date:

Detailed name:

BagIt - BagIt File Packaging Format

Status:

Proposal

Other versions:

Description:

BagIt is a hierarchical file packaging format designed to support disk-based or network-based storage and transfer of arbitrary digital content. A bag consists of a 'payload' and 'tags'. The content of the payload is the custodial focus of the bag and is treated as semantically opaque. The 'tags' are metadata files intended to facilitate and document the storage and transfer of the bag. The name, BagIt, is inspired by the 'enclose and deposit' method, sometimes referred to as 'bag it and tag it'. [Citation: <http://tools.ietf.org/id/draft-kunze-bagit-09.txt>]

Reference:



Figure 5. Standards Register Record Page - excerpted

As current PrestoCentre tools are updated to incorporate the new schema for the Tools Catalogue links will be added where necessary to strengthen the relationship between tools and standards.

4.3 Tools Catalogue Landing Page

The Tools Catalogue [18] is currently part of the PrestoCentre Library, within the section header “Library items by type” [19]. Layout of the page mirrors that of the Standards Register, with columns and facets tailored to information relevant to the discovery of tools.

As well as the ability to “Search site” to query all records in the PrestoCentre website, the landing page allows you to “Search tools” to query records relating only to the Tools Catalogue.

The column view presents the user with a high-level snapshot of information on the tools listed, with columns chosen based on their perceived importance to the user of the tool:

- Name*
- Type*
- Function*

The catalogue displays 10 records per page with the functionality to browse back and forwards through pages. Facets (discussed further in “Discovery Mechanisms for the Tools Catalogue”) present the ability to browse the collection, based on controlled vocabularies and types relevant to audiovisual digital preservation.

