27 November 2012
Status: Draft
Author: Jean-Claude Dauphin

JISIS and Web Technologies

I. Introduction

This document does aspire to explain how J-ISIS is related to Web technologies and how to use J-ISIS formats (CDS/ISIS formatting language) for creating HTML/XHTML display formats that contain hypertext. It's important to realize that the **J-ISIS application has an embedded Web Browser** that is used for displaying the display formats, **as well as a Web server which is used for serving static web resources** that can be html pages or any document file.

I.1 J-ISIS Web Browser

The new J-ISIS Web browser component is based on the JavaFX embedded browser, a user interface component that provides a web viewer and full browsing functionality through its API. The JavaFX embedded browser component is based on WebKit, an open source web browser engine. It supports Cascading Style Sheets (CSS), JavaScript, Document Object Model (DOM), and HTML5.

Supported Features of HTML5

The current implementation of the WebView component supports the following HTML5 features:

- Canvas
- Media Playback
- Form controls (except for `<input type="color">`)
- Editable content
- History maintenance
- Support for the `<meter>` and `<progress>` tags.
- Support for the `<details>` and `<summary>` tags.
- DOM
- SVG
- Support for domain names written in national languages

I.2 J-ISIS Web Server

J-ISIS contains now an embedded Jetty http server listening on port 8585, thus whenever J-ISIS is running on a machine, you have the J-ISIS database server listening on port 1111

and also a Web Server listening on port 8585. The document root of the Web server is defined by the first DEF_HOME value defined in the *dbhome.conf* file.

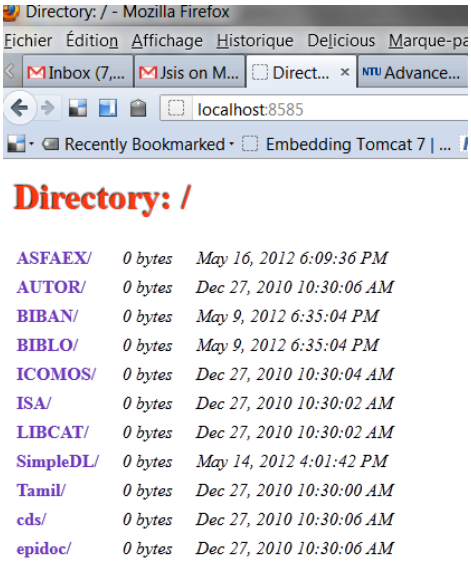
For example, if J-ISIS is installed in folder "`C:\jisis_suite 26 June 2012\jisis_suite`" and the *dbhome.conf* content is:

```
# Upper/lower case is important under unix
DEF_HOME=./home_example_db
```

Then the J-ISIS Web server document root is:

```
C:\jisis_suite 26 June 2012\jisis_suite\home_example_db
```

And typing http://localhost:8585 will display all databases which are children of the document root.



I.3 XHTML CSS JavaScript

J-ISIS display component is an embedded Web Browser that expect to get a Web document for rendering. In order to incorporate the ability of outputting Web documents from J-ISIS, it is needed to master the task of creating Web documents.

The basic building blocks of the Web are —HTML (or XHTML), CSS and JavaScript. They are also called the three layers of Web Development:

Content or Structure - The content or structure layer is what users will see when the Web document is displayed in the Web browser. Content can consist of text and/or images and includes links to navigate. In Web development, **XHTML** makes up the content layer and it also structures the Web document.

Style or Presentation -The style or presentation layer is how the document will look to the readers. This layer is defined by the **CSS** or styles that indicate how the document should be displayed and on what media types.

Behavior -The behavior layer is the layer of a Web page that has some interactivity. Ajax or DHTML can be used to send/receive data to the server side. For most Web pages, the first level of behavior is the **JavaScript** interactions on the page.

a) The Basic Structure of an (X)HTML Document

A number of basic structural elements are required to make a valid (X)HTML page. Basically, everything should fit into a structure outline that looks something like this:

```
<!DOCTYPE html PUBLIC "-//W3C//DTD XHTML 1.0 Strict//EN"
"http://www.w3.org/TR/xhtml1/DTD/xhtml1-strict.dtd">
<html xmlns="http://www.w3.org/1999/xhtml" xml:lang="en">
<head>
<title></title>
</head>
<body>
</body>
</html>
```

b) Hypertext

The **HT** in **HTML** stands for *Hypertext* . Hypertext is the ability to click on a link in one page and have another page open. There are two parts to a hypertext link: the location of the new page and the link text that appears in the browser. The location of the pages is specified as a *Uniform Resource Locator* (URL), which contains four parts: protocol, server, path and name. The protocol could be http, ftp, telnet or others. The protocol is followed by a colon and two slashes (://). After the protocol is the server. The server is followed by a slash and the path of the directory that contains the resource. The name of the resource follows the path. protocol://server/path/name The URL of the hypertext link is not displayed in the browser, but it is associated with the underlined text on the web page. Another way to say this is that the URL has to be included in the markup, since it does not appear as plain text.

URL
A Universal Resource Locator is the location you want to link to (in the case of links, the value of the ***href*** attribute), be it a website, page, or any other file.

