

ISKO Media

KO for management, dissemination and preservation

Simon Wilson
si.j.wilson@gmail.com

The value of an AV archive

- **Reference**
 - A record of what an organization has achieved, the ideas it has fostered, the organisations it has collaborated with
 - A valuable source for others working on research in the field
- **Engaging new audiences**
 - Content relevant to fellow-workers around the world
 - Raise profile of ISKO UK

ISKO Media

- ISKO Media is ISKO UK's repository of event recordings
- Collection began in 2007
- A rich record of a rapidly changing knowledge environment
 - New technologies: GPT-2
 - New threats: fake news
- Migration for the future
- <http://c2264562.myzen.co.uk/content/iskomedia>

Automatic Metadata Generation - A Better Alternative to Controlled Vocabularies

Automatic metadata generation for resource discovery. This research exercise was carried out for JISC in 2006. Aimed to establish the state of the art regarding machine extraction and generation of metadata. Considered the whole range of metadata: descriptive, technical, rights, preservation, subject, LOM, etc. Embraced both intrinsic and extrinsic metadata. Extrinsic metadata tools show rapid development. The majority of tools use a controlled vocabulary or other subject authority.

Presentation Type:

Talk

Language:

English

Presenter:

Vanda Broughton

Presentation Audio:



[ISKO_VandaBroughton_040907.mp3](#)

Presentation Visual:



[broughton_ISKOUKseminar1.pdf](#)

Tags (comma-separated):

[metadata](#), [generation](#), [tools](#), [digital resource description](#), [machine extraction](#)

Planning an AV migration project

- Audit material
 - Identify requirements
 - Compare solutions
 - Prepare data & metadata
 - Complete migration
- 
- Maximise value of content

The content

- **Audio**
 - Drupal: 313 Files, 2007-2018
 - Wild Apricot: 2018 onward
 - Total: c.400
- **Presentations**
 - Drupal: 235
- **PDFs**
 - Drupal: 488 (not all ISKO Media)
- **Some video**
 - Drupal: 4 (.mp4)



→ **The chosen solution**

The current state of ISKO Media

- **Dispersed**
 - Drupal site ('Presentations' and 'Event archive')
 - Recent material on Wild Apricot
 - More recent content not online
- **Silo**
 - Material primarily findable from ISKO web pages
 - Value of cross-searchability
- **Preservation risk**
 - Content backed up on external HDDs vulnerable to corruption
 - Web hosting is not long-term preservation

Requirements:

- Content is consolidated
- Content is cross-searchable
- Content is preserved

Requirements

- In the chosen solution the collection will be:
 - Consolidated
 - Searchable
 - Browsible
 - Accessible
 - Shareable
 - Preserved

Options considered

- There are a superfluity of digital collection management solutions out there

- Ashley Blewer maintains a useful list:

https://docs.google.com/spreadsheets/d/1cXOug3qM0pNNeD_wssiVEv9c0W1Y5I1VDTnSPTk7fb4/edit#gid=0

Platform	Model	Support for types of content	Streaming delivery	Cross-searchable	Website
Zenodo	Free service	AV and documents	Not supported	Yes	https://zenodo.org/
Figshare	Free service with paid option	AV and documents	Supported	Yes	https://figshare.com/
Dspace	Free software with entailed support cost	AV and documents	Supported with customization	No	https://duraspace.org/dspace/
Internet Archive	Free service	AV and documents	Supported	Yes	https://archive.org/
Avalon Media System	Free software	AV	Supported	No	https://www.avalonmediasystem.org/
Greenstone	Free software	AV and documents	Supported with customization	No	http://www.greenstone.org/

Recommended options: 'Preservation and Access'

- Zenodo

- Archival repository maintained by CERN
- Rich, flexible metadata schema
 - Supports ISKO's 'Classification System for Knowledge Organization Literature'
- Handles all content types
- Complexly searchable
- OAI-PMH
- Assigns DOIs to content – persistent identifiers
- *Content is not streamable*

Communities recommended

Start typing a community name...

International Society for Knowledge Organization (ISKO)

Upload type required

Publication Poster Presentation Dataset Image Video/Audio Lesson Other Software

Publication type

Subjects optional

Specify subjects from a taxonomy or controlled vocabulary. Each term must be uniquely identified (e.g. a URL). For free form text, use the keywords field in basic information section.

