

Experiencing the Testbed

Matthew Barr
HATII at the University of Glasgow



How to Proceed 1

- ❑ In this practical session you and your group should work through the six-step experiment process for one of the listed experiment scenarios
- ❑ You should work together as a group using one laptop, but if you wish to view the Testbed on other laptops in your group you may log in with the same user details as used on your main laptop
- ❑ Login details will be given on the handout



How to Proceed 2

- ❑ Before you begin your experiment, please take about 10 minutes to discuss and make a note of the sorts of results you would expect to receive for your chosen experiment and the sorts of metrics you would use to analyse the results



How to Proceed 3

As you progress through the stages you should:

- ❑ Discuss the decisions you reach and make a note of the reasoning behind your decisions
- ❑ Try to supply as much relevant information as possible at each stage, such as the purpose, focus and scope of your experiment in Stage 1
- ❑ To help us improve the Testbed, please make a note of aspects you like or dislike and any improvements or additions you would like to see
- ❑ If you would like to send us feedback after the event email: helpdesktb@planets-project.eu



How to Proceed 4

- ❑ You will have 50 minutes to work through your experiment scenario. If you manage to complete the scenario with time to spare please choose another scenario to work through
- ❑ Following on from the session, one member of your team should then present your findings to the wider group. You will be given the opportunity of showing your experiment and discussing your experience



Reporting Back

Each group should cover the following in their report back:

- Scenario selected
- Expected results and potential metrics you would use
- Metrics applied
- Overview of experiment outcomes
- Impressions of using the Testbed



Experiment Input Files

- ❑ Unless otherwise specified in the scenario description you should browse through the Testbed Corpora for appropriate input files for your experiment. Using the Testbed Data Browser, the Corpora can be found here: **[/planets-ftp/corpus/](#)**
- ❑ When choosing files for your experiment you should open them to see if you consider them to be appropriate for your current experiment.



Scenario 1: File format Identification

- ❑ **Services:** JHOVE and DROID
- ❑ **Scenario:** Your organisation has been given a CD containing a selection of files in a variety of formats. You have been given the job of identifying the formats of these files and you would like to compare how effective JHOVE and DROID are at this task
- ❑ **How to Proceed:**
 - Perform an identification experiment using JHOVE
 - Select a number of files of different formats from the corpora (choose 5 or 6 files)
 - In order to check the identified Pronom IDs go to this page:
<http://www.nationalarchives.gov.uk/PRONOM/PUID/proPUIDSearch.aspx?status=new>
 - After evaluating your experiment use the 'Save As...' option to create a new experiment and tailor it for the DROID service
 - Compare the outcomes of the second experiment to the first



Scenario 2: Migration of Word files to PDF

- ❑ **Services:** NZ Metadata Extractor, XenaOO
- ❑ **Scenario:** Your organisation would like to be able to batch migrate its extensive archive of Microsoft Word files to PDF and it would like to be able to automatically extract metadata from the input and output files during this process
- ❑ **How to Proceed:**
 - Perform a migration experiment using the NZ metadata extractor and the Xena Open Office service
 - Select 5-6 Word files from the corpora (look in the DOC directory)
 - Discuss the outcomes of the experiment process, the measurements that were made and your evaluation of the results
 - If you have time re-run the experiment with PDF/A as an output file rather than PDF1.4 and compare the results
 - Open the input and output files on your PC and consider how you might manually evaluate the migration



Scenario 3: Migration of HTML to XHTML

❑ **Services: JTidy**

❑ **Scenario:** Your organisation has inherited a number of old static websites from projects that are no longer active. The websites were encoded in HTML 4.01 and some elements are no longer displaying correctly in modern browsers. It is your job to test the effectiveness of a tool that can automatically migrate HTML 4.01 to XHTML

❑ **How to proceed:**

- Perform a migration experiment using the JTidy service
- Consider whether you would also like to characterise the input and output files
- Select 1-5 HTML 4.01 files from the corpora (look in the HTML directory) or find appropriate pages online and attach them to your experiment
- Discuss what possible manually measured properties you might wish to focus on
- Download the input and output files to your PC and run them through the W3C validator (http://validator.w3.org/#validate_by_upload) to check how effective the migration process has been
- Open the input and output files in your browser to manually compare them



Scenario 4: Migration of SVG to PDF / JPG

- ❑ **Services:** DROID, Inkscape, Imagemagick
- ❑ **Scenario:** Your organisation hosts a digital art repository that contains a number of vector based images in SVG format. As SVG support isn't guaranteed in web browsers you would like to explore alternative formats for these files. You would also like to ensure the tools you are using are producing valid output files
- ❑ **How to Proceed:**
 - Perform a migration experiment using the Inkscape service and DROID to characterise the input and output files
 - Find and select SVG files from the following location in the data browser: /planets-ftp/ftp/matt/svg/
 - Consider what properties should be manually measured for this experiment
 - Open the input and the output files on your computer and manually compare them
 - After evaluating your experiment if there is still time use the 'Save As...' option to edit your experiment, selecting Imagemagick instead of Inkscape and JPG instead of PDF
 - Compare the outcomes of this second experiment to those from the first



Scenario 5: Migration of GIF to PNG

- ❑ **Services:** GIMP, XCDL Extractor
- ❑ **Scenario:** Although the patent for the GIF compression algorithm has now expired, your organisation is considering migrating its extensive collection of GIF files to an alternative format. It is your job to identify the pros and cons of possible formats such as PNG when combined with certain image tools. You intend to automatically extract significant properties from the files in order to compare the applicability of different formats.
- ❑ **How to Proceed:**
 - Perform a migration experiment using the GIMP service and the XCDL Extractor
 - Select about 5 or 6 GIF files from the corpora (look in the GIF directory)
 - Consider how effective you consider the automatically extracted properties to be as a means of assessing the applicability of PNG and the GIMP service
 - Consider what manually measured properties you think would be useful for further analysis
 - Open the input and output files to manually compare them
 - After evaluating your experiment and if there is still time use the 'Save as...' option to repeat your experiment with a different output format and compare the results



Accessing the Testbed

- ❑ To access the Testbed load the following URL into your browser:
 - **<https://testbed.planets-project.eu/testbed/>**
- ❑ (be sure to enter **https** not just **http**)
- ❑ Your team will be assigned a number and you will be able to log into the Testbed with the following credentials:
 - Username: *****
 - Password: *****

