Unified Digital Format Registry (UDFR)

User’s Guide

Version 1.0.0 – 2012-06-04

UC Curation Center
California Digital Library
University of California
Copyright © 2012, The Regents of the University of California
All rights reserved

This document is licensed under the Create Commons Attribution-Share Alike 3.0 United States license. To view a copy of this license, visit http://creativecommons.org/licenses/by-sa/3.0/us/ or request a copy from Creative Commons, 171 Second Street, Suite 300, San Francisco, California, 94105, USA.

Acknowledgements

The Unified Digital Format Registry (UDFR) is a product of the University of California Curation Center (UC3), http://www.cdlib.org/uc3, at the California Digital Library (CDL). The project was funded by the Library of Congress as part of its National Digital Information Infrastructure Preservation Program (NDIIPP).

The UDFR project team gratefully acknowledges the support of the UK National Archives (TNA) and the Agile Knowledge Engineering and Semantic Web (AKSW) research group at the University of Leipzig.

UC3 project team
Stephen Abrams
Lisa Dawn Colvin
John Kunze
Margaret Low
Mark Reyes
Abhishek Salve
Marisa Strong

Library of Congress
Martha Anderson
Leslie Johnston

UK National Archives
Tim Gollins
Tracey Powell
Spencer Ross

AKSW, Univ. Leipzig
Philipp Frisch
Norman Heino
Sebastian Tramp

Version history

<table>
<thead>
<tr>
<th>Version</th>
<th>Date</th>
<th>Note</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.0.0</td>
<td>2012-06-04</td>
<td>Initial publication</td>
</tr>
</tbody>
</table>
Contents

Acknowledgements ........................................................................................................ 1
Version history ................................................................................................................. 1
Contents .......................................................................................................................... 2
Introduction ....................................................................................................................... 4
1  Getting started ............................................................................................................. 5
1.2  User roles .................................................................................................................. 5
1.3  User interface ............................................................................................................ 5
1.3.1  OntoWiki pane...................................................................................................... 6
1.3.1.1  User menu ........................................................................................................ 7
1.3.1.2  Extras menu ..................................................................................................... 9
1.3.1.3  Help menu ....................................................................................................... 10
1.3.1.4  Debug menu .................................................................................................. 10
1.3.2  Knowledge Base pane ....................................................................................... 11
1.3.3  Navigation pane .................................................................................................. 12
1.4.4  Login pane ........................................................................................................... 12
1.3.5  Workspace pane ............................................................................................... 14
1.3.6  Contextual menu ............................................................................................... 14
2  Initial data loads ......................................................................................................... 18
2.1  MIME media types ................................................................................................. 18
2.2  PRONOM ................................................................................................................ 18
3  Search and browse .................................................................................................... 20
3.1  Browse .................................................................................................................... 20
3.2  Search ..................................................................................................................... 20
4  Format representation information ............................................................................ 21
4.1  Review status .......................................................................................................... 22
4.1.1  Reviewing properties ....................................................................................... 22
4.2  Provenance history ............................................................................................... 23
4.3  Community comments ......................................................................................... 24
4.4  RDF source ............................................................................................................. 24
Introduction

A deep understanding of digital formats is necessary to support the long-term preservation of digital assets, as it facilitates the preservation of the information content of those assets, rather than just their bit stream representations. A format is the set of syntactic and semantic rules that govern the mapping between information and the bits that represent that information. The Unified Digital Format Registry (UDFR) is a new open source, semantically-enabled platform for the collection, long-term management, and dissemination of the significant properties of formats of interest to the preservation community [5]. The UDFR builds upon and “unifies” the function and holdings of two existing registry solutions: PRONOM, from the UK National Archives [13]; and GDFR (Global Digital Format Registry), from Harvard University [8]. While these services rely on older relational database technology, the UDFR uses a semantic database in which all information is represented in RDF form [23] and exposed as Linked Data [11]. Use of the UDFR is open to the public, although contribution or editing of information requires prior self-service account registration.

