Connecting the Dots

SCAPE and other projects

David Tarrant
Open Planets Foundation
davetaz@ecs.soton.ac.uk
Summary

• Top 5 things you should know from EU Projects. There are many more!

• Each of the 5 will be followed by a random slide. This is partly there to break up the items, but also because I lost a bet at iPres2012. These slides I have never seen before and I can only apologize for their content.

I blame Kara
The content of this slide is in no way associated with the Open Planets Foundation and may contain copyrighted or material of prior art.
Preservation Watch

- Alerting services for collection managers.
- Hunter gatherer, normalizer, processor, and intelligence for digital preservation decisions.
- Provides you the evidence.
How I feel after finding out we didn’t keep the originals

The content of this slide is in no way associated with the Open Planets Foundation and may contain copyrighted or material of prior art.
Preservation Workflows

- Publishing workflows to help you with common preservation tasks and problems.

NEXT TIME

LISTEN WHEN I TELL YOU NOT TO KEEP ALL YOUR BACKUPS IN 1 PLACE

The content of this slide is in no way associated with the Open Planets Foundation and may contain copyrighted or material of prior art.
3 The Results Evaluation Framework

File Identification for DOCX Files
10 Files of supposed Type – MS Word Document (Open XML) – identified using 65 versions of 5 tools

Microsoft Word for Windows
Microsoft Office Open XML
ZIP Format
Microsoft Word
Microsoft Office Open XML – Word

http://data.openplanetsfoundation.org/ref/

www.openplanetsfoundation.org
Random Slide #4

PRESERVES YOUR VIDEO

ON DVD AND THROWS AWAY THE ORIGINAL

The content of this slide is in no way associated with the Open Planets Foundation and may contain copyrighted or material of prior art.
The SPRUCE/Hackathon Wiki

Datasets

These are the Datasets or collections that relate to specific preservation issues. They are categorized by their media type.

Audio datasets

- User-generated audio field recordings
- Audio Collection (York)
- LAVC audio
- Endangered Archives Programme (EAP)

Disk image datasets

- Disk Images
- Realistic Disk Image Collection for Research and Education

Issues

These are the preservation or other business-driven issues that relate to specific Datasets. There may be one or more specific Solutions developed to solve them. The aim of an issue page is to provide a detailed description of preservation challenge and requirements of the issue owner that will help to inform development of a Solution that solves the issue.

Click this link to create a new Issue, then edit the italicised text.

Unsolved issues

Issues that do not have linked solutions. Why not suggest or contribute a solution?

- Classification of files within a disk image
- Extraction of keywords (and images) from large collections of text-based files
- Checking significant properties of images have been retained after migration
- Correct File Formats
- Checking significant properties of documents have been retained after migration

Solutions

These are Solutions that address specific issues encountered in particular Datasets. Solutions are typically quite specific to a particular issue and Dataset but may have a wider application. For details of tools utilized in a Solution, either follow links from individual Solution pages or see the Tool Registry.

Click this link to create a new Solution, then edit the italicised text.

Appraisal and assessment solutions

Solutions for assessing or appraising datasets.

- Analysis of Lucene Index Word Frequency
- Characterising Externally Generated Content

Bit rot detection and repair solutions

Solutions for detecting and possibly repairing bit rot issues in Datasets.

- Corrupted JPEG and JPEG2000 files solution

www.openplanetsfoundation.org
The content of this slide is in no way associated with the Open Planets Foundation and may contain copyrighted or material of prior art.
The content of this slide is in no way associated with the Open Planets Foundation and may contain copyrighted or material of prior art.
An OPF Goal

To help develop, sustain and find permanent homes for services critical to our members.*

* The services have to be sustainable in the first place!

www.openplanetsfoundation.org
Thank-You

David Tarrant
Open Planets Foundation
davetaz@ecs.soton.ac.uk