The Anchor Element

One of the most important elements in XHTML is of course the hyperlink element. This element allows to create the hyperlinks within XHTML documents.
The anchor element is represented by the `<a>` element. Anchor elements can be placed just about anywhere and house absolute URLs such as `http://www.unesco.org` or relative URLs, such as `./mydatabase/work/printsort.html`.
The tag for a hypertext link is the paired tag <a> , which is short for *anchor* . Note that the text that is visible in the browser is not inside a tag, but that the URL of the file is. This is an example of a tag that has additional information stored in it. The additional information is called an *attribute* . The URL of the page is stored in an attribute named *href* . Attributes in HTML tags provide extra information that is not visible in the browser. This agrees with the basic definition of HTML as having plain text and tags. The tags contain extra information

about how to display the plain text. In this case, when the user clicks on the plain text, the browser will read the URL from the *href* attribute and request the resource from the server.

The image Element

The image element in XHTML allows for the insertion of an image within the Web content. An image, for example, can be placed into the document with:

```
<img src="./mydatabase/images/portrait.jpg" alt="Self Portrait"/>
```

The `src` attribute accepts the URL that references the image, and the `alt` attribute displays an alternative message.

Universal Resource Locators (URLs)

Every resource available on the Web --- HTML document, image, video clip, program, etc. --- has an address that may be encoded by a *Universal Resource Locator*, or "URL" (defined in [RFC1738]).
URLs typically consist of three pieces:

1. The name of the protocol used to transfer the resource over the Web.
2. The name of the machine hosting the resource.
3. The name of the resource itself, given as a path.

Consider the URL that designates the current HTML specification:

http://www.w3.org/TR/WD-html4/cover.html

This URL may be read as follows: Use the HTTP protocol to transfer the data residing on the machine www.w3.org in the file /TR/WD-html4/cover.html

URLs in general are case-sensitive (with the exception of machine names). There may be URLs, or parts of URLs, where case doesn't matter, but identifying these may not be easy. Users should always consider that URLs are case-sensitive. /u

The character set of URLs that appear in HTML is specified in [RFC1738].

Relative URLs

A *relative* URL (defined in [RFC1808]) doesn't contain any protocol or machine information. Its path generally refers to a resource on the same machine as the current document. Relative URLs may contain relative path components (".." means one level up in the hierarchy defined by the path), and may contain fragment identifiers.

Relative URLs are resolved to full URLs using a base URL. [RFC1808] defines the normative algorithm for this process.

As an example of relative URL resolution, assume we have the base URL "http://www.acme.com/support/intro.html". The relative URL in the following markup for a hypertext link:

```
<A href="suppliers.html">Suppliers</A>
```

would expand to the full URL "http://www.acme.com/support/suppliers.html", while the relative URL in the following markup for an image

```
<IMG src="../icons/logo.gif" alt="logo">
```

would expand to the full URL "http://www.acme.com/icons/logo.gif".

Full URLs

Briefly, a full URL is derived from a relative URL by attaching a "base" part to the relative URL. The base part is a URL that may come from any or all of the following sources:

- HTTP transfer protocol header information (see [RFC2068]).
- Metadata (the META element) in the HEAD section of an HTML document.
- Explicit base path information (the BASE element) in the HEAD section of an HTML document, or the CODEBASE attribute of the APPLET element.

II Building Web documents with J-ISIS Print Formats

II.1 How J-ISIS Process a Format:

When J-ISIS processes a format, it works with three objects: a data base **record**, the ***format and a work area*** where the output produced by the format is stored. The commands are executed sequentially in the order they are listed in the format. Some commands produce data (e.g. the contents of a given field), while others produce actions (such as skipping to a new line, leaving one or more blank lines, etc.). The data produced uses HTML as markup language, and is stored in the work area, which is then passed to the J-ISIS embedded Web Browser for further processing, e.g. for interpreting and displaying.
It's important to realize that the **J-ISIS application has an embedded Web Browser** that is used for displaying the display formats, **as well as a Web server which is used for serving static web resources** that can be html pages or any document file.

II.2 How to define static resources in J-ISIS Web documents

J-ISIS Web documents can be generated using ISIS Print Formatting Language and J-ISIS records . A resource is anything that is addressable over the Web, i.e. resources that can be accessed and transferred between clients and servers.

The rule is rather simple:

- Use absolute links if the resource is somewhere on the web or not a descendant from the J-ISIS server document root (defined by 1st DB_HOME value).
- Use relative links if the resource is on the J-ISIS server machine where the database is hosted.

J-ISIS will automatically replace relative URLs (or links) by absolute URLs (or links)
The base URL default value is define as = "http://localhost:8585/"
Thus for example, if the image portrait.jpg is stored in the /images folder of mydatabase. The image link must be defined as follow:

```
<img src=”./mydatabase/images/portrait.jpg” alt=”Self Portrait”/>
```

and it will be replaced by:

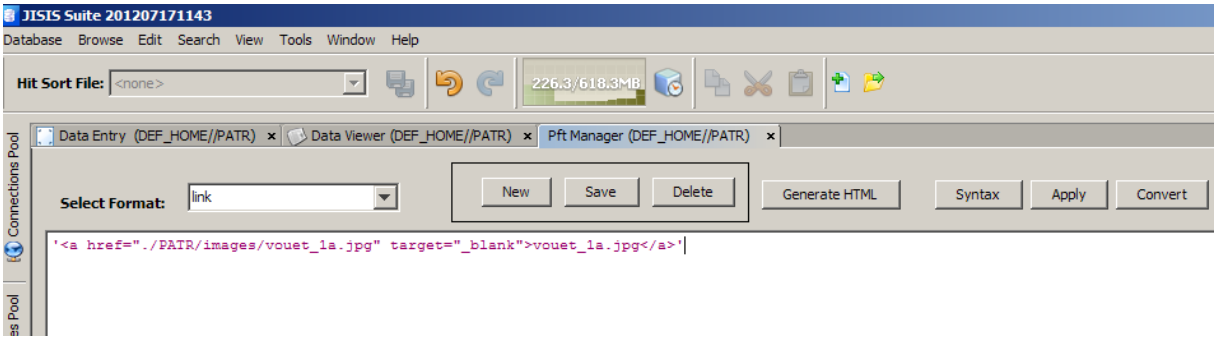
```
<img src=” http://localhost:8585/mydatabase/images/portrait.jpg”  alt=”Self Portrait”/>
```