The UDFR was developed by the University of California Curation Center (UC3) [6] at the California Digital Library (CDL) [3] with funding from the Library of Congress as part of its National Digital Information Infrastructure Preservation Program (NDIIPP) [10].
1 Getting started

The UDFR portal is found at http://udfr.org (see Figure 1). This page provides links to the registry itself at http://udfr.org/registry, the UDFR OWL ontology file at http://udfr.org/onto/onto.rdf, and the online UDFR documentation at http://udfr.org/docs.

1.2 User roles

The UDFR defines four distinct user roles:

- **Consumer.** An anonymous read-only user of the UDFR.

- **Contributor.** A registered user of the UDFR permitted to add and edit format representation information.

- **Reviewer.** A contributor who is additionally permitted to review contributor-supplied information for technical veracity.

- **Administrator.** A UDFR system administrator, with all privileges.

This document provides a general introduction to the UDFR features available to consumers and contributors.

1.3 User interface

The general layout of the UDFR user interface, http://udfr.org/registry is shown in Figures 2 and 3. There are five main screen areas, or panes.
• **OntoWiki pane.** The OntoWiki pane (uppermost-left, labeled “O” in the Figures) contains menu options of general utility throughout the use of UDFR. (OntoWiki is the underlying semantic wiki platform on which the UDFR is based [19]. For more implementation details, see Appendix B.)

• **Knowledge Base pane.** The Knowledge Base pane (upper-left, “K”) offers a choice of knowledge bases (KBs), or RDF models [23], for selection.

**NOTE**  The terms “knowledge base”, “model”, and “graph” are used synonymously within the semantic web community.

![Figure 2 – Initial UDFR user interface layout](image)

• **Navigation pane.** The Navigation pane (lower-left, “N”) provides an RDF class browser for navigating the UDFR format representation information. This pane is not displayed until a knowledge base is selected in the Knowledge Base pane as described in Section § 1.3.2, **Knowledge Base pane**.

• **Login pane.** The Login pane (bottom-left, “L”) lets contributors login to their user account. Potential contributors can also follow a link for new account registration. Once logged in, this pane is no longer displayed.

• **Workspace pane.** The Workspace pane (right, “W”) is function dependent. Its appearance will vary based on the operation being performed.

### 1.3.1 OntoWiki pane

The OntoWiki pane contains menu options of general utility throughout the use the UDFR. There are four main menus:

• **User.**
• **Extras.**
Once a knowledge base has been selected, as described in Section § 1.3.2, the OntoWiki pane also displays a search box.

![UDFR User Interface Layout](Figure3.png)

**Figure 3 – UDFR user interface layout following knowledge base selection**

### 1.3.1.1 User menu

The “User / Register New User” menu option permits contributors to register for user accounts (see Figure 4).

**NOTE** Account registration is necessary *only* for the contribution of new data or editing of existing data in the UDFR. No account registration is necessary for read-only access to the UDFR.

The details of user account registration are provided in Section § 1.3.4, *Login pane*. 

- *Help.*
- *Debug.*
Once a contributor is logged in, the “User” menu will also provide options for setting user preferences and logging out (see Figure 5).

The “User / Preferences” option permits a contributor to update his or her user account preferences, including email address and password (see Figure 6).

NOTE Although the username appears on the form, it is not editable.
1.3.1.2 Extras menu

The “Extras” menu offers three options:

- File Manager, for uploading sample files.
- Queries, for issuing SPAQRL queries [16].
- Select Language, for selecting the user interface language.

As advanced topics, file upload and SPARQL queries will be discussed later on in Sections § 5.2 and 5.2.

The “Extras / Select Language” menu option selects the language used in the UDFR user interface. Customized language is available for:

- Chinese.
- English.
- French.
- German.
- Russian.

Additional language customizations may become available in the future.

![Figure 7 – OntoWiki pane Extras menu](image_url)

Changing the user interface language does not affect the format representation information displayed by the UDFR, only the user interface. For example, selecting the German language would change the appearance of the OntoWiki pane as follows:
1.3.1.3 Help menu

The “Help” menu provides links to help information and documentation about the OntoWiki semantic wiki underlying the UDFR (see Figure 9).