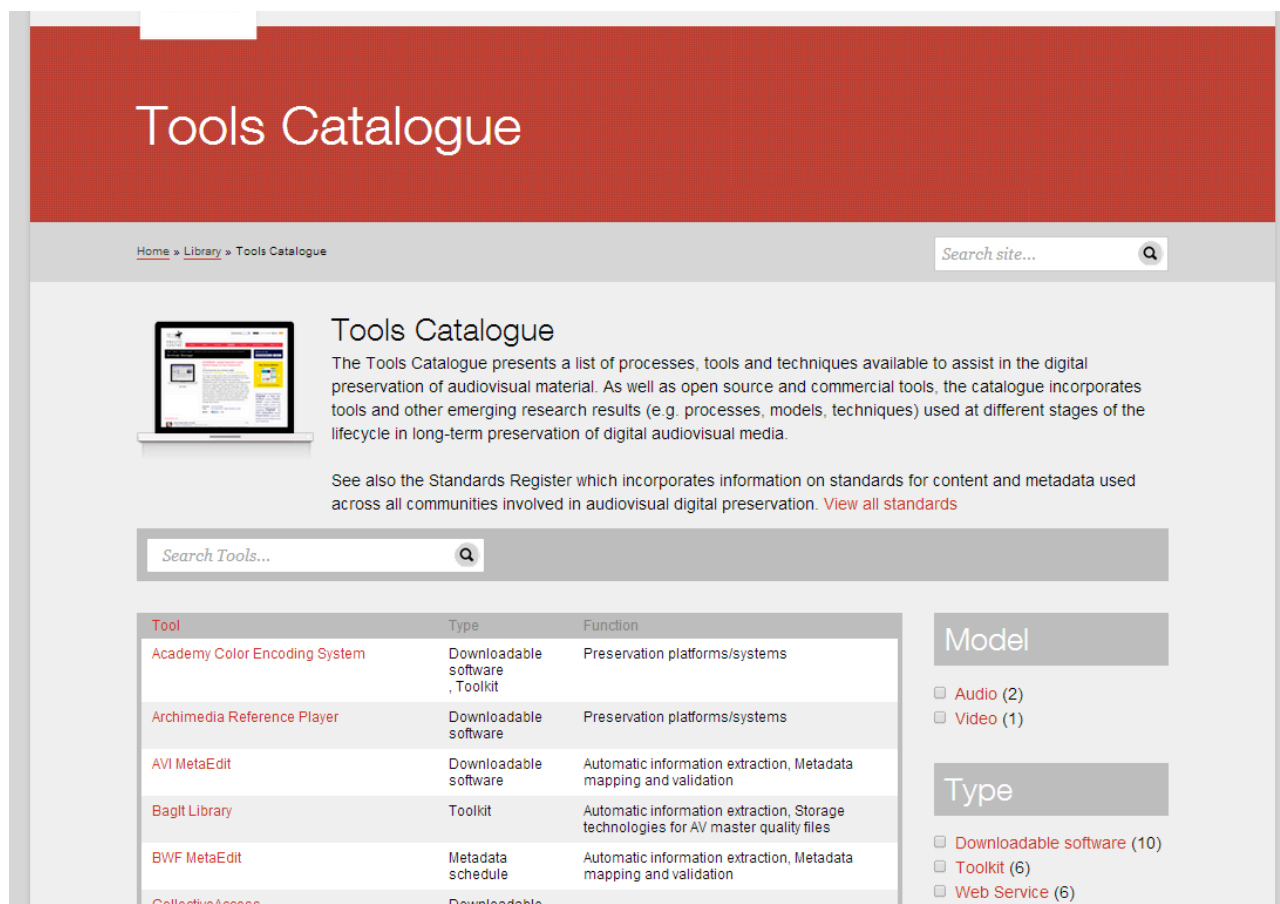


Figure 6. Tools Catalogue Landing Page - excerpted

4.4 **Tools Catalogue Record Page**

The record page has been specified and built to present two sections of information to users:

Description fields lists fields of information about the tool.

Sustainability Factors lists the dependencies around the use of the tool.

All fields display in the interface regardless of whether they contain data or not. Further user testing and feedback from the Communities of Practice will inform whether all fields display in this way.

4.4.1 **Description Fields**

Name

The formal name of a tool with any common abbreviations in brackets

Commercial type

Signifies whether tool is commercial or non-commercial

Type

List of categories that a tool is associated with

Function

General topic or functional purpose of a tool

Audiovisual lifecycle

Classification of records based on the OAIS lifecycle model

Version

Where applicable, the version release of a tool

Version date

Where applicable, the date of a version release

Description

Full description of the tool with citations, where relevant, to information reproduced from other authority sources.

Reference

Link to the authority website, or specific page within that website, with information on the tool

Tags

Non authority-controlled tagging of tools, facilitated primarily by admin users at PrestoCentre

Description

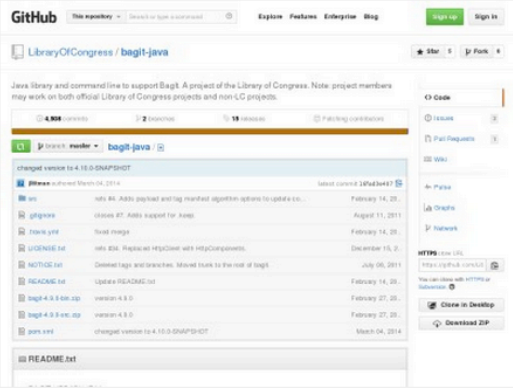
BagIt Library

Commercial type: Non-Commercial
Type: Toolkit
Function: Automatic information extraction, Storage technologies for AV master quality files
Audiovisual lifecycle: Ingest
Version: 4.9.0
Version date: January 2, 2014

Description:
BagIt is a specification for packaging of digital content for transfer and storage. The content of the package is known as a "bag" and includes content files and a checksum for each file in the bag. Machine readable text is included that facilitates receipt and validation of bags.

BagIt is a content-neutral specification, and as such it is equally applicable to use for storage and transfer of audiovisual content as it is for any other type of content. A number of tools exist to support bag creation and validation according to the BagIt specification, including a platform-independent GUI called Bagger, a command line Java library, and a Ruby library.