The base URL default value can be changed by adding a line to the server.conf file
For example:
jetty.webserver.baseurl=http://www.j-isis.org

II.3 Web Documents with ISIS Print Formatting Language

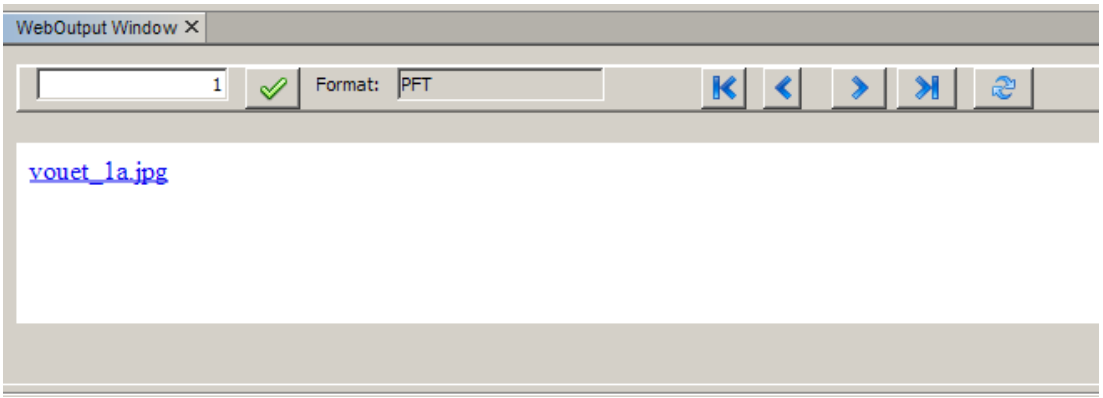
a) PFT Hard Coded URL

We can use the PFT Manager to create the following small PFT. Everything is written as unconditional literal.



Important Note: This is user's responsibility to generate the HTML formatting using PFT literals. J-ISIS doesn't generate the HTML for you except for the RAW format.

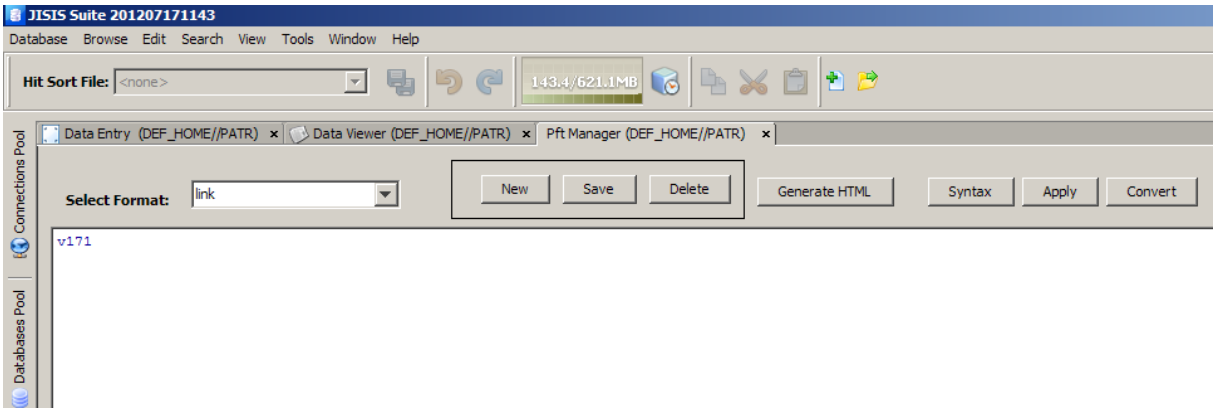
Clicking on Apply will display the interpreted PFT in the WebOutput Window:



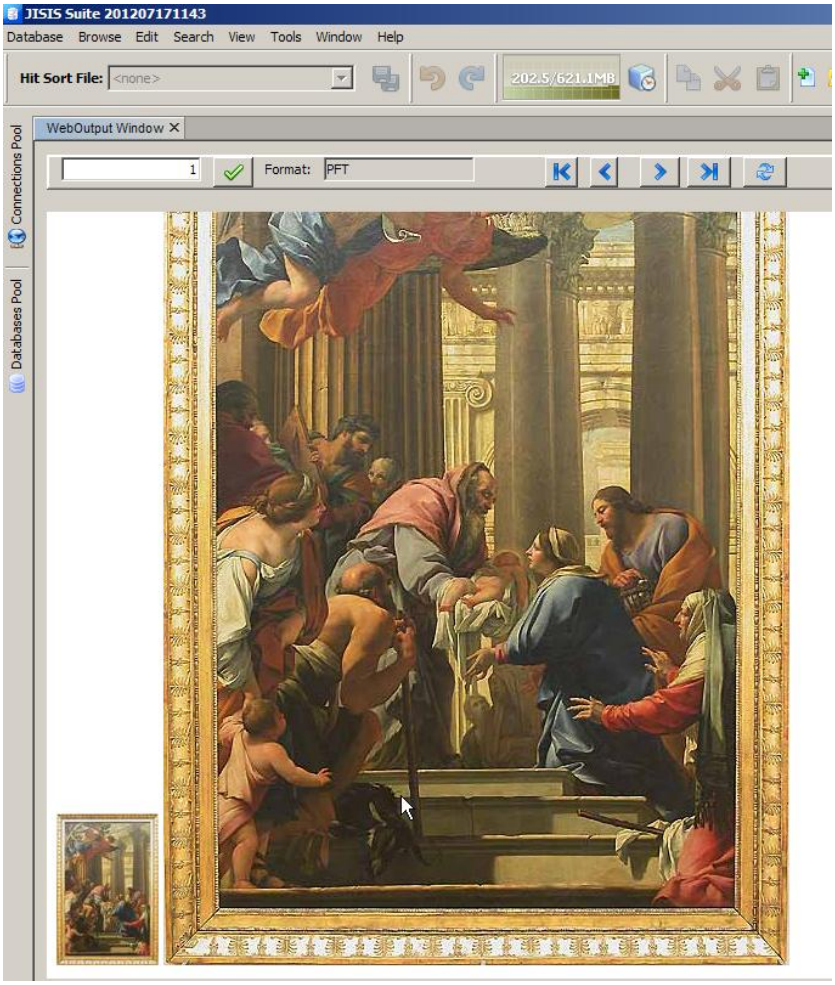
You can test the link by clicking on “vouet_1a.jpg”

b) Displaying images from BLOB fields in a PFT

Using the PFT Manager, we can enter this simple format:



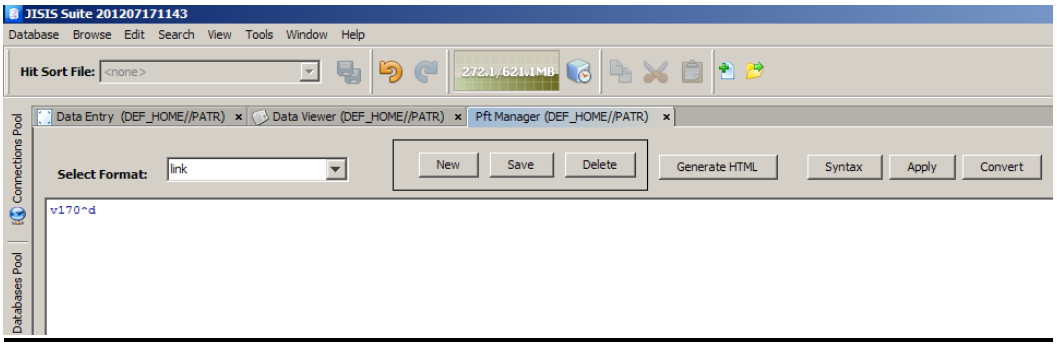
Clicking on “Apply” will interpret the PFT against the 1st record and display it in the WebOutput window:



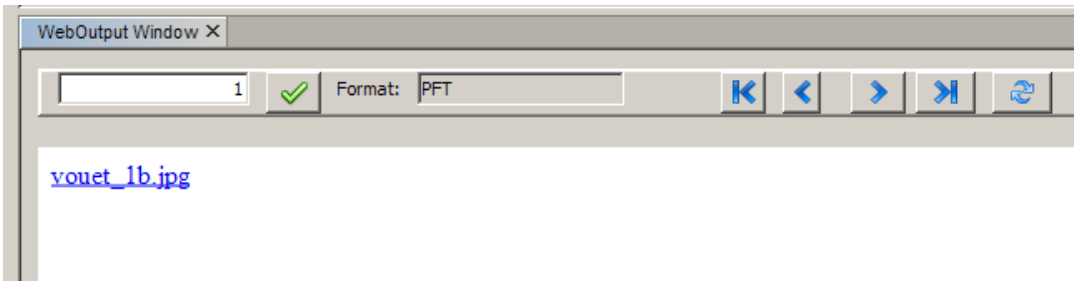
We can see the two occurrences of field V171

c) URL stored in a Field/Subfield

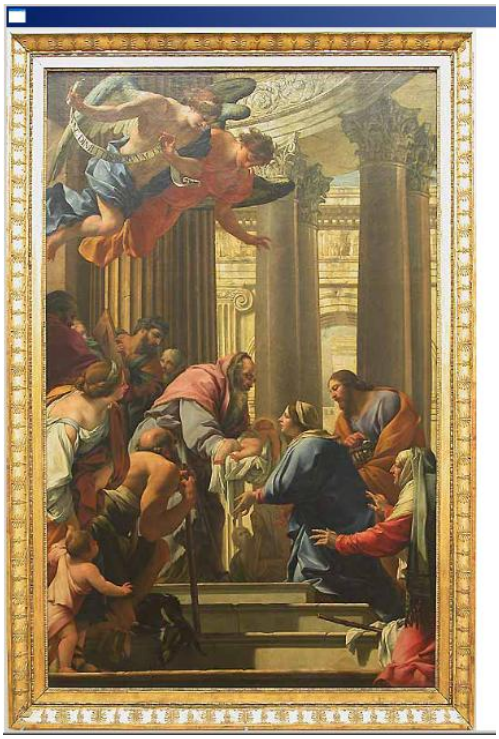
We can use a more flexible approach if the URLs are stored in record fields. For example, we can use field with tag 170 of database PATR as follow:



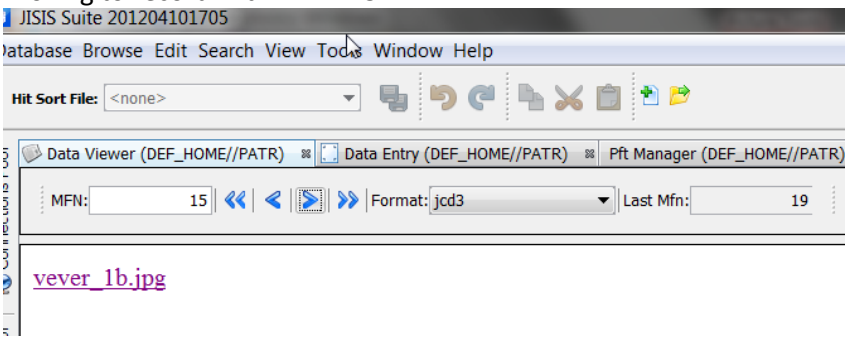
By clicking on "Apply", we can see the result of the PFT apply to the 1st record:



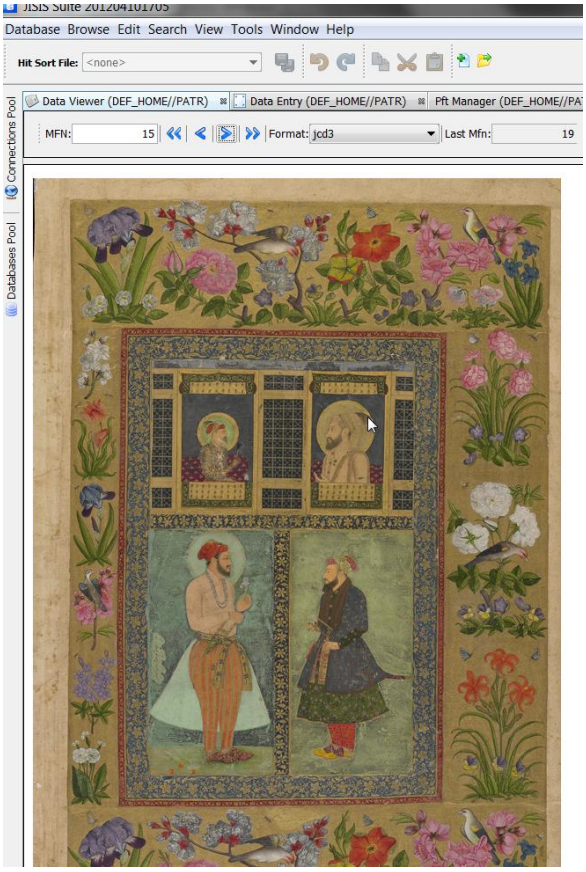
And clicking on the link popup the following window



Moving to record with MFN=15



And clicking on the link



HTML tag <img> can also be used if you want to include the image inside text provided by other fields/subfields or the PFT.

The <**img**/> element allows you to insert an image straight into your HTML.

```
<img src="./PATR/images/vever_1b.jpg" alt="Vever Collection" />
```

III. Images and HTML in J-ISIS.

J-ISIS records are displayed according to print format definitions written in the ISIS formatting language. The "RAW" format is always available and used by default

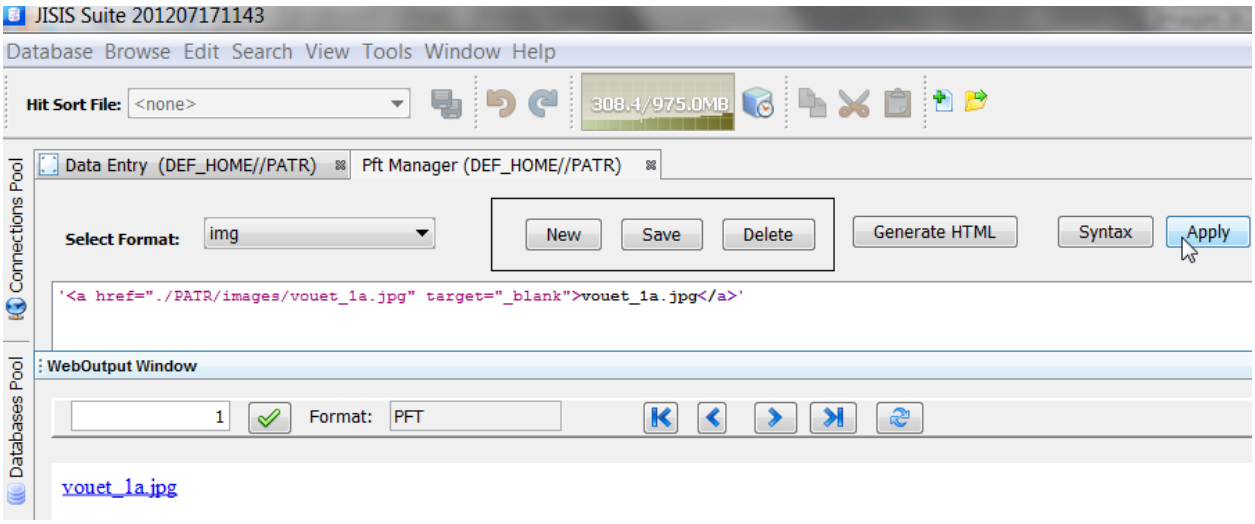
J-ISIS offers two methods for storing images:

1) Images are stored separately in external files and are accessed through URLs

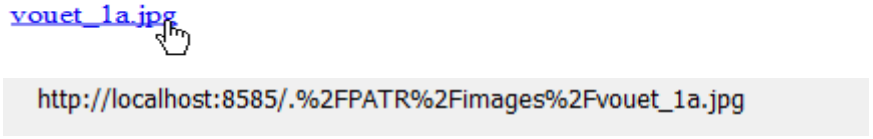
2) Images are stored in database records as Binary Large Object (BLOB)

III.1 Images are stored separately in external files

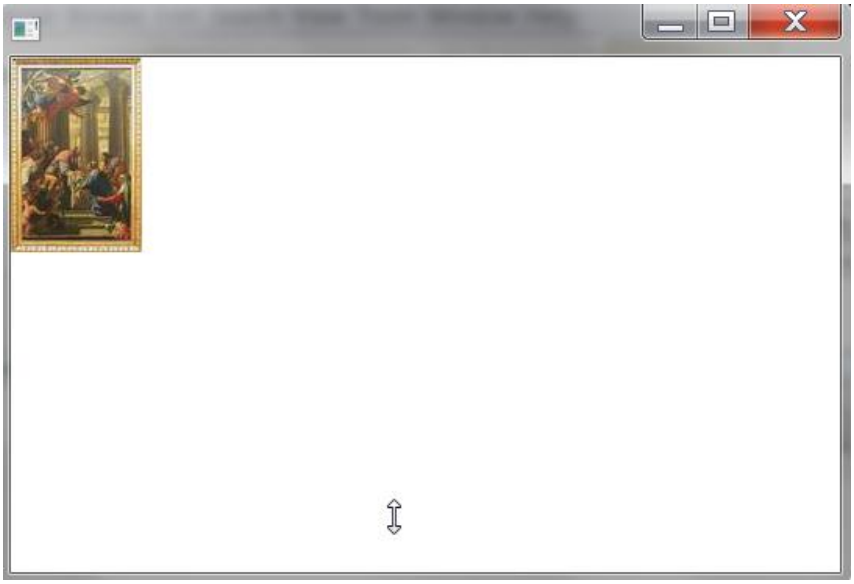
They can be accessed through urls such as : ./PATR/images/vouet_1a.jpg , PATR is the database name.
Thus a simple PFT such as:
'vouet_1a.jpg'
will provide a link to the image.
It's quite easy to test it inside the PFT Manager after opening the PATR database:



The absolute url is displayed at the left bottom corner when the mouse cursor is hover the link:



And clicking on the link will display the image:



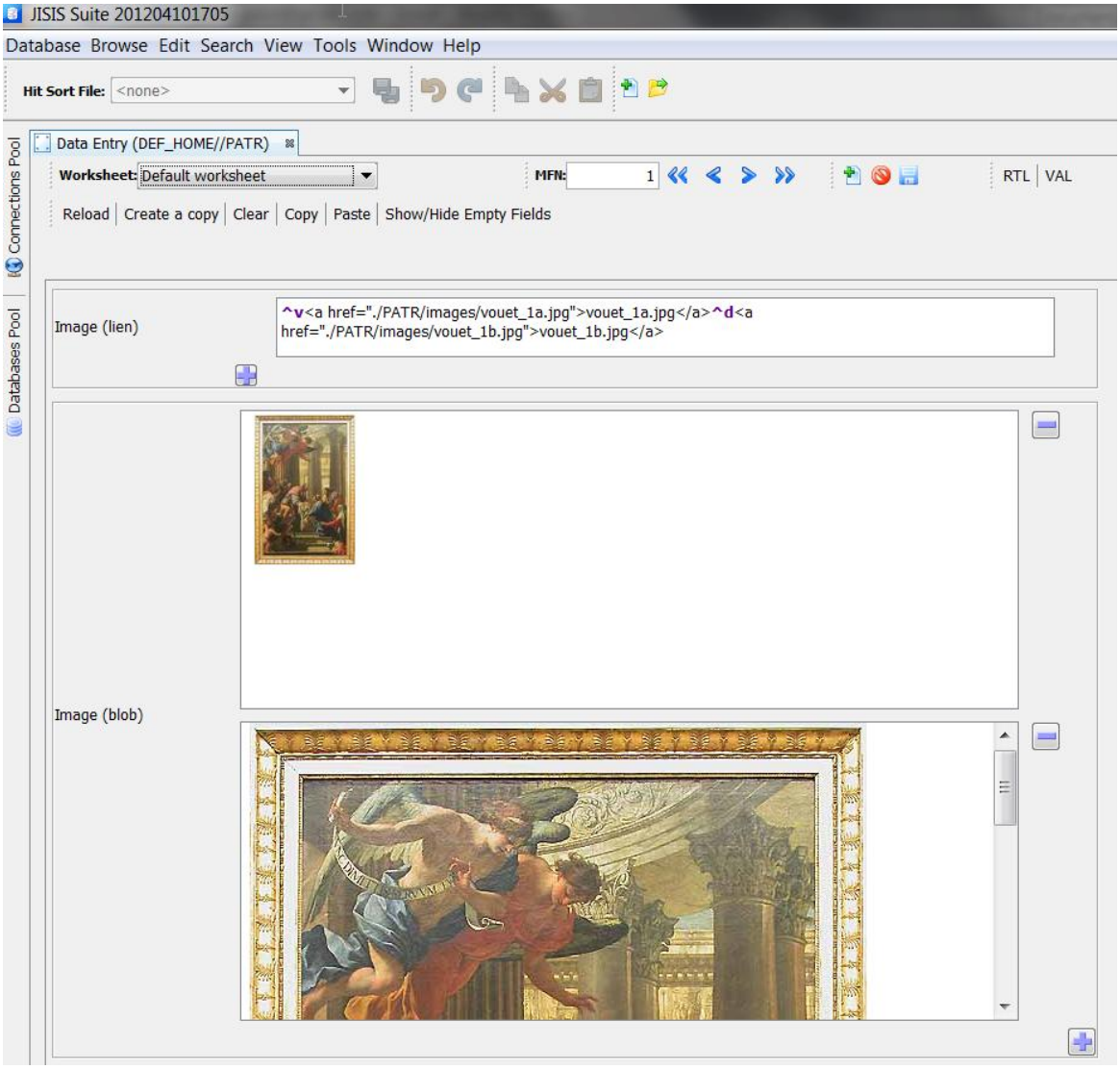
Please note that the image is served by the embedded Jetty server