1.3.1.4 Debug menu

The “Debug” menu provides options for clearing various internal caches and restarting a UDFR session, which clears all caches and logs out the user (see Figure 10).
1.3.2 Knowledge Base pane

The Knowledge Base permits the selection of the knowledge base (KB), or RDF *model*, as the focus for subsequent searches and edits (see Figure 11). The UDFR defines three main knowledge bases:

- **UDFR data.** The UDFR data knowledge base holds the public format representation information. To select this KB click on the “UDFR – Start Here” link. Hereinafter this KB will be referred to as the “data” KB.

- **UDFR ontology.** The UDFR ontology knowledge base defines the RDF ontology, or *schema*, for expressing format representation information. To select this KB click on the “http://udfr.org/onto#” link.

- **UDFR user profiles.** The UDFR user profile knowledge base holds public information about UDFR users, including name, title, and organizational affiliation. To select this KB click on the “http://udfr.org/profile/” link.
1.3.3 Navigation pane

The Navigation pane provides an ontology browser for navigating the UDFR format representation information (see Figure 12). This pane is not displayed until a knowledge base is selected as described in Section § 1.4.2, Knowledge Base pane.

<table>
<thead>
<tr>
<th>Navigate: Classes</th>
</tr>
</thead>
<tbody>
<tr>
<td>View</td>
</tr>
<tr>
<td>Abstract Base</td>
</tr>
<tr>
<td>Controlled Vocabulary</td>
</tr>
</tbody>
</table>

Figure 12 – Navigation pane for data knowledge base

The UDFR classes are defined in a hierarchical structure. The full hierarchical structure of the data knowledge base is shown in Appendix § C, UDFR class structure. In the Navigation pane the left and right arrow icons and ascend and descend one level of the class hierarchy, respectively. The smaller double left arrow icon returns to the root class.

Clicking on any class name will display all of the instances of that class in the Workspace pane (see Figure 13).

1.4.4 Login pane

The Login pane permits contributors to register for and login to their user accounts (see Figure 14).

**NOTE** Account registration and login are only necessary for the contribution of new, or editing of existing, representation information. Read-only access to all representation information can be performed anonymously.

To register for a user account, click the “Register” button and complete the form in the Workspace pane (see Figure 15). Form fields marked with a red asterisk * are required; all other fields are optional.
Figure 13 – Displaying class instances in the Workspace pane

Figure 14 – Login pane
NOTE By completing the user account registration form you agree to contribute format representation information to the UDFR under the terms of the Creative Commons Attribution license (CC-BY). That attribution is fulfilled by your further agreeing that your full name, title, institutional affiliation, and additional profile information and website URL, if specified, will be publicly displayed in the “History” tab for all representation information that you contribute or edit. This information is made public in furtherance of the legitimate UDFR business need to provide strong provenance information for all contributed content.

Passwords are never publicly displayed and are stored by the UDFR only in a hashed form that is opaque to UDFR administrators. Passwords can be reset with the “User / Preferences” menu option of the OntoWiki pane, as described in Section § 1.3.1.1, User menu.

A forgotten password can be recovered by completing the “Login” operation (see Figure 14) and clicking the “Forgot your password?” link (see Figure 16) and supplied a account username or email address (see Figure 17).

The registration form incorporates a Captcha that must be entered into the form in a case sensitive manner (see Figure 15). If the Captcha is too indistinct or ambiguous, a new one can be generated by clicking the “refresh” icon .

Once a contributor is logged in this pane is no longer displayed. Once logged in, a contributor can logout via the “User / Logout” or “Debug / Reset Session” menu options in the OntoWiki pane, as described in Section § 1.4.1.

1.3.5 Workspace pane

The Workspace pane holds the results of UDFR search and browse operations. The appearance of its displayed information is operation dependent. Many of these operations will result display an ordered list of results. The Workspace pane supports paged browsing for lists that cannot be accommodated on a single screen (see Figure 18).

