



Preservation of Interoperability and Interoperability of Preservation

**Seamus Ross, University
of Toronto**

4 October 2010

DL.org Autumn School – Athens, 3-8 October 2010



What is a digital library

“A digital library is the **infrastructure, policies and procedures**, and organisational, political and economic mechanisms necessary to enable access to and preservation of digital content”.

Source: Ross, S., Digital Library Development Review. Final report, National Library of New Zealand, July 2003,
http://eprints.erpanet.org/50/01/ross_report.pdf

Trains, Cars and DLs

- Some folks teaching with
- Complexity,
- Ian Ayres, SuperCrunchers
- Processable articles

Interoperability: Yet Another Definition

“**Interoperability** is a property referring to the ability of diverse systems and organizations to work together (inter-operate). The term is often used in a technical systems engineering sense, or alternatively in a broad sense, taking into account social, political, and organizational factors that impact system to system performance.”

Source: <http://en.wikipedia.org/wiki/Interoperability>

But what is interoperability?

- Is it a representation problem?
- Is it a semantic problem?
- Is it a process problem?
- Is it possibility to define generic interoperability objectives?
- Can we create transformation services to enable interoperability across time?



Seven Key Interoperability Issues

- *Process* – what is the boundary between static content, representations, linkages
- *Authenticity* – how do we (people and machines) know ‘it’ is authentic
- *Quality* – how do we measure quality and does it change overtime
- *Change over time* – how do we create ‘dynamic interoperability’ frameworks
- *Policy* – how do we reconcile policies in a contemporary context and how do we handle policy drift
- *Legal* – how can we address issues related to legal aspects
- *Preservation* – how do we preserve ‘interoperability potentiality’ what do we preserve.

Interoperability

Value and Benefits of addressing lack of interoperability

- Layered Approach across systems, space and time
- Levels of Abstraction – functionality, data
- Interoperability Parameters
 - Syntactical
 - Semantic
 - Content
 - Functionality
 - Context
- Object binding, boundaries and change

Which priorities for interoperability?

- Understand the uses that those at the leading edge are making or want to make of DLs
- Use qualitative as well as quantitative approaches
- Partner with researchers in the social sciences
- Study users in various disciplines

Who is creating digital libraries?

- Traditional libraries (universities, governments, etc.)
- Researchers
- Students
- General public

