

# Connecting Preservation Planning and Plato with Digital Repository Interfaces

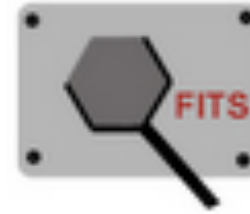
David Tarrant  
[davetaz@ecs.soton.ac.uk](mailto:davetaz@ecs.soton.ac.uk)



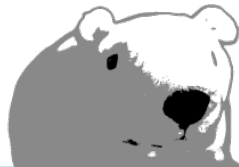
# The digital preservation landscape



DROID



DRAMBORA

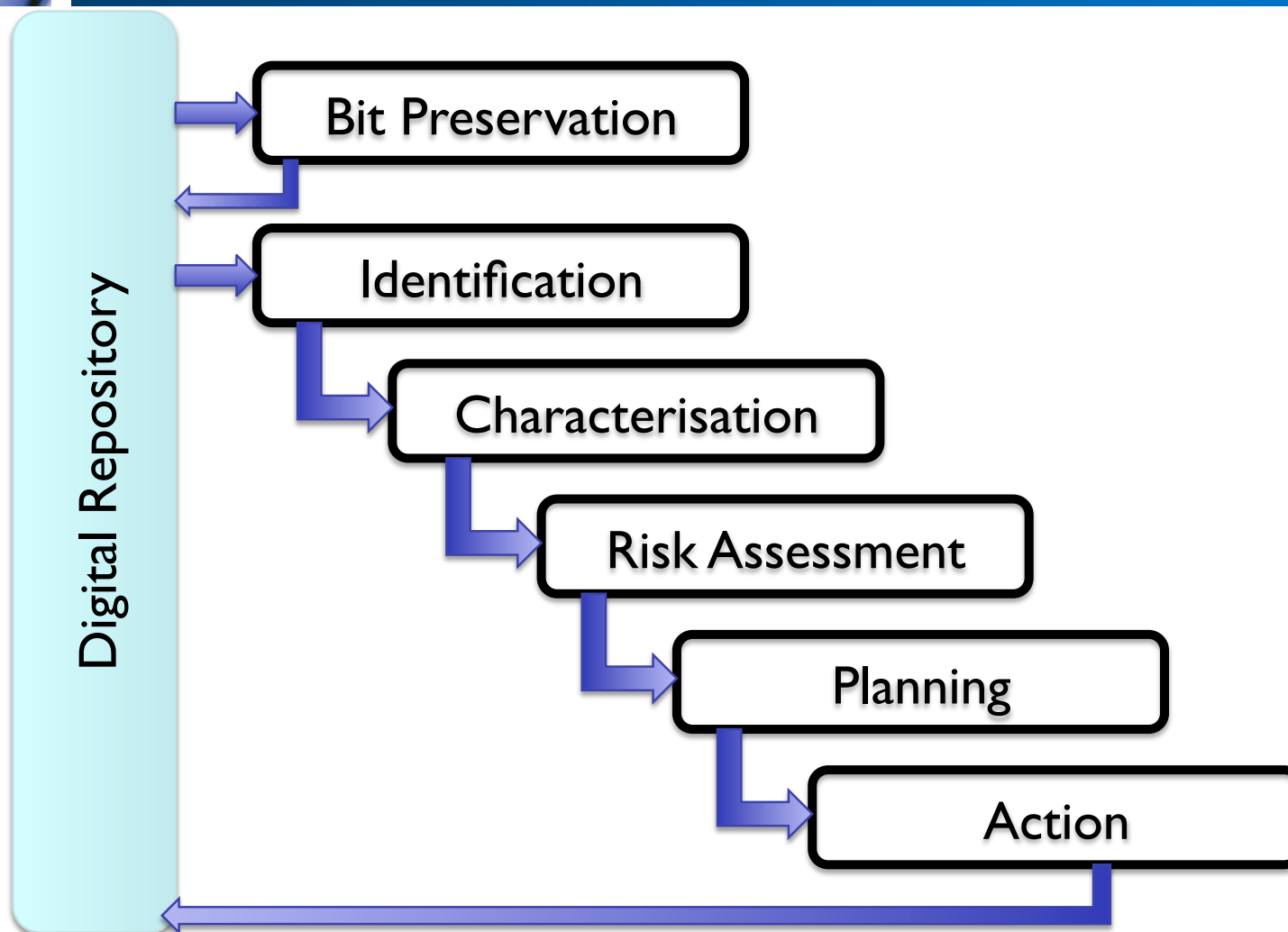


DAF





# Overview





# Interfaces...

*“75 million people already own iPod Touches and iPhones. That's all people who already know how to use the iPad.”*