Subjects

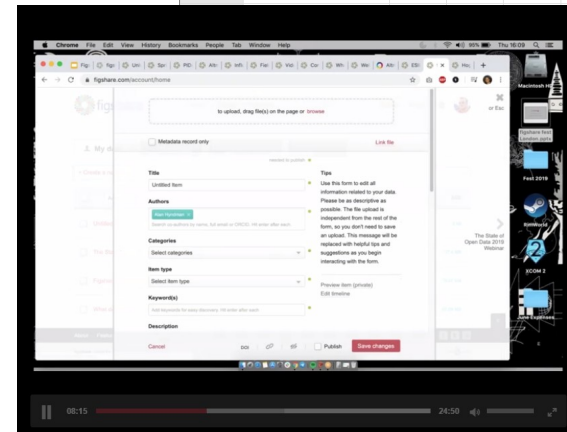
+ Add another subject

Recommended options: ‘Preservation and Access’

- Figshare

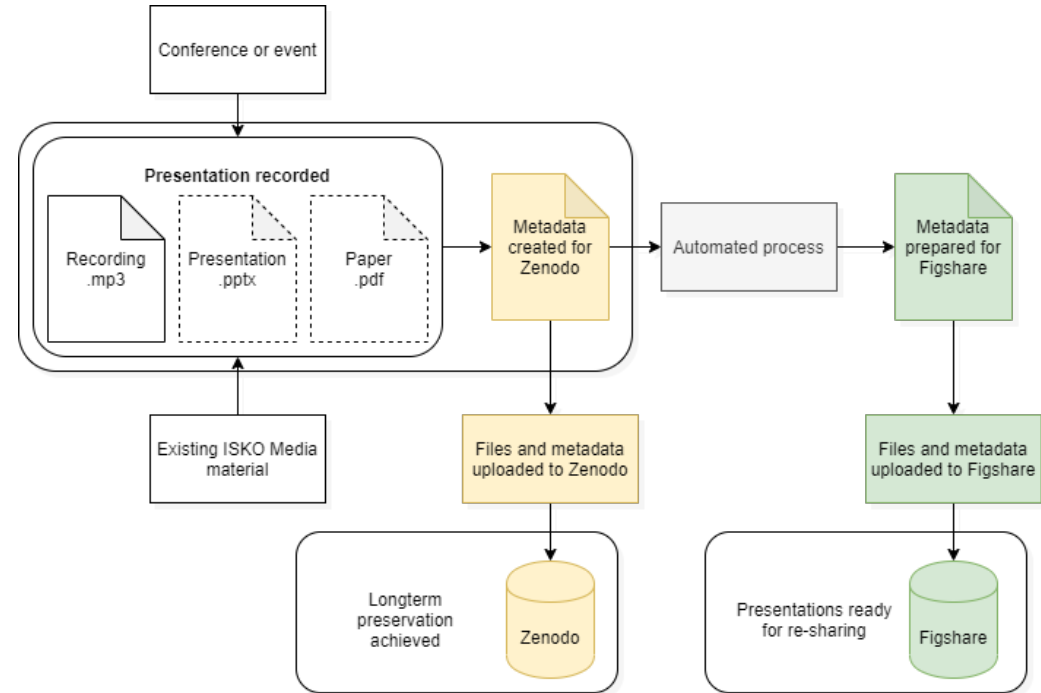
- Designed as a research dissemination platform
- Supports in-browser display of AV media
- Simpler metadata schema than Zenodo, narrower range of subjects
- Content can be browsed by collection or event, e.g.
 - *Conference*
 - *Paper 1*
 - *Paper 2*

	A	B	C	D	E	F	G	H	I
1	Clinical Scores								
2	Day Post Infection	RM1	RM2	RM3	RM4	RM5	RM6	RM7	RM8
3	0	3	3	3	3	3	3	3	3
4	1	25	18	18	10	8	10	20	13
5	2	25	10	18	18	15	13	20	18
6	3	18	13	13	15	18	13	20	18
7	4					20	15	28	18
8	5					18	18	25	13
9	6					20	20	25	13
10	7					8	8	18	13
11	8					8	8	15	8
12	9					3	8	10	5
						3	3	8	3



Next steps

- **Standardize metadata for Zenodo**
 - ORCID IDs for authors
 - ‘Classification System for Knowledge Organization Literature’ IDs
- **Prepare metadata for Figshare**
 - Merge appropriate Zenodo fields into a general description
 - Assign broad subject codes



Metadata preparation

- Digital collection management and media migration depend on high-quality metadata
- Properly structured metadata will:
 - Facilitate automation using public APIs
 - Enable fixity-checking
 - Create linked data (ORCID, KOLIT)
- **Some prepared metadata**

1	Resource description
2	Resource discovery
3	Administration and management of resources
4	Record of intellectual property rights
5	Documenting software and hardware environments
6	Preservation management of digital resources
7	Providing information on context and authenticity

- Day's model of metadata purposes



Q&A