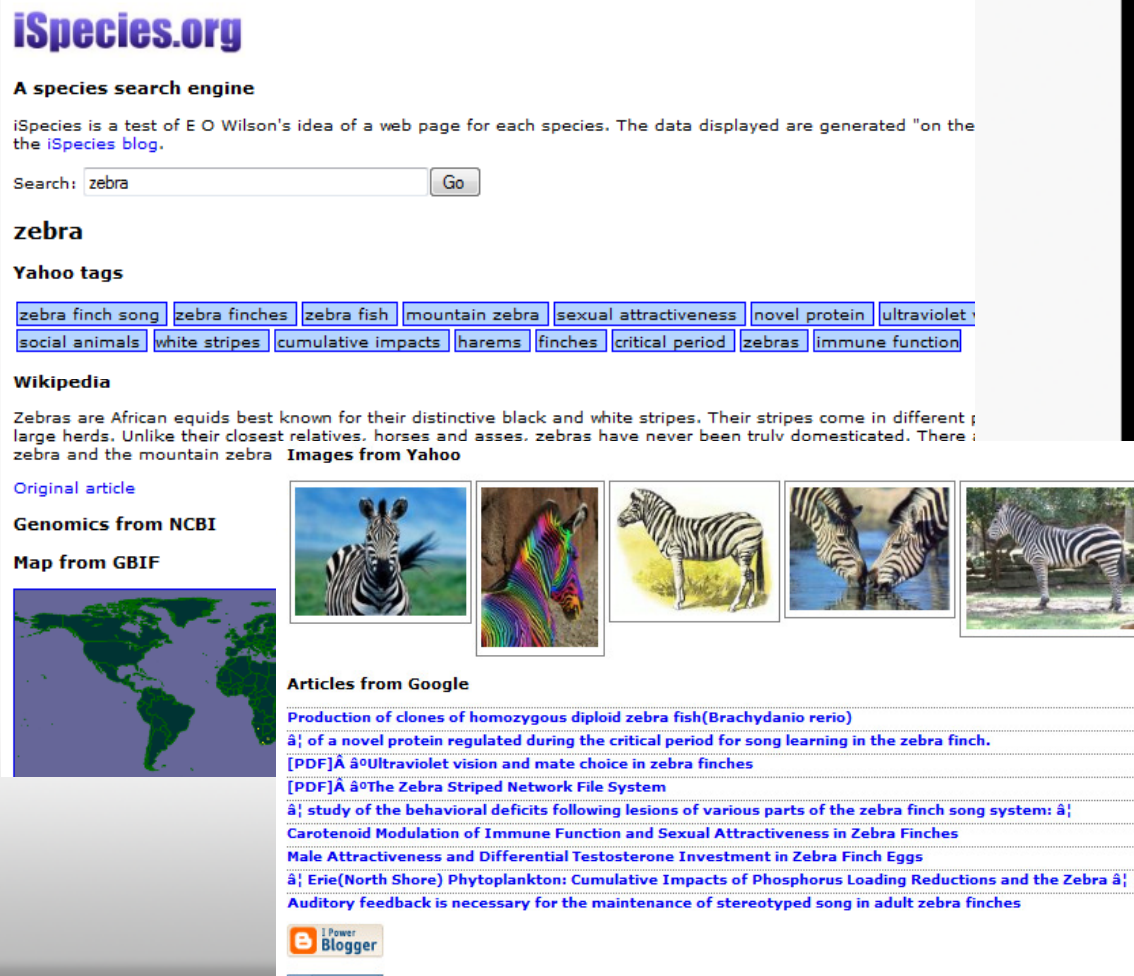


Actual and potential interoperability

- Interoperate with:
 - People
 - metadata
 - simple and compound digital items (text, still and moving images, audiovisual files, 3D files)
 - data
 - services and environments
 -

Interoperate with biological journals, reference tools, species collections?

- iSpecies.org
- A researcher at University of Glasgow creates a species search engine yielding many types of data (text, GIS, pics..)



iSpecies.org
A species search engine

iSpecies is a test of E O Wilson's idea of a web page for each species. The data displayed are generated "on the fly" from the [iSpecies blog](#).

Search:

zebra

Yahoo tags

[zebra finch song](#) [zebra finches](#) [zebra fish](#) [mountain zebra](#) [sexual attractiveness](#) [novel protein](#) [ultraviolet](#) [social animals](#) [white stripes](#) [cumulative impacts](#) [harems](#) [finches](#) [critical period](#) [zebras](#) [immune function](#)


Wikipedia

Zebras are African equids best known for their distinctive black and white stripes. Their stripes come in different patterns and colors. They live in large herds. Unlike their closest relatives, horses and asses, zebras have never been truly domesticated. There are three species of zebra and the mountain zebra. [Images from Yahoo](#)

[Original article](#)

Genomics from NCBI

Map from GBIF

Articles from Google

[Production of clones of homozygous diploid zebra fish\(Brachydanio rerio\)](#)

[A novel protein regulated during the critical period for song learning in the zebra finch.](#)

[\[PDF\] Ultraviolet vision and mate choice in zebra finches](#)

[\[PDF\] The Zebra Striped Network File System](#)


[study of the behavioral deficits following lesions of various parts of the zebra finch song system:](#)

[Carotenoid Modulation of Immune Function and Sexual Attractiveness in Zebra Finches](#)

[Male Attractiveness and Differential Testosterone Investment in Zebra Finch Eggs](#)

[Erie\(North Shore\) Phytoplankton: Cumulative Impacts of Phosphorus Loading Reductions and the Zebra](#)