Picture by curiouslee <http://www.flickr.com/photos/curiouslee/4320074421/>

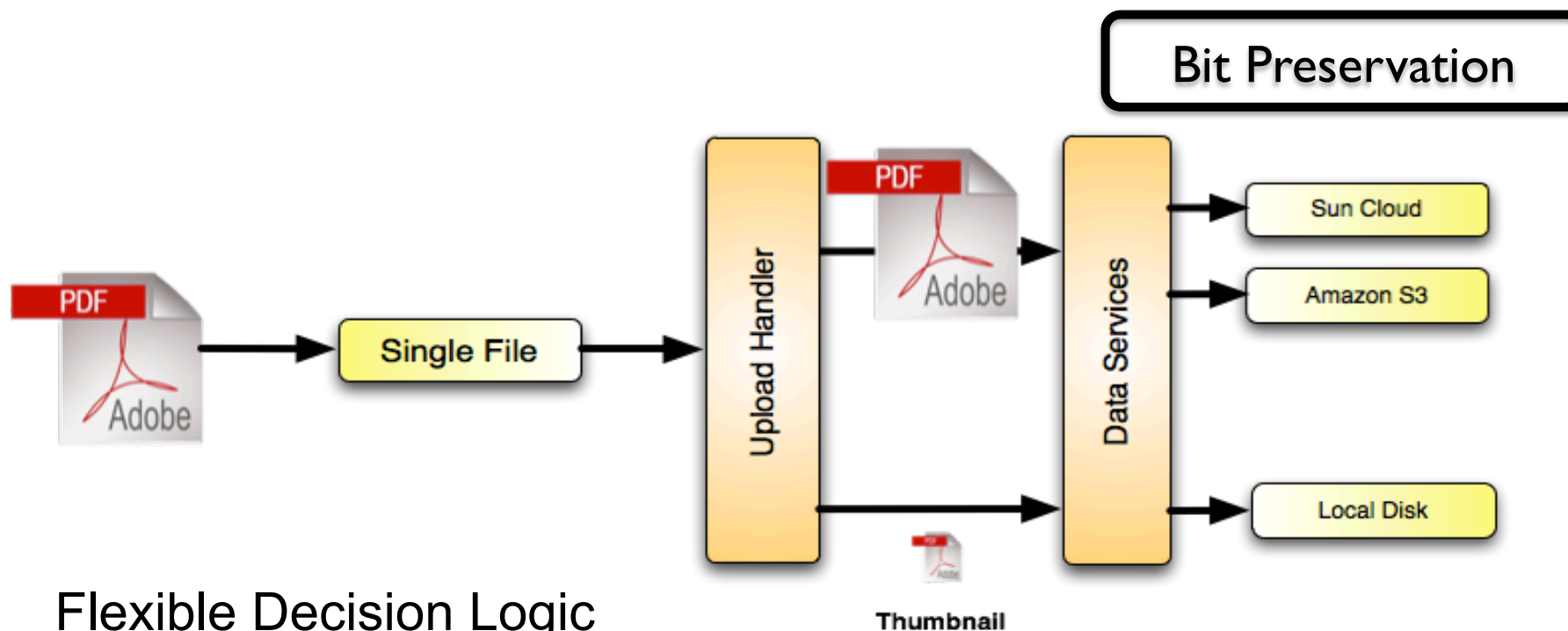


## Message 1

“Digital preservation can be simplified by integrating tools and services with existing digital repository interfaces #ipres2010”