Reference:



Tags: <https://github.com/LibraryOfCongress/bagit-java>
Metadata Trusted Digital Repositories Storage Ingest
Open Source Fixity File Copy

Share




Figure 7. Tools Catalogue BagIt Library record page – Description

4.4.2 Sustainability Factors

Sustainability covers a wide range of information relating to a tool from the author to the license to user documentation. Where available, it presents the users with links to source code and guidelines to allow them to make informed choices about open source versus vendor-based tools with commercial licenses. To better guide users through this section, the information has been divided into 3 sub-sections:

Development fields lists the provenance and technical specifications of the tool

License and Adoption fields list the license, potential costs and relevance to audiovisual communities

Documentation fields lists resources to investigate or install a tool

4.4.2.1 Development

Author

The organisation or developer responsible for the creation and development of a tool

TRL Level

An assessment calculation of the technology readiness of a tool based on a value list of 1-9, with 1 representing the raw idea of a tool and 9 representing a system or tool that has been released into the market and in use

Platform

Operating system on which the software operates

Platform Specifics

Description of exclusions to the use of a tool on a certain platform, such as a tool only operating on certain versions of Windows.

Software dependencies

Description of any additional software that is integral to the functioning of a tool

Developing community

For open-source tools, the name of the communities continuing to develop and support the tool

Further developments expected

Description of any further development expectations, with dates where applicable

4.4.2.2 License and Adoption

License Model

Expresses whether a license is attached to a tool, with a controlled vocabulary of license types

License fee

Indicates whether a fee is associated with download or use of a tool

License URL

Where applicable, a hot-link URL to the license terms associated with a tool

License citation

Where appropriate, citation information relating to the use of the license URL will be expressed

Domain

List of Communities of Practice a tool will be of interest to in digital preservation workflows

Community adoption

The representation of a sliding scale from light to dark to illustrate those Communities of Practice who have actively adopted use of a tool, with a controlled vocabulary to express the formal names of the communities.

Implemented at

List of organisations who have implemented a particular tool, with a hot-link (where applicable) back to that organisation.

4.4.2.3 Documentation

Available

Yes/No representation of the availability of documentation, such as schemas or source code relating to a tool

Fee

Indication if a fee is applicable to access documentation relating to a tool

Resources

List of related resources such as tutorials on use or implementation of a tool

Standards compliance

The name(s) of standards employed within a tool, with hot-links back to the relevant record page(s) in the PrestoCentre Standards Register

Sustainability Factors

Development

Author: Library of Congress

TRL level:

Platform: Windows

Platform specifics:

Software dependencies: Java 6

Developing community:

Futher developments expected: Yes

License and Adoption

License model: Open


Licence fee:


License URL: <https://github.com/LibraryOfCongress/bagit-java/blob/master/LICENSE.txt>

Licence citation:

Domain: [Film Collections and Filmmakers](#)
[Footage Sales Libraries](#)
[Learning and Teaching Repositories](#)
[Music and Sound Archives](#)
[Personal Audiovisual Collections](#)
[Research and Scientific Collections](#)
[TV, Radio and New Media Broadcasting](#)
[Video Art, Art Museums and Galleries](#)
[Video Production and Post-Production](#)

Community adoption:

TV, Radio and New Media Broadcasting  **low** **high**

Learning and Teaching Repositories  **low** **high**

Implemented at:

Documentation

Available: Yes

Download java source code here:
<https://github.com/LibraryOfCongress/bagit-java> Download python source code here: <https://github.com/LibraryOfCongress/bagit-python>

Fee:

Resources: Introductory video from the Library of Congress:
<http://www.digitalpreservation.gov/multimedia/videos/bagit0609.html>

Standards compliance: [BagIt 0.97](#)

Figure 8. Tools Catalogue BagIt Library record page – Sustainability Factors

4.4.3 Research outputs

The fields displayed, and the layout of those fields, represent the preliminary user interface. Several fields have been created in the backend in order to support information relevant to Research Outputs. For now they have not been exposed in the user interface as year one focused on a small preliminary set of research outcomes, the results of which

can be found in the D3.2 “Research Outputs Assessments v1” report [6]. Year two continues the investigation into emerging research outputs based on Community of Practice needs. The results of these investigations will be reported in D3.3 “Research Outputs Assessment v2,” due for publication at the end of year two, December 2014. At that point the data gathered from the year two testing will be added to the Tools Catalogue and exposed in the user interface.