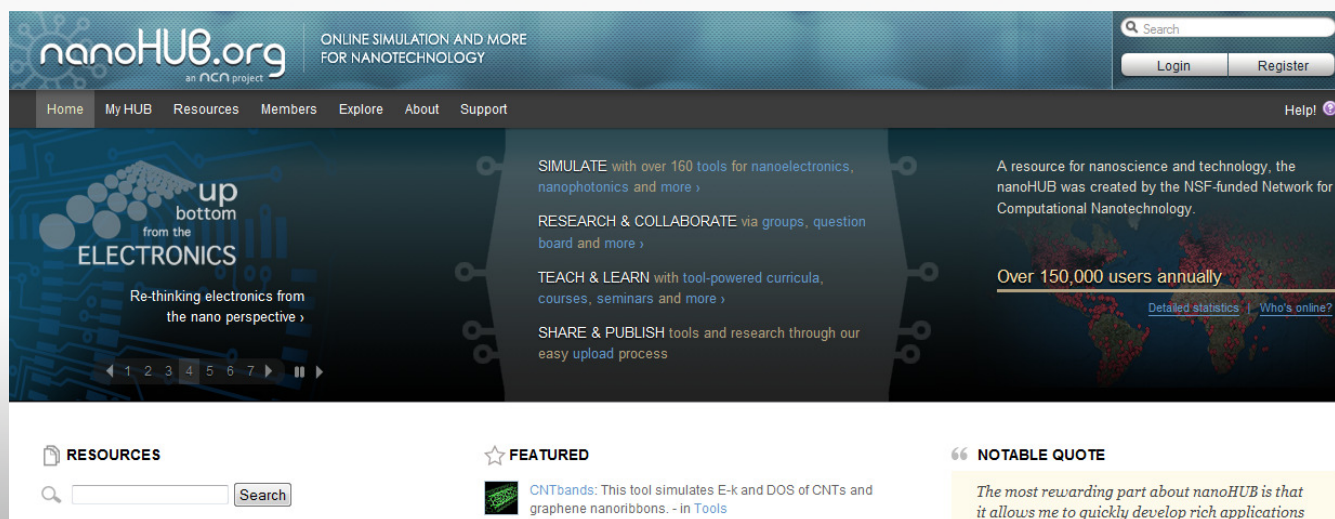
[Auditory feedback is necessary for the maintenance of stereotyped song in adult zebra finches](#)

 [Blogger](#)

Hub Zero and nanoHUB at Purdue

- Access to collaborative simulation tools
- Access to Grid environment
- Collaborative Web 2.0 environment
- Oriented to teaching and learning

<http://nanohub.org/>




The screenshot shows the nanoHUB.org website. The header includes the logo, the text "ONLINE SIMULATION AND MORE FOR NANOTECHNOLOGY", and a search bar with "Login" and "Register" buttons. The navigation menu contains "Home", "My HUB", "Resources", "Members", "Explore", "About", "Support", and "Help!". The main content area features a large banner for "up bottom from the ELECTRONICS" with the tagline "Re-thinking electronics from the nano perspective". Below the banner are four columns of text: "SIMULATE with over 160 tools for nanoelectronics, nanophotonics and more", "RESEARCH & COLLABORATE via groups, question board and more", "TEACH & LEARN with tool-powered curricula, courses, seminars and more", and "SHARE & PUBLISH tools and research through our easy upload process". To the right, there is a world map and the text "Over 150,000 users annually" with links for "Detailed statistics" and "Who's online?". The footer contains three sections: "RESOURCES" with a search bar, "FEATURED" with a star icon and a link to "CNTbands", and "NOTABLE QUOTE" with a quote: "The most rewarding part about nanoHUB is that it allows me to quickly develop rich applications".