# EPrints Storage Controller



Flexible Decision Logic

Multiple Copies

Periodic  
md5 sum checks

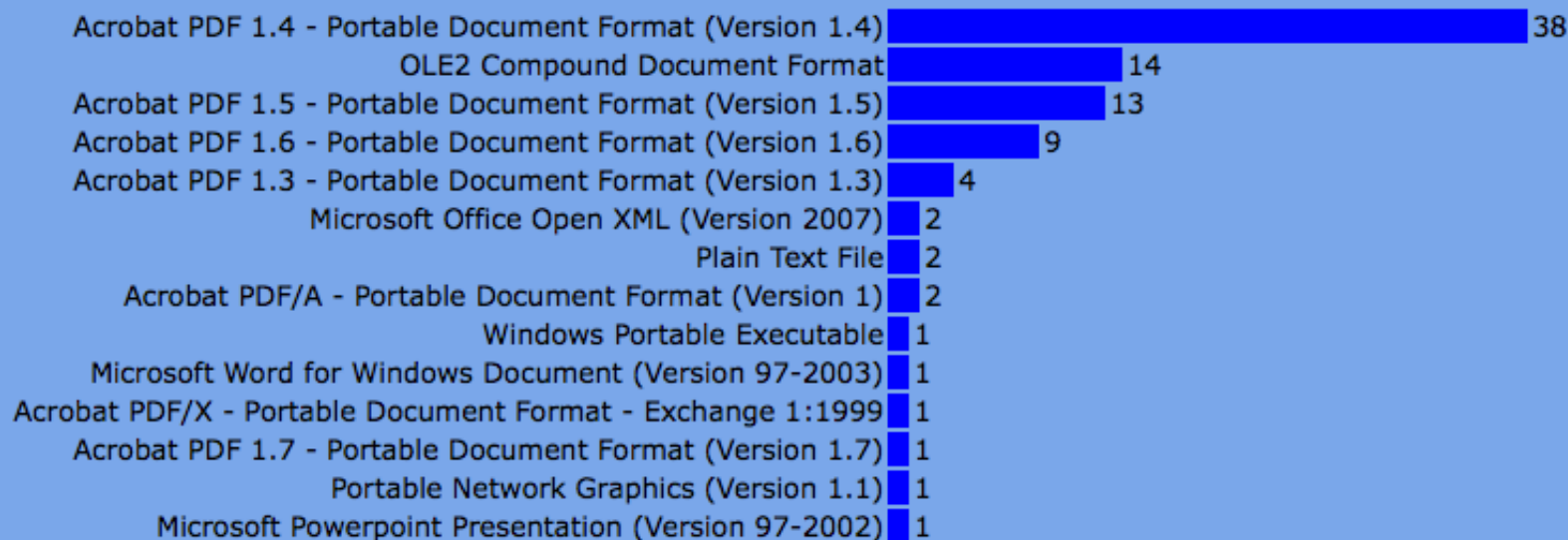
Multiple Services



# DROID in the Repository

## Identification

### Format Breakdown





# JHove in Plato

Done as part of Preservation Planning

Characterisation

[+] Sample Records

Description of sample records: The sample pages are taken from the 1965, 1971, 1977, and 1983 year books. The quality of these pages is quite similar and so is it for the rest of this part of the collection as it was scanned as one batch. However, the sample images were selected from different years in order to increase the probability for differences in the image quality, as the paper and/or

What are sample records and why do I need them?

**JHove Tree View**

- Maserati\_Karif.gif
  - Extraction date: 2010-03-11T21:19:42+01:00
  - Size: 81612
  - Mime-Type: image/gif
  - Format: GIF
  - Version: 89a
  - Status: Well-Formed and valid
  - Module
    - Profiles
      - GIF 89a
    - Properties
      - GIFMetadata
        - GraphicRenderingBlocks [1]
          - Blocks
            - LogicalScreenDescriptor
              - LogicalScreenWidth [732]
              - LogicalScreenHeight [502]
              - ColorResolution [7]
              - BackgroundColorIndex [1]
              - PixelAspectRatio: type: Short/ [0]
              - GlobalColorTableFlag [Global color table follows; background color index n
              - GlobalColorTableSortFlag [Not ordered]
              - GlobalColorTableSize: type: Short/ [3]

**Object Format**

PUID: fmt/4  
Name: Graphics Interchange Format  
Version: 1989a  
Mime-type: image/gif  
Extension: gif

Identify format

DROID  
Digital Object Identification

View Characteristics

**JHOVE**  
JHOVE (JHOVE) is a Java-based tool for the automatic identification and characterization of digital objects.

DL Description:

PUID: fmt/4  
Name: Graphics Interchange Format  
Version: 1989a  
Mime-type: image/gif  
Extension: gif





# Risk Analysis in EPrints

## Factual vs Subjective Information

Risk Assessment

Preserv 2



Home | About | Browse by Year | Browse by Subject |


Logged in as Mr David C Tarrant | [Manage deposits](#) | [Profile](#) | [Saved searches](#) | [Review](#) | [Admin](#) | [Logout](#)

### Formats/Risks



This EPrints install is referencing a trial version of the risk analysis service. None of the risk scores are likely to be accurate and thus should not be used as the basis for a program of action.