Buttons at the bottom of the pane provide direct access to the first, previous, next, last page of results, as well as to an arbitrarily-numbered page. Additional buttons control the maximum number of results displayed on any given page (see Figure 19).

1.3.6 Contextual menu

Whenever mousing over any active link in the UDFR user interface, a contextual menu becomes available as indicated by the graph icon . Clicking the graph icon will
display a context-sensitive menu offering appropriate options.

Figure 15 – User registration form

Figure 16 – Bad password
Figure 17 – Password recovery

Figure 18 – File format list

9. 7-bit ASCII Text
   File Format

10. 8-bit ANSI Text
    File Format

Figure 19 – Paged results buttons
Figure 20 – Contextual menu
2 Initial data loads

The UDFR was initialized with data loads from two external data sources:

- MIME media types [9].
- PRONOM technical registry [13].

2.1 MIME media types

MIME types were loaded from [http://mediatypess.appspot.com/](http://mediatypess.appspot.com/) [12], which itself manages data “routinely scrapped form IANA [Internet Assigned Numbers Authority]”. The data export was taken on February 22, 2012, and encompassed the following types:

<table>
<thead>
<tr>
<th>Count</th>
<th>Media type</th>
</tr>
</thead>
<tbody>
<tr>
<td>809</td>
<td>application/*</td>
</tr>
<tr>
<td>125</td>
<td>audio/*</td>
</tr>
<tr>
<td>39</td>
<td>image/*</td>
</tr>
<tr>
<td>14</td>
<td>message/*</td>
</tr>
<tr>
<td>14</td>
<td>multipart/*</td>
</tr>
<tr>
<td>51</td>
<td>text/*</td>
</tr>
<tr>
<td>56</td>
<td>video/*</td>
</tr>
<tr>
<td>1,127</td>
<td></td>
</tr>
</tbody>
</table>

Table 1 – Initial MIME media types

2.2 PRONOM

PRONOM data were loaded from [http://nationalarchives.gov.uk/PRONOM/Default.aspx](http://nationalarchives.gov.uk/PRONOM/Default.aspx), the technical registry operated by the UK National Archives [13]. The data export was taken on February 23, 2012, and encompassed the following entities:

<table>
<thead>
<tr>
<th>Count</th>
<th>Resource</th>
</tr>
</thead>
<tbody>
<tr>
<td>846</td>
<td>File formats</td>
</tr>
<tr>
<td>28</td>
<td>Character encodings</td>
</tr>
<tr>
<td>17</td>
<td>Compression algorithms</td>
</tr>
<tr>
<td>1,237</td>
<td>Identifiers</td>
</tr>
<tr>
<td>1,006</td>
<td>External signatures</td>
</tr>
</tbody>
</table>
Table 2 – Initial PRONOM resources

The PRONOM data export is made freely available under the terms of the Open Government License (OGL), [http://www.nationalarchives.gov.uk/doc/open-government-licence/](http://www.nationalarchives.gov.uk/doc/open-government-licence/). Unless otherwise specified, all other contributed data all made freely available under the terms of the Creative Commons Attribution license (CC-BY), [http://creativecommons.org/licenses/by/3.0/](http://creativecommons.org/licenses/by/3.0/).

Additional data may have been contributed since the time of these initial data loads.
3  Search and browse

UDFR format representation information can be retrieved via search or browse.

3.1  Browse

Use the ontology browser in the Navigation pane to navigate to the class of interest, then click the class name to display all instances of that class in the Workspace pane, as described in Section § 1.3.3, Navigation pane (see Figure 13).

3.2  Search

All textual data in the UDFR is indexed for search. Enter terms in the search box in the OntoWiki pane and press the Return or Enter key to display matching search results in the Workspace pane (see Figure 21).

Figure 21 – Search results
4 Format representation information

As described in Section § 3, UDFR format representation information can be retrieved via search or browse. The unit of display in the UDFR is the resource, which corresponds to an RDF class (see Appendix C, UDFR class structure.). The Workspace pane display for a given format (or other resource) will be similar to that shown in Figure 22.