Librarian-created video at iSchool, University of Washington

You Tube | [Browse](#) | [TV Shows](#)

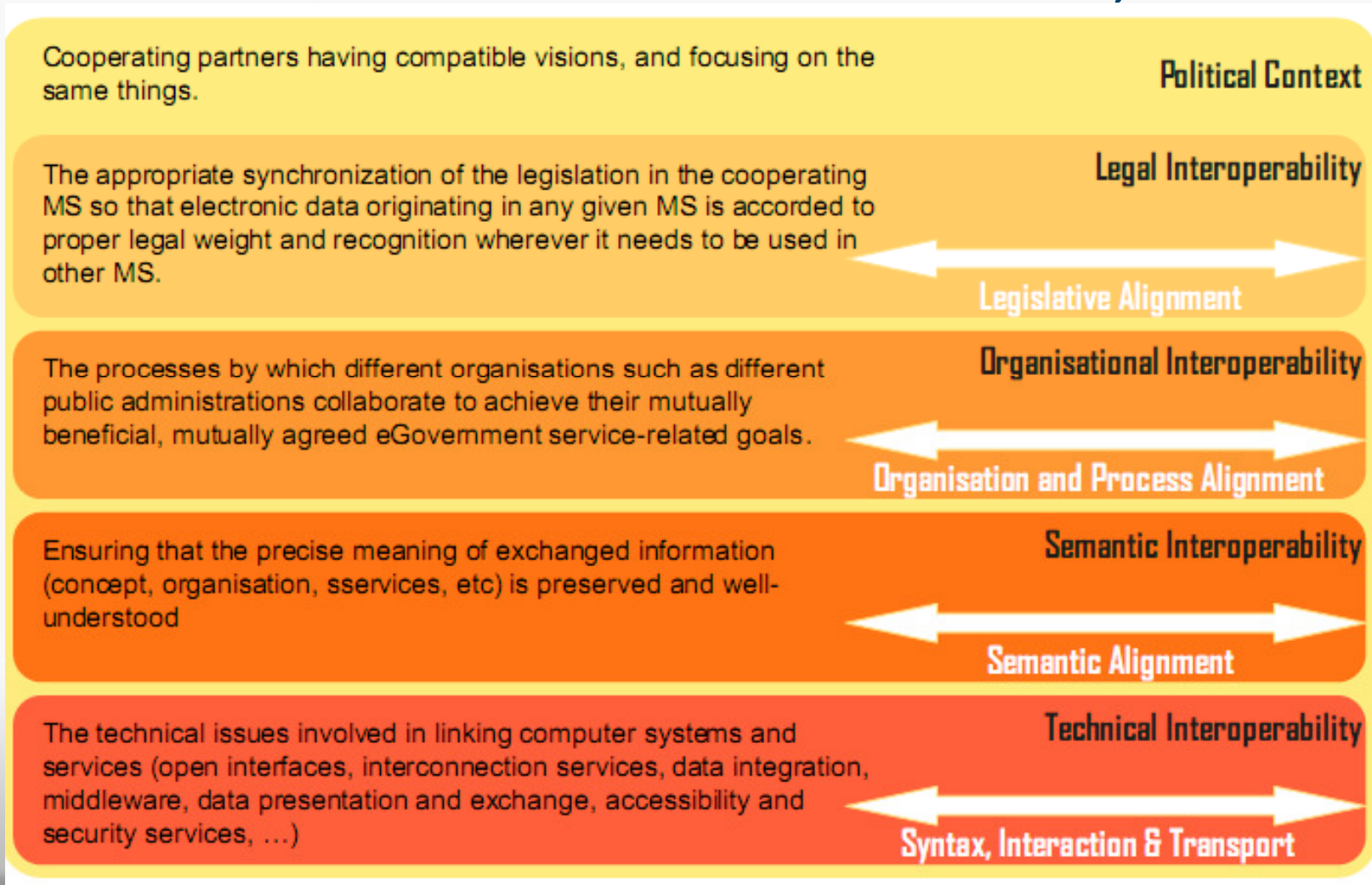
Librarians Do Gaga http://www.youtube.com/watch?v=a_uzUh1VT98

