PREMIS PRESERVATION METADATA MAINTENANCE ACTIVITY

OBJECTS EXERCISE - BACKGROUND

A preservation repository has ingested an article submitted by a faculty member. The article takes the form of a PDF file, and is accompanied by a Dublin Core record (see below), as well as a Submission Agreement (see below). The repository's preservation strategy is to store the bits as submitted.

Other relevant information:

- All files submitted to the repository include a message digest calculated by the digital content management system used throughout the institution, known to the repository by the designation "LocalDCMS".
- The repository will only preserve the document itself; it makes no attempt to preserve links to resources external to the document. Therefore, the only attribute considered significant for preservation is "Internal content only".
- Files submitted to the repository are password-protected, and can be accessed with the username/password key "guest/guest".
- The repository uses the PRONOM format registry for detailed format-specific information. The unique PRONOM ID for PDF 1.5 is fmt/19.
- The repository only records one recommended environment per file. For PDF, the environment is Adobe Reader 6.0. The recommended memory requirement for this program is 64 MB; 32 MB are the minimum requirement.
- The namespace for all repository-supplied identifiers is "LocalRepository".

The following Dublin Core record was included with the article:

```
Record Number: dcms222333
dc.Title: "DPC Technology Watch Report on Preservation
Metadata"
dc.Creator: Brian Lavoie
dc.Creator: Richard Gartner
dc.Subject: digital preservation
dc.Subject: metadata
dc.Publisher: Digital Preservation Coalition
dc.Date: 2005-09 [publication]
dc.Type: Text
dc.Format: PDF 1.5
dc.Identifier: dpctw05-01.pdf
dc.Language: eng
dc.Rights: Public
```

Here is an excerpt from the Submission Agreement associated with the article:

```
Transaction Number: 00445578
```

```
Date of Submission: 2006-07-04
Submitter: Professor Melvil Dewey
SubmitterID: 1870
File Name: dpctw05-01.pdf
Size: 209,000 bytes
Technical Information: created in Adobe Acrobat 5.0
```

Given the information above, complete Objects Exercise Part 1 by filling in values for the semantic units listed in the template. In most cases, values for semantic units can be taken directly from the information provided above. Otherwise, use your general knowledge to create reasonable metadata values.

Let's make the previous exercise slightly more complicated. Suppose that in addition to the PDF file described above, the submitter also includes a Microsoft Excel spreadsheet file, containing a data set that was published along with the article text. We now have two files: the PDF file and the spreadsheet. The spreadsheet can be accessed by clicking on a link within the PDF file. This is considered an internal link (i.e., it points to another document within the repository), so the repository will preserve its actionability.

Relevant information:

- The PRONOM ID for Excel 2003 is fmt/62.
- The spreadsheet file will be stored and accessed as "read-only".
- Memory requirements for Excel 2003 are 128 MB.

Here are the Dublin Core record and Submission Agreement information associated with the spreadsheet:

```
Record Number: dcms222334
     dc.Title: "Data to Accompany Technology Watch Report"
     dc.Creator: Brian Lavoie
     dc.Subject: digital preservation
     dc.Subject: metadata
     dc.Publisher: Digital Preservation Coalition
     dc.Date: 2005-09 [publication]
     dc.Type: Dataset
     dc.Format: Excel [Office Professional Edition 2003]
     dc.Identifier: 050908lavoie.xls
     dc.Language: eng
     dc.Rights: Public
```

Here is an excerpt from the Submission Agreement associated with the presentation:

```
     Transaction Number: 00445579
     Date of Submission: 2006-07-04
     Submitter: Professor Melvil Dewey
```

```
Submitter ID: 1870
File Name: 050908lavoie.xls
Size: 146,000 bytes
Technical Information: created in Microsoft Excel,
part of Office Professional Edition 2003
```

Given the information above, complete Objects Exercise Part 2A by filling in values for the semantic units listed in the template. In most cases, values for semantic units can be taken directly from the information provided above. Otherwise, use your general knowledge to create reasonable metadata values.

Make sure you record the relationship between the spreadsheet and the PDF file. Also, go back and record this relationship in the metadata for the PDF file (see worksheet for Exercise 1).

Taken together, the PDF file and spreadsheet comprise a distinct *representation* of the article. It is the policy of the repository to record preservation metadata at the file and representation levels.

Using the information from the previous two examples, complete Objects Exercise 2B by creating preservation metadata for the article representation, consisting of the PDF and spreadsheet files.

Make sure you record the relationship between the representation and the PDF file, and the relationship between the representation and the spreadsheet file. Also, go back and record these relationships in the metadata for the PDF file (see worksheet for Exercise 1) and the spreadsheet file (see worksheet for Exercise 2A).