![Figure 22 – Representation information for JPEG JFIF 1.02](image)

The representation information takes the form of a list of property names and values associated with the resource in alphabetical order. Note that all of the property names, and many of the values, are actionable links. Clicking the link will display the properties pertaining to that property or value.

In addition to the property names and values, the following additional information is made available:

- **Review status.**
- **Provenance history.**
- **Community comments.**
- **RDF source.**
- **Instance linking.**
4.1 Review status

Every representation information property is associated with a review status set by UDFR users with the Reviewer role. Reviewed properties are indicated by the presence of a green checkmark icon before the property name (see Figure 23).

![Figure 23 – Review status](image)

4.1.1 Reviewing properties

UDFR users holding the Reviewer role will see active review checkboxes on all representation information displays (see Figure 24).

![Figure 24 – Review checkboxes enabled](image)

Once the appropriate boxes are checked, the review status can be updated by clicking the “Review” button.
4.2 Provenance history

UDFR tracks the full provenance of every representation information property. This information is made available via the “History” tab (see Figure 26).

---

4.2.1 Provenance history

UDFR tracks the full provenance of every representation information property. This information is made available via the “History” tab (see Figure 26).

**Figure 26 – History tab**

All actions related to properties are tracked:

- Adding, editing, and deleting.
- Reviewing.

The “History” display shows the type of action (and its underlying RDF form), the user who performed the action, and the timestamp of its performance. User names are actionable links. Clicking this link will display additional information about the user in a popup window so that an informed decision can be made regarding the veracity of the contributed information (see Figure 27).

**Figure 27 – User profile information**
4.3 Community comments

The “Community” tab displays comments contributed by register users holding the Contributor (or Reviewer) role (see Figure 28).

![Figure 28 – Community tab](image)

4.4 RDF source

The RDF underlying the displayed properties can be seen in the “Source” tab (see Figure 29).

![Figure 29 – Source tab](image)
4.5 Instance linking

The “Instance linking” sub-pane of the Workspace displays active links to other properties whose values are the property being shown. For example, Figure 30 displays the representation information properties defined for the C-Cube Microsystems corporate agent. The “Instances linking here” sub-pane to the right of the display indicates that C-Cube is the creator of the JPEG JFIF 1.00, 1.01, and 1.02 formats; as well as indicating the presence of a user contributed comment.

Figure 30 – Instance linking
5 User contribution

UDFR users holding the Contributor (or Reviewer) role are permitted to add, edit, and delete representation information for UDFR resources, such as File Format, Agent, etc. (Resources correspond to RDF classes. See Appendix C, UDFR class structure, for a comprehensive list of resources.)

5.1 Create a resource

Use the Navigation pane ontology browser to find the desired resource and then select the “Create Instance” option in the resource’s contextual menu (see Figure 31).

![Navigate: Classes](image)

**Figure 31 – Create instance menu option**

This will display a popup window with input fields for all properties defined for the resource, either directly or by inheritance from ontological super-class relationships (see Figure 32).

UDFR properties are of three types:

- **Text strings**.
- **Closed enumerations**.
- **Open enumerations**.

Text properties can be entered directly in the form text box. Closed enumerations have a fixed set of values defined in the UDFR ontology. Clicking on the “Select Predefined Value” link for property, for example, the “byteOrderType” shown in Figure 33, will
Figure 32 – Create instance popup window

Figure 34 – Byte order type closed enumeration
display a popup window with all of the enumeration values. Once a value is selected, its RDF URI will be automatically displayed in the form box.

Open enumerations can be selected from a pre-defined list in similar fashion, or if the appropriate value is not available, entered directly in to the form box (see Figure 35) and clicking the “Create New Value” link (Figures 36 and 37).

Figure 35 – Open enumeration, step 1

"ICC.1:2010-12" has been saved successfully with the identifier: http://udfr.org/udfr/u1r9247

Figure 36 – Open enumeration, step 2
Additional instances of repeatable properties can be defined by clicking the “+” icon to the right of the property form box (see Figures 38 and 39).