[Athenasbanquet](#) 3 videos

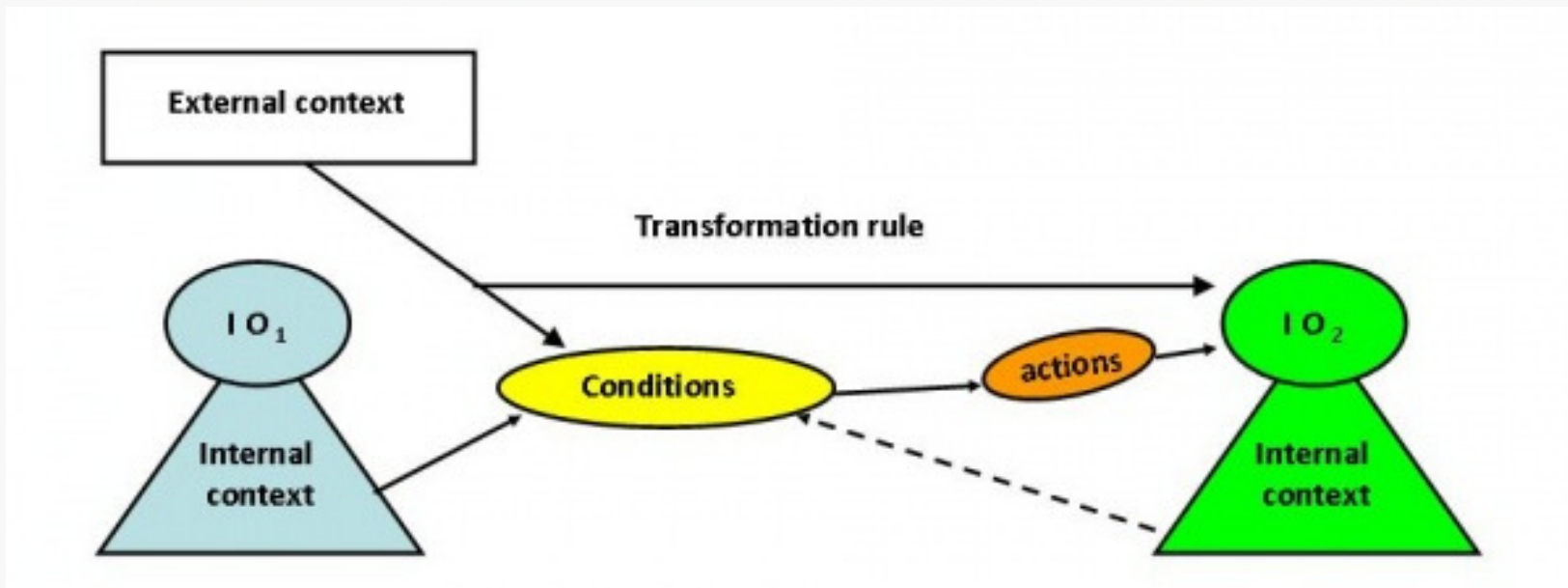


Cause when it comes to search if it's not
tough it isn't fun (fun)