The preliminary list of fields ported over from PrestoKAT and seen as most relevant to Research Outputs are:

Start year

The date a tool is first recognised as a development idea

Finish year

The date development ends for a tool

Output Type

The type of research output such as hardware, software or project. This has been mapped from PrestoKAT “Early Research Output”

Throughout year two of the project, we will be working with the team responsible for the D3.3 report, to ensure we track and expose any further fields that are deemed relevant to express data relating to the development lifecycle of a tool.

4.5 Discovery Mechanisms for the Tools Catalogue

Facets provide an immersive environment for users to explore the Tools Catalogue. Both the pre-existing PrestoCentre Tools section and PrestoKAT employed the use of facets.

One of the preliminary tasks of the project was to map those facets to identify areas of commonality and difference and to create a taxonomy that best suited the needs of the emerging Tools Catalogue. For example, PrestoCentre organised tools by both type and keyword, but differed in controlled vocabulary from Research Outputs. There was little commonality between PrestoCentre keywords and those used for “Software Type” and “Tool Type” in the Research Outputs database. All value lists have now been harmonised to create a unified browsing structure across different outputs on the website.

As work continues on Research Outcomes in year two of the project, some facets may be adapted to better illustrate Type and Function for tools, or create new facets to better describe Research Outputs. In addition, extensive research has been undertaken into the taxonomies applied to other Tools Catalogues, such as COPTR [4], Library of Congress Tools [5] and POWRR [20].

As year two progresses the PrestoCentre Tools Catalogue will be re-assessed and adapted to incorporate additional terms suitable for describing digital audiovisual tools.

4.5.1 Facets as an integration point with Research Outcomes

Incorporation of the PrestoKAT structure involved the creation of new facets to accommodate the status of Research Outputs in the development lifecycle. Specifically this involved the creation of facets for Software Function Type and TRL-Level.

Software Functional Type

It is the general topic or functional purpose of a tool. Among others, the current list is made up of:

1. Metadata mapping and validation
2. Storage technologies for AV master quality files
3. Automatic information extraction
4. Quality assessment technologies
5. Manual content annotation
6. Rights management technologies and formats
7. Preservation platforms/systems

“Software Functional Type” was incorporated into the Tools Catalogue as renamed as “Function.”

TRL Level

Technology Readiness Level, i.e. the user can provide an initial estimate of the TRL level of the software. In the following the 9 possible levels listed in the facet.

1. The tool is described as an idea, not proven and no implementation. Software can be demonstrated that shows feasibility of tool, but no optimisation or other practical considerations
2. Some proving has been performed suggesting that the tool is practical (algorithm or other analysis) but no implementation
3. Software can be demonstrated that shows feasibility of tool, but no optimisation or other practical considerations
4. Software has been created and run in a laboratory environment demonstrating the tool architecture, scalability, fitness for purpose, reliability of algorithms
5. Key components of the tool have been tested in the laboratory with real input data and the performance has been successfully demonstrated
6. All components run together in an integrated system in a laboratory environment using real test data and performance, including scalability and robustness has been successfully demonstrated
7. A prototype version of the system has been tested by a representative user in a production environment using real input data (alpha test)
8. Sales-ready system in the final form has been released for beta test to users for use in real production environments
9. Commercial system derived and in use by users in their day to day operations

“TRL” was incorporated into the Tools Catalogue as a numerical scale. Further deployment may include a description of TRL levels to afford users better understanding of this development rating scale.