5.2 Editing a resource

Existing resources can be edited by selected the “Edit” button on the resource display (see Figure 40). The behavior of the input boxes is same as described in Section § 5.1, Creating a resource (see Figure 41).

Editing can also be initiated via the “Edit Resource” option on the contextual menu for a resource in a results list (see Figure 42).
### Properties of ICC color profile

#### Resource

<table>
<thead>
<tr>
<th>Properties</th>
<th>History</th>
<th>Community</th>
<th>Source</th>
</tr>
</thead>
<tbody>
<tr>
<td>Review</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Alias identifier</td>
<td>ICC:1:2018-12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Byte order</td>
<td>Big endian</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Composition</td>
<td>Unitary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Creation date</td>
<td>2010-12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Description</td>
<td>Cross-platform profile format for the creation and interpretation of colour data. Such profiles can be used to translate between different colour encodings, and to transform colour data created using one device into another device's native colour encoding. The acceptance of this format by application and operating system vendors allows end users to transparently move profiles, and images with embedded profiles, between different systems. For example, this allows a printer manufacturer to create a single profile for multiple applications and operating systems.</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disclosure</td>
<td>Full-disclosure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Form</td>
<td>Binary</td>
<td></td>
<td></td>
</tr>
<tr>
<td>MIME type</td>
<td>application/vnd.iccprofile</td>
<td></td>
<td></td>
</tr>
<tr>
<td>release date</td>
<td>2010-12</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Status</td>
<td>Active</td>
<td></td>
<td></td>
</tr>
<tr>
<td>UDRF Local Identifier String</td>
<td>u1982</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nIf Type</td>
<td>File Format</td>
<td></td>
<td></td>
</tr>
<tr>
<td>nIf Label</td>
<td>ICC color profile</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Figure 40 – Resource edit button**

**Figure 41 – Resource edit display**
5.3 Deleting a resource

Existing resources can be deleted by selecting the “Delete” button on the resource display (see Figure 43). Deletion can also be initiated via the “Delete Resource” option on the contextual menu for a resource in a results list (see Figure 44).

5.4 Deleting a property

Individual properties can be deleted using the “−” icon to the right of the property form box in the edit resource display (see Figures 45 and 46).
Figure 43 – Resource delete button

Figure 44 – Delete resource menu option
<table>
<thead>
<tr>
<th>Properties of ICC color profile</th>
<th>Resource</th>
</tr>
</thead>
<tbody>
<tr>
<td>Properties</td>
<td>History</td>
</tr>
<tr>
<td></td>
<td>Community</td>
</tr>
<tr>
<td></td>
<td>Source</td>
</tr>
</tbody>
</table>

**Review**

- **Alias identifier**
  - http://udfr.org/udfr/u1r9247
  - Select Predefined Value
  - Create New Value

- **Byte order**
  - http://udfr.org/onto#BigEndian
  - Select Predefined Value

- **Composition**
  - http://udfr.org/onto#Unitary
  - Select Predefined Value

- **Creation date**
  - 2010-12

- **Description**
  - Cross-platform profile
  - Format for the creation and interpretation of colour data. Such profiles...
6 Advanced features

This section describes a number of advanced UDFR features.

- File upload.
- SPARQL queries [16].

6.1 File upload

The OntoWiki pane “Extras / File Manager” menu option enables the upload of reference files that can be associated with particular formats (see Figure 23).

![Figure 23 – OntoWiki pane file manager option](image)

The file manager display shows all uploaded files by resource URI and MIME type (see Figure 24).

![Figure 24 – File manager](image)

Clicking the “Upload” button displays the upload form (see Figure 25). The “Resource URI” field should not be modified; UDFR will automatically assign the appropriate URI to the uploaded file.

The file can be associated with a format by selecting the “Edit Resource” contextual menu option for the format resource, clicking the “Select Predefined Value” link for the “File” property, and selecting the file from the pop-up window (see Figures 26 and 27).
6.2 SPARQL queries

Direct SPARQL queries against the underlying RDF in the Virtuoso quadstore [14] can be performed via two mechanisms.