European Interoperability Framework 2.0. EC, 2008



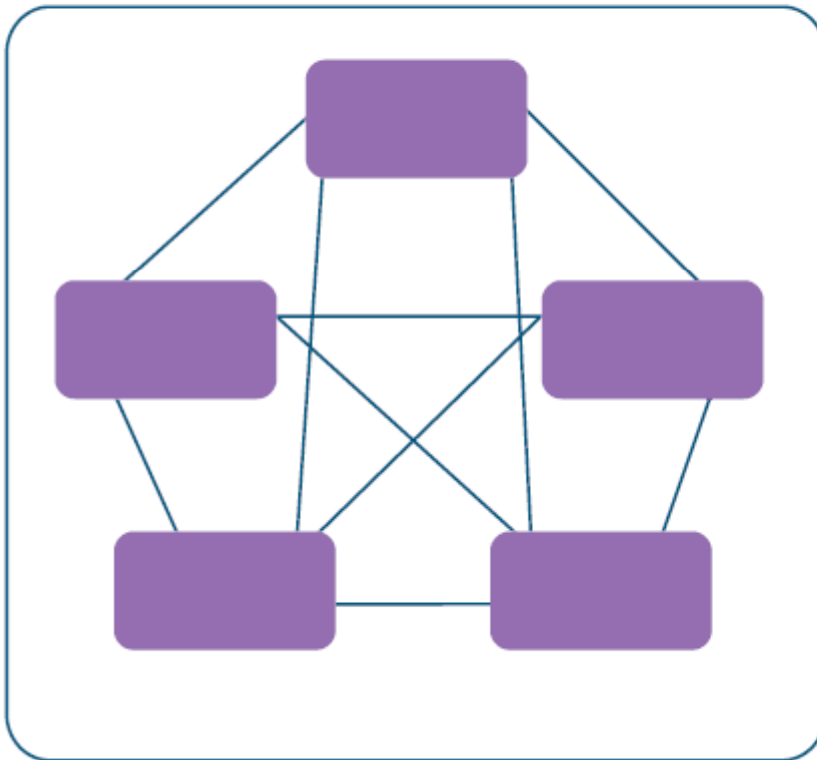
Which workflow for interoperability? (1)



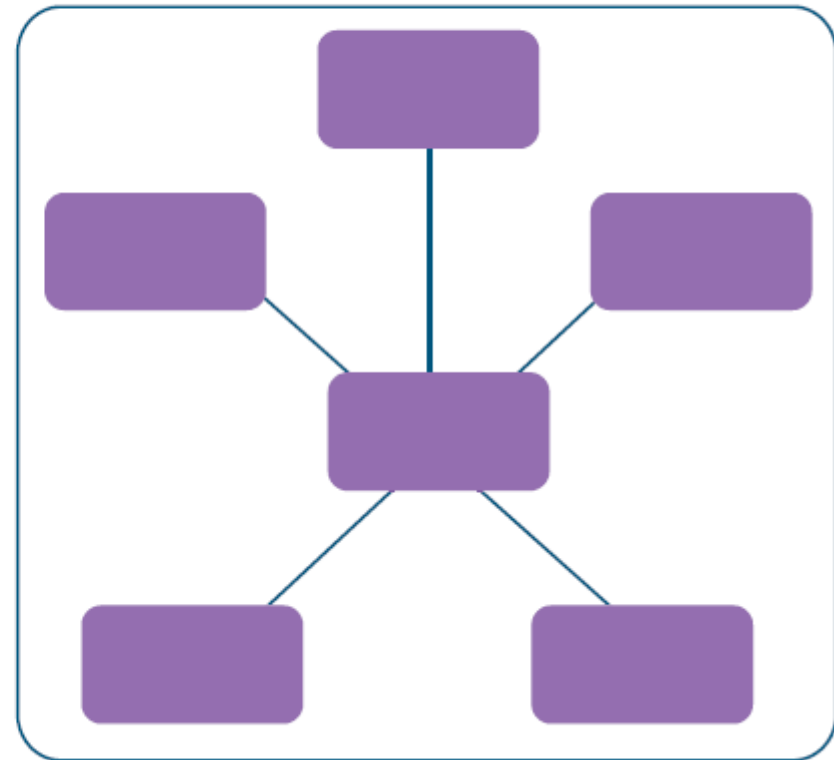
Draft of possible interoperability workflow,
Rome DL.ORG Mtg Dec 2009

Which workflow for interoperability? (2)