#### High Risk Objects

OLE2 Compound Document Format  1


#### Medium Risk Objects

Microsoft Powerpoint Presentation (Version 97-2002)  3

#### Low Risk Objects

Portable Document Format (Version 1.4)  3

Portable Document Format (Version 1.3)  2

ZIP Format  2



## Risk Assessment

### High Risk Objects

Graphics Interchange Format (Version 1987a) ☐ 9

<a href="#">Lada 1200 E.gif</a> (82Kb)	
Title: PLANETS GIF collection	
EPrint ID: <a href="#">21</a>	User: Unnamed user with email <a href="mailto:davetaz@ecs.soton.ac.uk">davetaz@ecs.soton.ac.uk</a>
<a href="#">Lamborghini Countach LP 500.gif</a> (76Kb)	
Title: PLANETS GIF collection	
EPrint ID: <a href="#">21</a>	User: Unnamed user with email <a href="mailto:davetaz@ecs.soton.ac.uk">davetaz@ecs.soton.ac.uk</a>
<a href="#">Matra Simca Rancho.gif</a> (79Kb)	
Title: PLANETS GIF collection	
EPrint ID: <a href="#">21</a>	User: Unnamed user with email <a href="mailto:davetaz@ecs.soton.ac.uk">davetaz@ecs.soton.ac.uk</a>
<a href="#">Morris Marina de luxe MK II.gif</a> (55Kb)	

User	No of Files
Unnamed user with email <a href="mailto:davetaz@ecs.soton.ac.uk">davetaz@ecs.soton.ac.uk</a>	9

### Preservation Actions

---

#### Download File SeLECTION

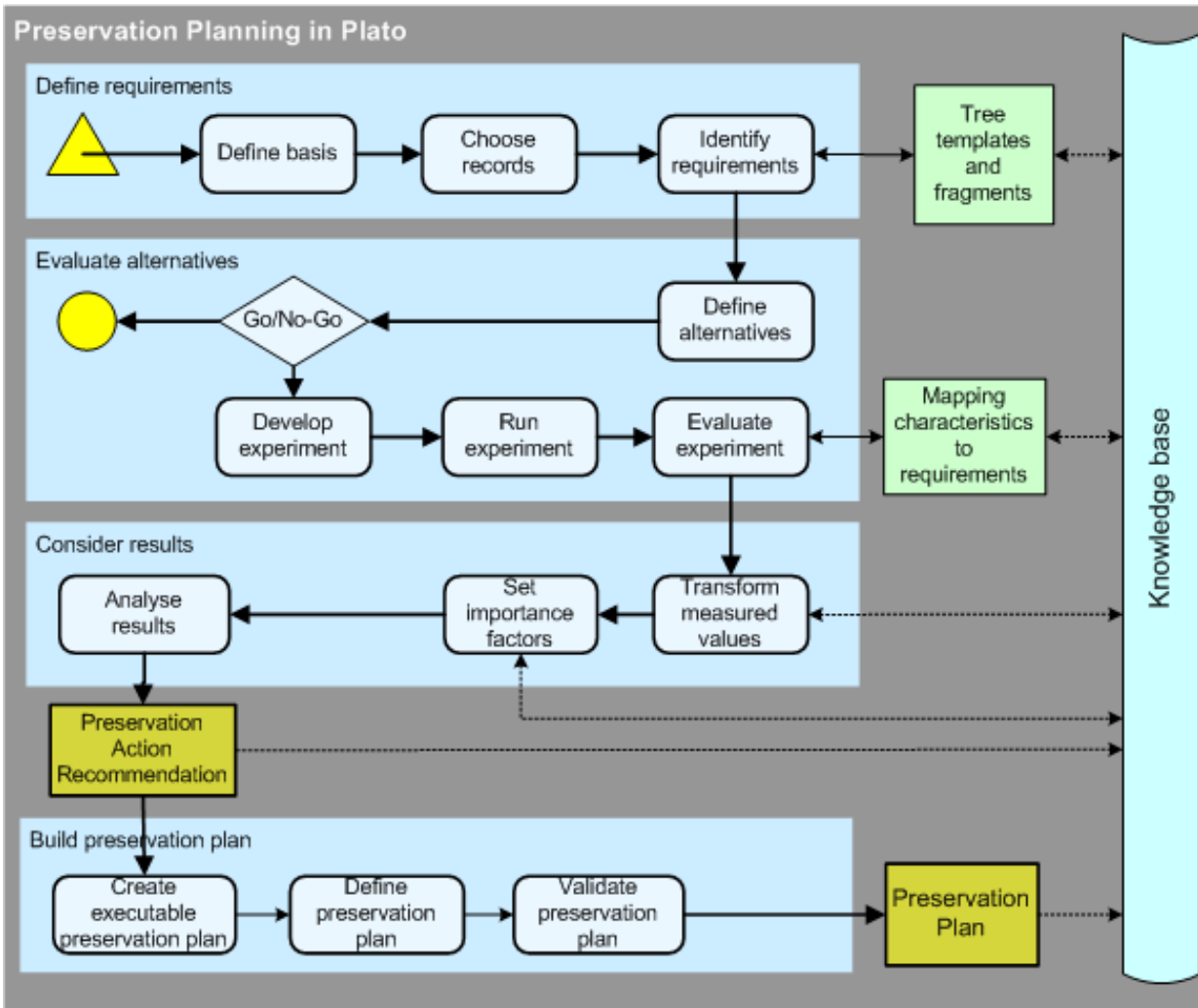
No. of Files:

---

#### Upload Preservation Plan



# The Planets Planning Workflow



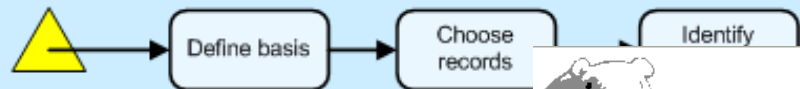
Planning



# Plato

## Preservation Planning in Plato

### Define requirements

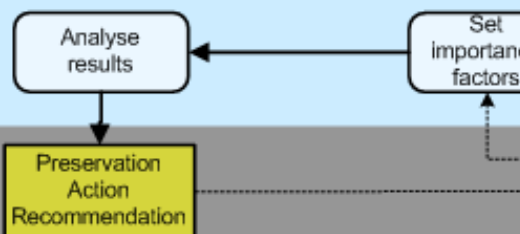


Tree templates

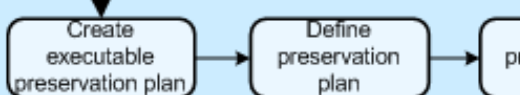
### Evaluate alternatives



### Consider results



### Build preservation plan



## Planning



## PLANETS Preservation Planning Tool (*Plato*)



[\[logout becker\]](#) [\[help\]](#)

Project | Define Requirements | Evaluate Requirements | Consider Results | PP4 workshop - The National Archive

### Identify Requirements

[Objective Tree](#)  
[Descriptive Information](#)

[How can I define the objective tree?](#)

#### [+] Objective Tree

[Expand All](#) | [Collapse All](#)  
[Website](#)

Focus	Node	Single	Scale	Restriction	Unit
	Website				
X	Record characteristics				
X	Technical characteristics				
X	Ubiquity		Ordinal	Ubiquitous/Widespread/Special	
X	Support		Positive Integer		number of tools
X	Documentation				
X	Stability				
X	Ease of identification		Ordinal	Automatic/Manual/None	
X	Ease of validation		Ordinal	Automatic/Manual/None	
X	Lossiness		Ordinal	Lossy/Lossless	
X	IPR		Boolean	Yes/No	
X	Complexity		Ordinal	High/Medium/Low	
	Event detectability			None/Detectable/Recoverable	

Release 1.1 - Institute of Software Technology and Interactive Systems: «off-ice bears»

Quick Access:





# The Action Plan

## ▲ A Small part of the Plato preservation plan

```
<eprintsPlan>  
  <tool>  
    <toolIdentifier uri="http://dbpedia.org/data/ImageMagick"/>  
    <parameters toolParameters="-verbose -compress None -quality 100 %INFILE% %OUTFILE%"/>  
    <targetFormat>PNG</targetFormat>  
  </tool>  
</eprintsPlan>
```



# Performing Actions

- Each set of “at risk” classified files can have a single related preservation plan.
- Once uploaded, any defined actions will be performed on **all** files of that classification.