4.5.2 Facets as an integration point with the Standards Register

To reinforce the relationship between the Standards Register and Tools Catalogue it was important to incorporate some of the links to both Communities of Practice and digital preservation workflows from the Standards Register into the Tools Catalogue. The following facets were implemented to fulfil this remit:

Domain

Signifies whether a particular tool should be of use to a particular audiovisual community or communities.

For example, a tool such as Archivematica [21] may be of more use to the Video Art, Art Museums and Galleries.

Presto4U Community Domains

It lists the current communities involved:

- Film Collections and Filmmakers
- Footage Sales Libraries
- Learning and Teaching Repositories
- Music and Sound Archives
- Personal Audiovisual Collections
- Research and Scientific Collections
- TV, Radio and New Media Broadcasting
- Video Art, Art Museums and Galleries
- Video Production and Post-Production

Audiovisual lifecycle

Signifies the five stages in the audiovisual lifecycle OAIS functional model [6]

- Access
- Archival Storage
- Data Management
- Ingest
- Management
- Preservation Planning

4.5.3 Implementation of Facets in Tools Catalogue

After careful consideration, the following facets were decided as being most powerful as a browse mechanism to present to users:

Model

Categorisation of the audiovisual model relevant to the tool, such as relevance to Audio or Video workflows

Type

List of categories that a tool is associated with

Function

General topic or functional purpose of a tool

License Model

Expresses whether a license is attached to a tool, with a controlled vocabulary of license types

Domain

List of Communities of Practice a tool will be of interest to in digital preservation workflows

Lifecycle

Classification of records based on the OAIS lifecycle model

Platform

Operating system on which the software operates

TRL Level

An assessment calculation of the technology readiness of a tool based on a value list of 1-9, with 1 representing the raw idea of a tool and 9 representing a system or tool that has been released into the market and in use

Tags

Non authority-controlled tagging of tools

Facets have been implemented as emerging discovery functionality. The first 3 facet categories are listed on the page, with the dialogue to “See more filters” and to collapse back to the display of the first three facet categories with the user dialogue of “Show less filters.” In this way users are not overwhelmed with a long list of facets to initially select from, and are presented with the option to explore facets in a deeper fashion.

The screenshot shows the 'Tools Catalogue' record page. At the top, there is a red header with the title 'Tools Catalogue'. Below the header, there is a navigation breadcrumb 'Home > Library > Tools Catalogue' and a search bar. The main content area features a small image of a laptop displaying the catalogue interface, followed by the title 'Tools Catalogue' and a descriptive paragraph. Below this, there is a link to 'View all standards'. A search bar for tools is positioned above a table of tool records. The table has three columns: 'Tool', 'Type', and 'Function'. To the right of the table, there are three filter sections: 'Model', 'Type', and 'Function', each with a list of filter options and counts. At the bottom of the table, there is a pagination control.

Tool	Type	Function
Academy Color Encoding System	Downloadable software , Toolkit	Preservation platforms/systems
Archimedia Reference Player	Downloadable software	Preservation platforms/systems
AVI MetaEdit	Downloadable software	Automatic information extraction, Metadata mapping and validation
BagIt Library	Toolkit	Automatic information extraction, Storage technologies for AV master quality files
BWF MetaEdit	Metadata schedule	Automatic information extraction, Metadata mapping and validation
CollectiveAccess	Downloadable database	
D10 SumChecker	Command Line, Library	
DCC Curation Lifestyle Model	Model	
DRAMBORA: Digital Repository Audit Method Based on Risk Assessment	Toolkit	
DROID (Digital Record Object Identification)	Downloadable software	

Model

- Audio (2)
- Video (1)

Type

- Downloadable software (10)
- Toolkit (6)
- Web Service (6)
- Manual (4)
- Calculator (3)

▼ Show more

Function

- Automatic information extraction (3)
- Metadata mapping and validation (2)
- Preservation platforms/systems (2)
- Manual content annotation (1)
- Storage technologies for AV master quality files (1)