- *SPARQL endpoint*, for scripted queries by automated agents.

An explanation of the SPARQL query language is beyond the scope of this document. Many comprehensive online tutorials are available.

6.2.1 Query form

The OntoWiki pane “Extras / Queries” menu option enables an Working pane form for submitting SPARQL queries (see Figures 28 and 29).
Figure 27 – Associating a file with a format

Figure 28 – OntoWiki pane query option
Inline results are displayed at the bottom of the Workspace pane (see Figure 30).

When formulating the SPARQL query, note that a number of fundamental namespaces are pre-defined, including:

- `rdf` : http://www.w3.org/1999/02/22-rdf-syntax-ns#
- `rdfs` : http://www.w3.org/2000/01/rdf-schema#
- `udfrs` : http://udfr.org/onto#

The delivery mode and format of the query results can be specified via the radio buttons in the upper right hand corner of the display (see Figures 31 – 33).

- Inline (default) or downloadable file.
- Plain text (default), XML, or JSON.

### 6.2.2 SPARQL endpoint

A SPARQL endpoint is a web service that conforms to the SPARQL protocol [16]. UDFR presents a read-only endpoint at [http://udfr.org/ontowiki/sparql/](http://udfr.org/ontowiki/sparql/).  

---

**Figure 29 – SPARQL query form**
Figure 30 – Query results

Figure 31 – Query results output format
Figure 32 – Query results in XML

```
<?xml version="1.0"?>
<sparql xmlns="http://www.w3.org/2005/sparql-results#">

  <head>
    <variable name="uri" />
    <variable name="format" />
    <variable name="udfr" />
  </head>

  <results>
    <result>
      <binding name="uri">
        <uri>http://udfr.org/udfr/ulf46</uri>
      </binding>
      <binding name="format">
        <literal>BROADCAST WAVES version 0</literal>
      </binding>
      <binding name="udfr">
        <literal>ulf46</literal>
      </binding>
    </result>
  </results>

</sparql>
```

Figure 33 – Query results in JSON

```
{
  "head": { "link": [], "vars": ["uri", "format", "udfr"] },
  "results": { "distinct": false, "ordered": true, "bindings": [}
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf46" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf47" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf48" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf49" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf50" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf51" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf52" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf53" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf54" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf55" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf56" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf57" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf58" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf59" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf60" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf61" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf62" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf63" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf64" }, "format": { "type": "l" },
  [ "uri": { "type": "uri", "value": "http://udfr.org/udfr/ulf65" }, "format": { "type": "l" }
```
Note that unlike the use of the SPARQL query form, in which the most important namespaces are defined, as described in Section § 6.2.1, all namespaces must be explicitly defined in the queries sent to the endpoint.

Assuming the SPARQL query:

```sparql
PREFIX rdf: <http://www.w3.org/1999/02/22-rdf-syntax-ns#>
PREFIX rdfs: <http://www.w3.org/2000/01/rdf-schema#>
PREFIX udfrs: <http://udfr.org/onto#>

SELECT ?uri ?format ?udfr
WHERE {
  ?uri rdf:type       udfrs:FileFormat ;
      rdfs:label ?format ;
      udfrs:udfrIdentifier ?udfr .
}
```

in the file “example-query.sparql”, the query can be sent to the UDFR endpoint via the “curl” command:

```
% curl --data-urlencode query@example-query.sparql \  
  http://udfr.org/ontowiki/sparql/
```

```
<?xml version="1.0"?>
<sparql xmlns="http://www.w3.org/2005/sparql-results#">
  <head>
    <variable name="uri" />
    <variable name="format" />
    <variable name="udfr" />
  </head>

  <results>
    <result>
      <binding name="uri">
        <uri>http://udfr.org/udfr/u1f46</uri>
      </binding>
      <binding name="format">
        <literal>Broadcast WAVE, version 0</literal>
      </binding>
      <binding name="udfr">
        <literal>u1f46</literal>
      </binding>
    </result>
  </results>
</sparql>
```
Appendix A     RDF and the Linked Data