Action

**Preservation Actions**

---

**Download File SeLECTION**

No. of Files:

---

**Upload Preservation Plan**

## Preservation Plan Upload Successful



Actions have been queued to be executed shortly and changes will be reflected below once completed. In order to view these changes please revisit or refresh this page later.



# Trust, Authenticity and Provenance

## Before



Image (GIF)

[Download \(76Kb\)](#) | [Preview](#)

## After



Image (PNG) (Migrated (Preservation) from Document ID: 41 (image/gif))

[Download \(76Kb\)](#) | [Preview](#)

- [Image \(GIF\)](#) (Original Version)



# The KeepIT Training Modules

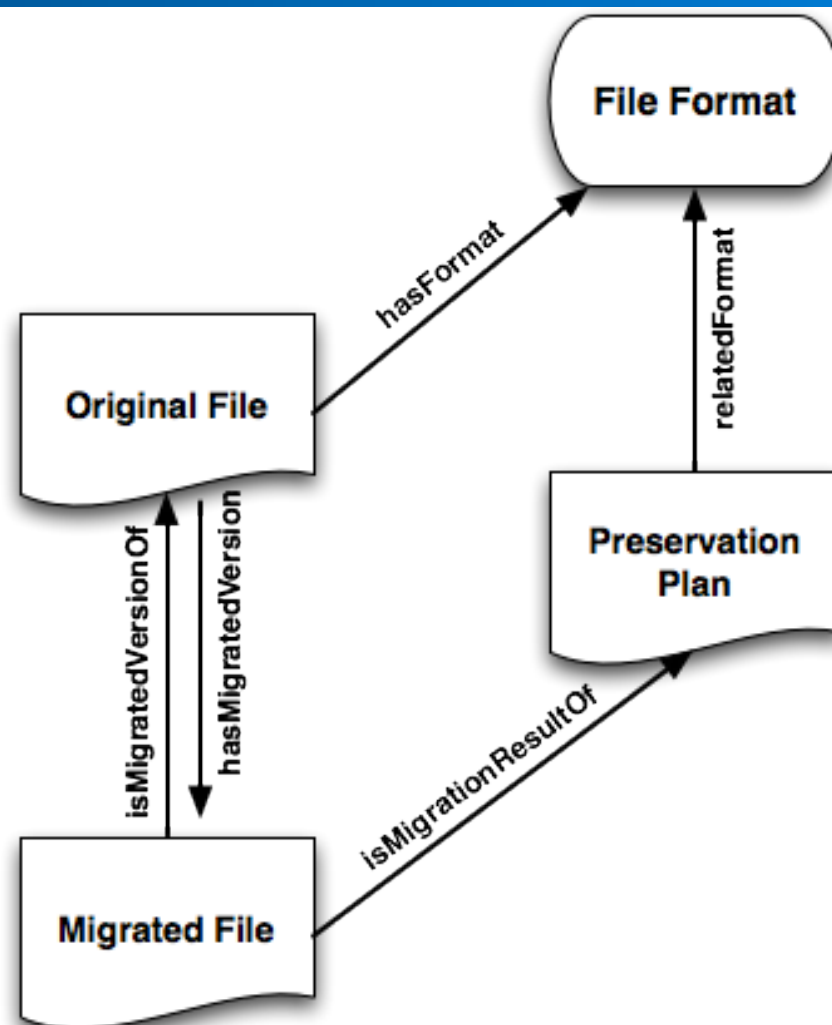
- **Module 1. Organisational issues**
  - Scoping, selection, assessment, institutional parameters
- **Module 2. Costs**
  - Lifecycle costs for managing digital objects, based on the LIFE approach, and institutional costs
- **Module 3. Description**
  - Describing content for preservation: provenance, significant properties and preservation metadata
- **Module 4. Preservation workflow tools**
  - Format management, risk assessment and storage, and linked to the Plato planning tool from Planets
- **Module 5. Trust**
  - The repository's approach to preservation; trust (by the repository) of the tools and services it chooses