Bilateral Solutions



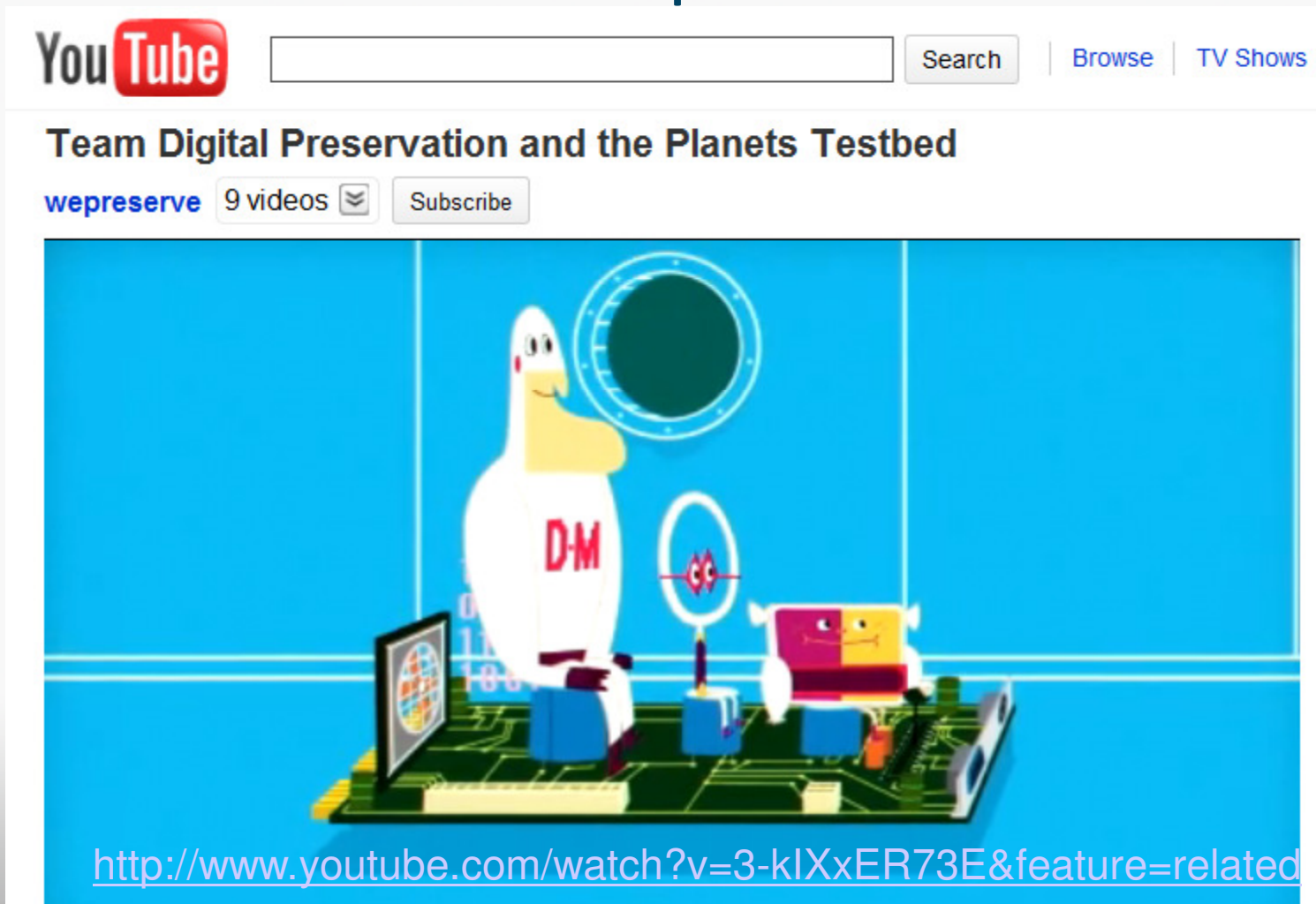
Multilateral Solutions



DELOS Digital Library Reference Model

- DL Ref Model makes a nod to preservation (Section II.3.2)
- It makes a nod to interoperability in terms of for example Section II.3.1 and C156, Interoperability Support
- It does not make a nod to the importance of preservation in terms of interoperability
- We look at content preservation – where content is data/docs

Interoperability and preservation



The image shows a screenshot of a YouTube video player. At the top left is the YouTube logo. To its right is a search bar with the word "Search" below it. Further right are links for "Browse" and "TV Shows". The video title is "Team Digital Preservation and the Planets Testbed". Below the title, it says "wepreserve" followed by "9 videos" and a dropdown arrow, and a "Subscribe" button. The video thumbnail shows a white cartoon bird wearing a white lab coat with "D-M" on it, standing on a green circuit board. Next to it is a small, colorful robot. In the background, there's a blue wall with a circular window showing a dark planet. The URL at the bottom of the video player is <http://www.youtube.com/watch?v=3-kIXxER73E&feature=related>.