▼ Show more filters

Figure 9. Tools Catalogue record page – Facets

4.6 Technical Implementation

The Tools Catalogue is one of the services that Presto4U has now published through the existing PrestoCentre online platform. The PrestoCentre online platform is built as an open-source Drupal7 platform and can handle core complex user permissions and workflows. By choosing Drupal as platform, the Tools Catalogue has at its disposal an expandable and flexible setup for the future of the service, which is especially important towards the Market Place integration later in the project. It also continued from the already implemented Community of Practice personalisation and filter features from Deliverable 2.3 (Community of Practice online workspaces).

The Tools Catalogue was implemented as a separate content type. Currently, only PrestoCentre administrators are able to create and maintain the tools records, but Drupal's flexibility to organise and distribute administrative and content access rights will enable others to publish and easily maintain their records in the future, contributing to the overall sustainability of the Tools Catalogue over time. Presto4U Tools assessment actions (primarily from the Partners involved in WP3, and Community of Practice leaders, will add to the correct description of records, fields and taxonomies in the Tools Catalogue. They will work with the PrestoCentre administrators on the content creation and maintenance. With the Drupal Workflow module, they PrestoCentre administrators can decide on the steps needed for content to become public. And with Drupal's Organic Groups module the project can work on different versions of tools should this need arise in the future. Overall, more than 100 Drupal modules are currently active for the functioning of the Tools Catalogue (and all the related services). These range from very basic modules such as the "View" module, through for the Catalogue specifically required modules like "Hierarchical Select Taxonomy," "Workbench Moderation," "PagePeeker Screenshots," to highly adapted and tweaked modules such as "Taxonomy revision." The field values of the Tools Catalogue can be easily exported to MySQL and mapped to other databases for creating possible future interoperability and crossovers with other database services, including PrestoKAT (although a direct data exchange will require some further work).

5. Conclusion

The launch of the Tools Catalogue represents the initial functionality for this service. As the core set of records are migrated forward into the new schema, and new tools are added, development will begin on the expansion of functionality and design to integrate and interact with the Marketplace. The project aims to create multiple avenues to explore the use of software tools. This may involve links to source code on open repositories and concurrent versioning systems such as SourceForge [11] and Github [13] for those wishing to do a deep dive into the backend of a tool; manuals or guidelines to explore the use of tools; blogs or use cases by individual institutions adopting a particular tool or suite of tools as a means of creating a community forum to discuss the use; or potentially creating a space where vendors or emerging research outputs can host demonstrations. Design and technical build will also reflect any integration points that emerge as a result of potential collaborations with other initiatives related to preservation tools.

The Tools Catalogue is an immersive portal allowing users of all skill-sets to browse and discover information and resources around tools. The focal point of the register is the community involvement in the use of tools. What tools are relevant to any community, as well as what tools have been adopted, are central concepts to the Tools Catalogue. The outputs of work on emerging Research Outcomes [6] in year one, as well as the expertise of the core project team in identifying an initial corpus of tools felt to be relevant to digital audiovisual preservation will form the basis of the software Tools Catalogue. By the end of year two of the project, further Research Outcomes will have been identified based on identification of Community of Practice needs. There will also be an on-going effort by the core experts on the project to continue to identify tools relevant to audiovisual preservation and to add them to the catalogue. By presenting these expert advices on what tools are relevant to particular communities, we can effectively guide users through what can sometimes seem the overwhelming options that are available within digital preservation.