# Provenance Information

- ▶ Open Provenance Model (OPM) compliant
- ▶ Stored in RDF triple form using the EPrints relation manager added in 3.2



OPM @ <http://openprovenance.org/>



# The KeepIT exemplars

Nectar

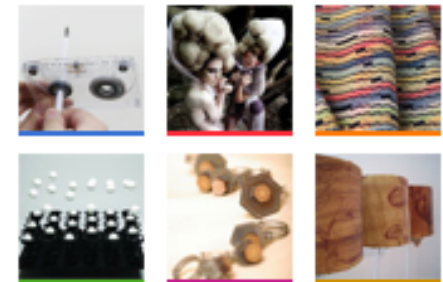


EdShare

UNIVERSITY OF  
Southampton



UAL Research  
Online



UNIVERSITY OF  
Southampton 



# UALResearch Online

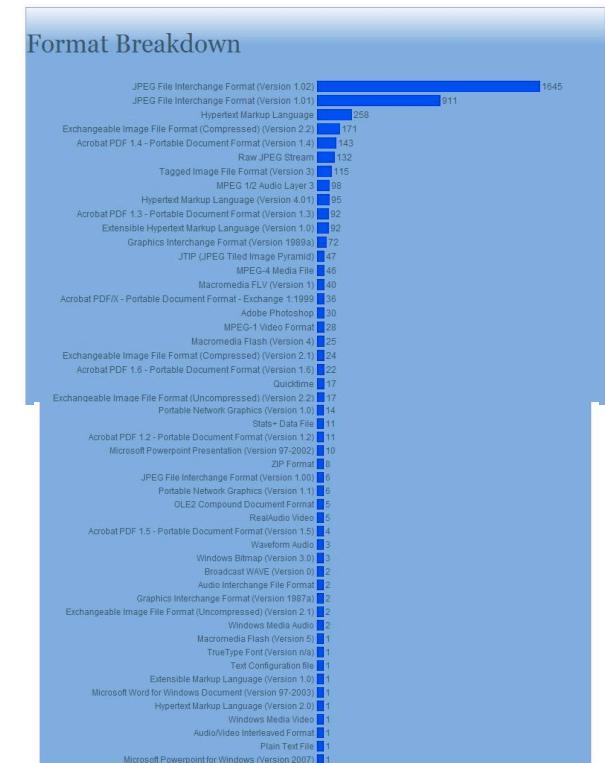
“Much more time-consuming and complicated”

“It's not just for the university/tech people to do  
it's part of the manager's job too”

“Somewhat less confident...but that's because  
I know what it really means now”

“I feel well-informed about what is available to  
help me”

“I would be quite confident teaching other  
managers, giving a talk etc. to share what I  
have learned”





# EdShare

“Participation in the KeepIt Project has provided me with an excellent practical grasp of the realities of preservation and brought an apparently enormous and daunting area into the realms of approachable and practical possibilities”

EdShare

UNIVERSITY OF  
Southampton



UNIVERSITY OF  
Southampton 



# Northampton University

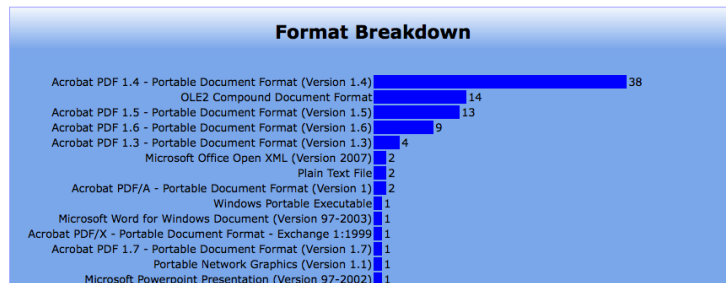


“The KeepIT project has shown me that Digital Preservation is not be the scary thing it can appear to be.”

“There isn't a quick fix.”