III.2 Images stored in BLOB Field Type

It's possible to define the field type of a particular field as BLOB in the FDT. A BLOB field allows to copy and paste a mixture of text and images, or a single image. The Data Entry module allows to enter images by doing Cut and Paste. In that particular case, the image is stored in the record itself.

III.3 Example

Let's have a look the PATR database which is part of the J-ISIS example databases. This database is particularly interesting because it uses BLOB fields and URLs.
We start J-ISIS, opening the default connection on the local host machine and we select the PATR database. To inspect how the techniques are used, we open the "Data Entry" module that displays an empty Data Entry form. Then we click on the  button to display the first record. We scroll down the data entry form until we can see the images as follow:

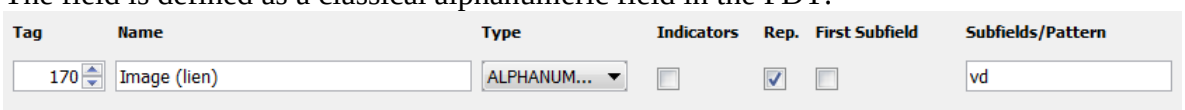


a) Images accessed through URLs

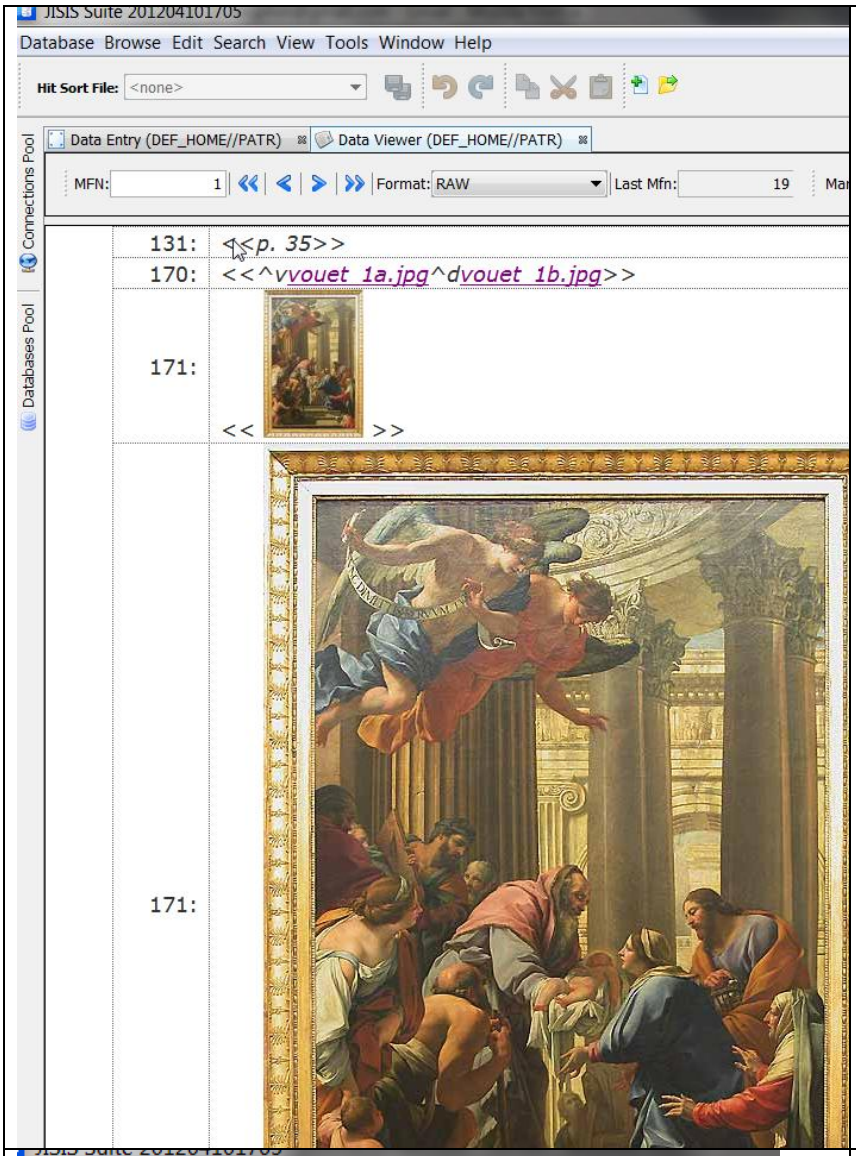
The field "Image (lien)" demonstrates how you can provide HTML hyperlinks as URLs to images which are stored in the PATR/images folder:

```
^v<a href="./PATR/images/vouet_1a.jpg">vouet_1a.jpg</a>^d<a
href="./PATR/images/vouet_1b.jpg">vouet_1b.jpg</a>
```