Digital Libraries like all Objects Break

- **Inaccessibility of digital object**
 - Object becomes lost
 - Degradation of storage medium means content can not be read.
 - Technological obsolescence
- **Syntactical interpretation or representation failures**
- **Semantic opaqueness**
 - Lack of contextual information (e.g. suitable metadata)
 - Loss of Process & dynamic nature
- **Legal impediments**
- **The organisation and its staff**
 - Lack of organisational will – visible benefits
 - Decentralised and node-based organisation



© HATII UofGlasgow, 2005

Historic Media on Display at the Launch of the
UK Digital Curation Centre (DCC), November 04
<http://www.dcc.ac.uk>

High-level Preservation View

- *bit stream*
(01100101101010010)
- *information content* (e.g. images, sounds, text)
- *Context of Information* (e.g. linkages, interrelatedness)
- *Experience* (e.g., speed, layout, quality of display device, input device characteristics)



Objectives of digital longevity

- Digital preservation aims to ensure that future users will be able discover, retrieve, render, manipulate, interpret and use digital information in the face of constantly changing technology
 - It involves conservation, renewal, restoration, selection, destruction, enhancing, updating, and annotating
 - It is a risk management activity at all stages of the longevity pathway
 - It is about translating uncertainties into manageable risks
 - In the digital age we are all digital curators whether in our work, in our community or in our personal life
 - Digital Preservation is an ongoing activity.



© HATII UofGlasgow, 2005

Charles Dollar visits HATII, 2004

Preservation Risk is Actual

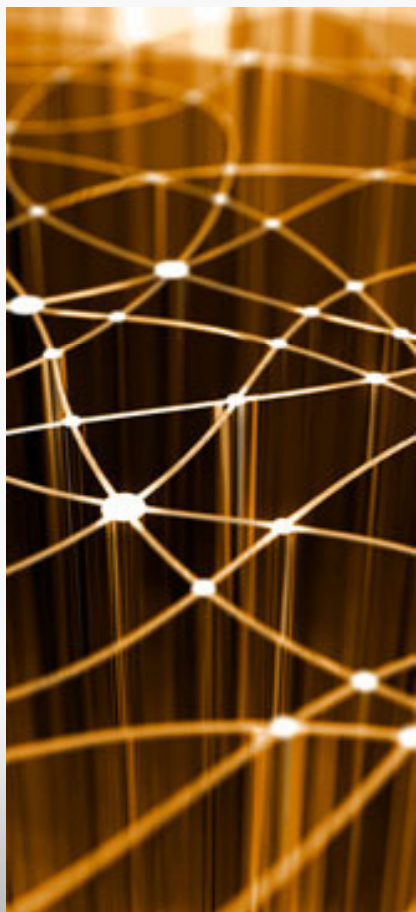
- It is technological.
- It is social.
- It is organisational.
- And it is cultural.
- Actual risks can be assessed and measured—
actual risks can be managed.

What needs to be considered

- Bit stream, information content, context, experience
- Syntactical
- Content
- Semantic
- Functionality
- Context

Seven Key Interoperability Issues

- *Process* – what is the boundary between static content, representations, linkages
- *Authenticity* – how do we (people and machines) know ‘it’ is authentic
- *Quality* – how do we measure quality and does it change overtime
- *Change over time* – how do we create ‘dynamic interoperability’ frameworks
- *Policy* – how do we reconcile policies in a contemporary context and how do we handle policy drift
- *Legal* – how can we address issues related to legal aspects
- *Preservation* – how do we preserve ‘interoperability potentiality’ what do we preserve.



Thank you