All format representation information in the UDFR is represented in RDF (Resource Description Framework) form [23]. The fundamental RDF data structure is the *triple*, an assertion of the form:

\[ \text{subject} \quad \text{predicate} \quad \text{object} \]

where subjects and predicates are represented by URIs (Uniform Resource Identifiers) and objects are represented by URIs or typed literals. For example:

\[
\text{udfrs:u1f46} \quad \text{rdfs:type} \quad \text{udfrs:FileFormat}.
\]
\[
\text{udfrs:u1f46} \quad \text{udfrs:udfrIdentifier} \quad \text{“u1f46”}.
\]
\[
\text{udfrs:u1f46} \quad \text{rdfs:label} \quad \text{“Broadcast WAVE, version 0”}.
\]

which collectively assert that the RDF resource represented by “udfrs:u1f46” is a UDFR file format, that it has the UDFR identifier “u1f46”, and that it represents the format, Broadcast WAVE, version 0.

Note that RDF has a number of different serialization formats. The example above is given in Turtle (Terse RDF Triple Language). Other equivalent forms are possible.

Linked Data is “a recommended best practice for exposing, sharing, and connecting pieces of data, information, and knowledge on the Semantic Web using URIs and RDF” [11].
Appendix B  UDFR implementation

The UDFR is implemented by a stack of open source products, primarily the OntoWiki semantic wiki, http://ontowiki.net/Projects/OntoWiki, developed by the Agile Knowledge Engineering and Semantic Web (AKSW) research group at the University of Leipzig, http://aksw.org/. The other main components of the technology stack include:

- **RDFauthor**. RDFa JavaScript library, http://aksw.org/Projects/RDFauthor [20].
- **Erfurt**. RDF API, http://aksw.org/Projects/Erfurt [18].
Appendix C    UDFR class structure

The two main class hierarchies of the http://udfr.org/onto# (udfrs) ontology are rooted at “Abstract Base” and “Controlled Vocabularies”. The “Abstract Base” hierarchy defines the main properties used to express UDFR format representation information (see Table 3).

Abstract Base >
    Abstract Product >
        Abstract Format >
            Compression
            Encoding
            File Format
        Document
        Hardware
        Media
        Software
    Abstract Signature >
        External Signature
        Internal Signature
    Agent
    Assessment
    Byte Sequence
    Digest
    File
    Grammar
    Holding
    IPR
    Identifier
    MIME
    Process

Table 3 – Abstract Base class hierarchy

The “Controlled Vocabulary” hierarchy defines controlled enumerations used as the values of various representation information properties (see Table 4).
Controlled Vocabulary >
  Administrative Status Type
  Agent Type
  Assessment Type
  Availability Type
  Basis Facet Type
  Byte Order Type
  Composition Facet Type
  Compression Lossiness Type
  Constraint Facet Type
  Country Code
  Digest Type
  Disclosure Type
  Document Intent Type
  Document Type
  Domain Facet Type
  File Purpose Type
  File Type
  Form Facet Type
  Format Role Type
  Genre Facet Type
  Grammar Type
  Hardware Type
  IPR Type
  Identifier Namespace Type
  Individual Agent Role Type
  Language Code Type
  Legal Jurisdiction Code Type
  Media Access Type
  Media Type
  Media Write Type
  Process Type
  Programming Language Type
  Role Facet Type
  Signature Obligation Type
  Signature Position Type
  Signature Type
  Software Type
  Transform Facet Type

Table 4 – Controlled Vocabulary class hierarchy
References


<http://www.w3.org/TeamSubmission/turtle/>.


[9] Internet Assigned Numbers Authority, *MIME Media Types* 


[11] *Linked Data – Connect Distributed Data across the Web* 
<http://linkeddata.org/>.


[13] National Archives [UK], *PRONOM* 

[14] Open Link Software, *Virtuoso RDF* 


[17] University of Leipzig, *Agile Knowledge Engineering and Semantic Web (AKSW)*