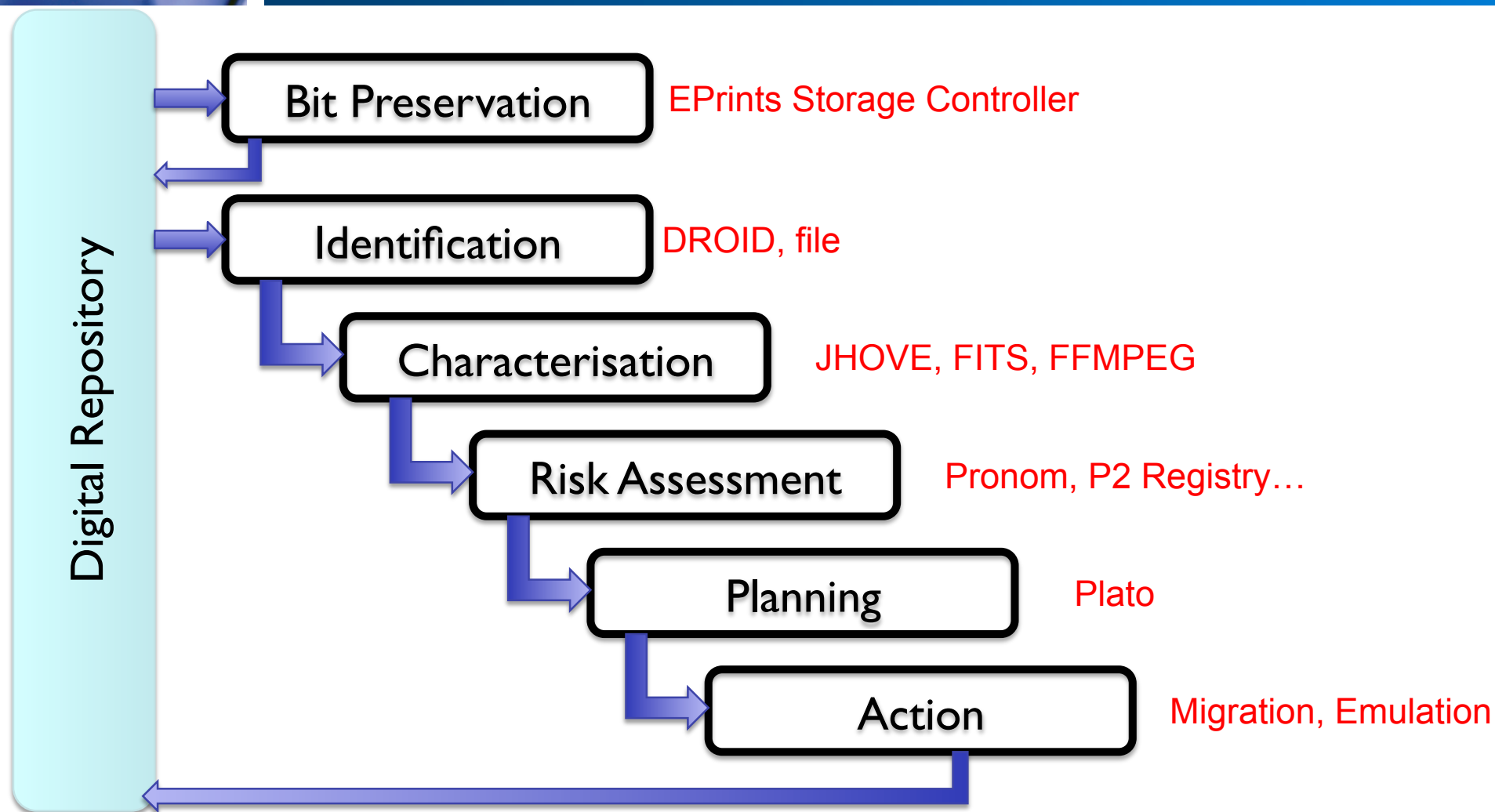
“The secret is to take small steps; to set preservation planning within the bigger institutional context; and to engage stakeholders through advocacy and training”

“My confidence has undoubtedly grown but the more I know, the more I realise there is to know.”





# Summary





## Message 2

“Can we pass XML between systems to enable breakthroughs in digital preservation?”





## Message 2

“Can we pass XML between systems to enable breakthroughs in digital preservation?”

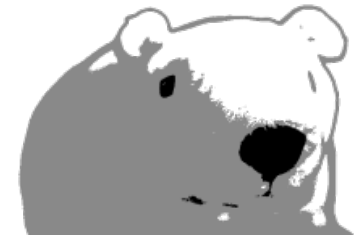






Thank You

JISC



UNIVERSITY OF  
Southampton eprints