References

- [1] Presto4U, D4.2 Interim Report on Audio visual and Preservation Standards
- [2] Presto4U, The Standards Register,
<https://www.prestocentre.org/standards>, Last accessed: 16 April 2014
- [3] APARSEN Standards
<http://www.alliancepermanentaccess.org/standards/menu.php>. Last accessed: 17 April 2014
- [4] Presto4U, Community Owned Digital Preservation Tools Registry,
http://coptr.digipres.org/Main_Page, Last accessed: 16 April 2014
- [5] Library of Congress Tools Showcase
www.digitalpreservation.gov/tools/, Last accessed: 16 April 2014
- [6] Presto4U, D3.2 Research Outputs Assessments v1
https://www.prestocentre.org/files/deliverable_d3.2_presto4u_21_12_2013_v1r.pdf, Last accessed: 16 April 2014
- [7] [Technology Readiness Level \(TRL\)](http://en.wikipedia.org/wiki/Technology_readiness_level),
http://en.wikipedia.org/wiki/Technology_readiness_level, Last accessed: 16 April 2014
- [8] [The NASA Technology Readiness Level Introduction](http://web.archive.org/web/20051206035043/http://as.nasa.gov/aboutus/trl-introduction.html),
<http://web.archive.org/web/20051206035043/http://as.nasa.gov/aboutus/trl-introduction.html>, Last accessed: 16 April 2014
- [9] CCDS, OAI preservation lifecycle.
http://en.wikipedia.org/wiki/Open_Archival_Information_System, Last accessed: 16 April 2014
- [10] Presto4U, D2.1 CoP Knowledge Schema
- [11] [Source Forge](http://sourceforge.net/)
<http://sourceforge.net/>, Last accessed: 16 April 2014
- [12] Google Code
<https://code.google.com/>. Last accesses 17 April 2014
- [13] [Github](https://github.com/), <https://github.com/>, Last accessed: 16 April 2014
- [14] Digital Curation Centre, DCC Tools & Services
<http://www.dcc.ac.uk/resources/external/tools-services>, Last accessed: 16 April 2014
- [15] Digital Curation Centre, DCC Tools & Services – The BagIt Library
<http://www.dcc.ac.uk/resources/external/bagit-library>, Last accessed: 16 April 2014
- [16] Presto4U, Tools Catalogue – The BagIt Library

<https://www.prestocentre.org/library/tools-catalogue/bagit-library>, Last accessed: 16 April 2014

[17] Presto4U, Standards Register – BagIt
<https://www.prestocentre.org/standards/bagit/0.97>, Last accessed: 16 April 2014

[18] Presto4U, Tools Catalogue
<https://www.prestocentre.org/library/tools-catalogue>, Last accessed: 16 April 2014

[19] PrestoCentre Library
<https://www.prestocentre.org/library>, Last accessed: 16 April 2014

[20] The Preserving Digital Objects with Restricted Resources Community, The Tool GRID,
<http://digitalpowrr.niu.edu/tool-grid/>, Last accessed: 16 April 2014

[21] Artefactual System, The Archivematica digital preservation platform,
https://www.archivematica.org/wiki/Main_Page, Last accessed: 16 April 2014

Document information

Delivery Type Other
Deliverable Number 4.6
Deliverable Title Software Tools Catalogue
Due Date 01/04/2014
Submission Date 30/04/2014
Work Package 4
Partners KINGS COLLEGE LONDON, IT INNOVATION, BEELD EN GELUID
Author(s) K. Colbron (KCL), Marius Snyders (B&G)
Reviewer(s) W. Allasia (EURIX)
Keywords Software tools, multimedia preservation
Document Identifier Deliverable_D4.6_presto4u_30_04_2012_v2_0(R).doc
Dissemination level Other
Document Status Released

Project Acronym Presto4U
Project Full Title European Technology for Digital Audiovisual Media Preservation
Grant Agreement 600845
Project Coordinator Beeld en Geluid
Contact Details Sumatralaan 45, 1217GP Hilversum, The Netherlands.
msnyders@beeldengeluid.nl

Document Status Sheet

Version	Delivery Date	Comment	Author
1.0	14/04/2014	Draft	K. Colbron
1.1	16/04/2014	Reviewed draft	W. Allasia
1.2	17/04/2014	Revised draft	K.Colbron
1.3	28/04/2014	Final review	M. Snyders
2.0	29/04/2014	Ready for release	K. Colbron
2.0(R)	30/04/2014	Release	