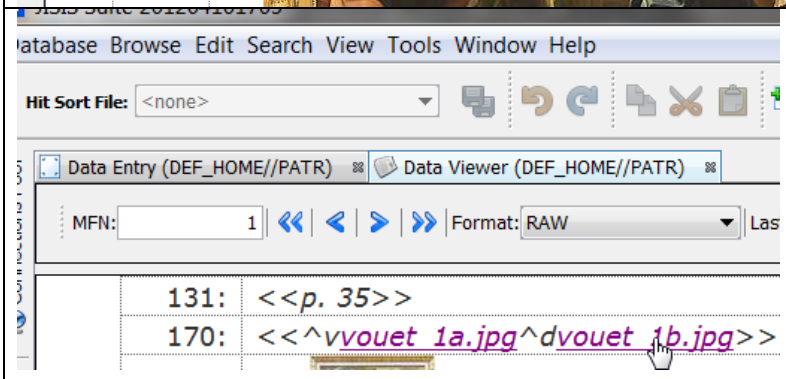
We have here two subfields where we store the relative URL of a thumbnail and the corresponding image. J-ISIS uses the machine native browser for displaying records and contains an embedded Web server. This is why we can use HTML either in record fields or in Print Formats.
The field is defined as a classical alphanumeric field in the FDT:



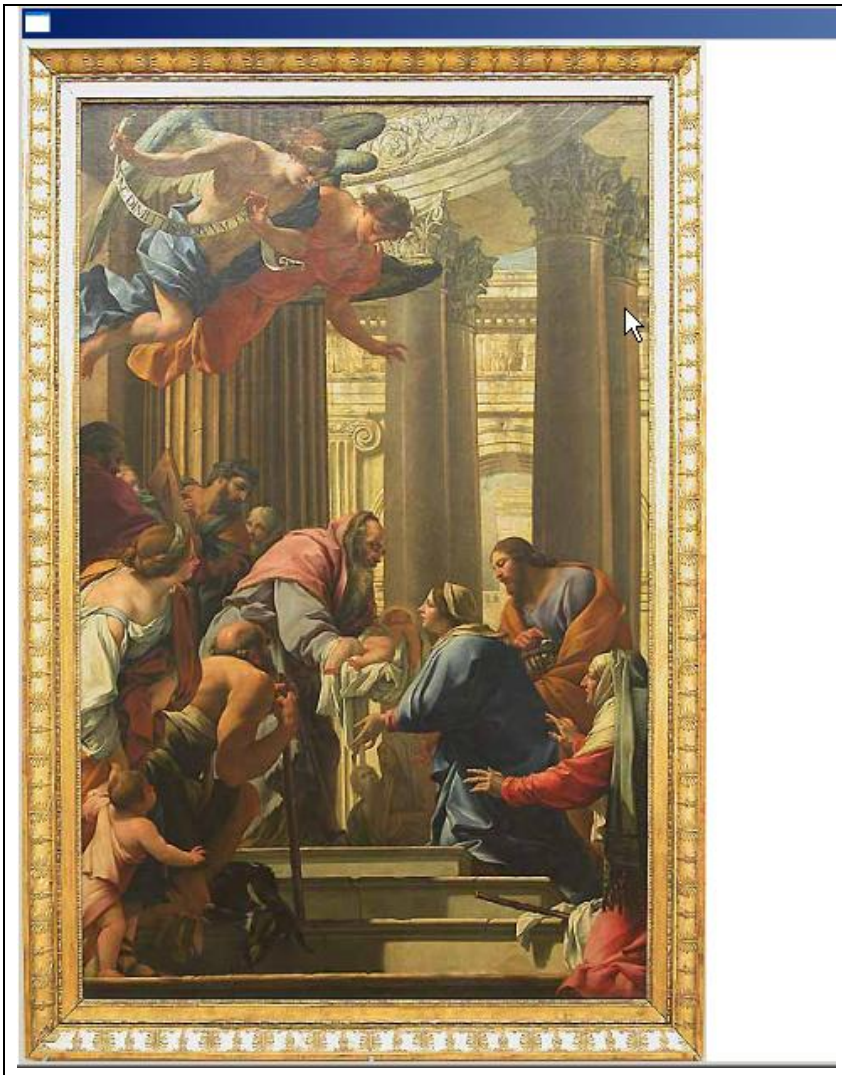
Looking at the record in the Data Viewer, we can see the record displayed once the HTML format interpreted by the browser:



We can see two hyperlinks on which we can click:
vouet_1a.jpg
vouet_1b.jpg



Clicking on vouet_1b.jpg will load the full image

	The full image is displayed in a popup window

IV. Other Web resources

You can also use links to documents in your URLs. Standard file extensions are those that are recognized by most web servers and web browsers.

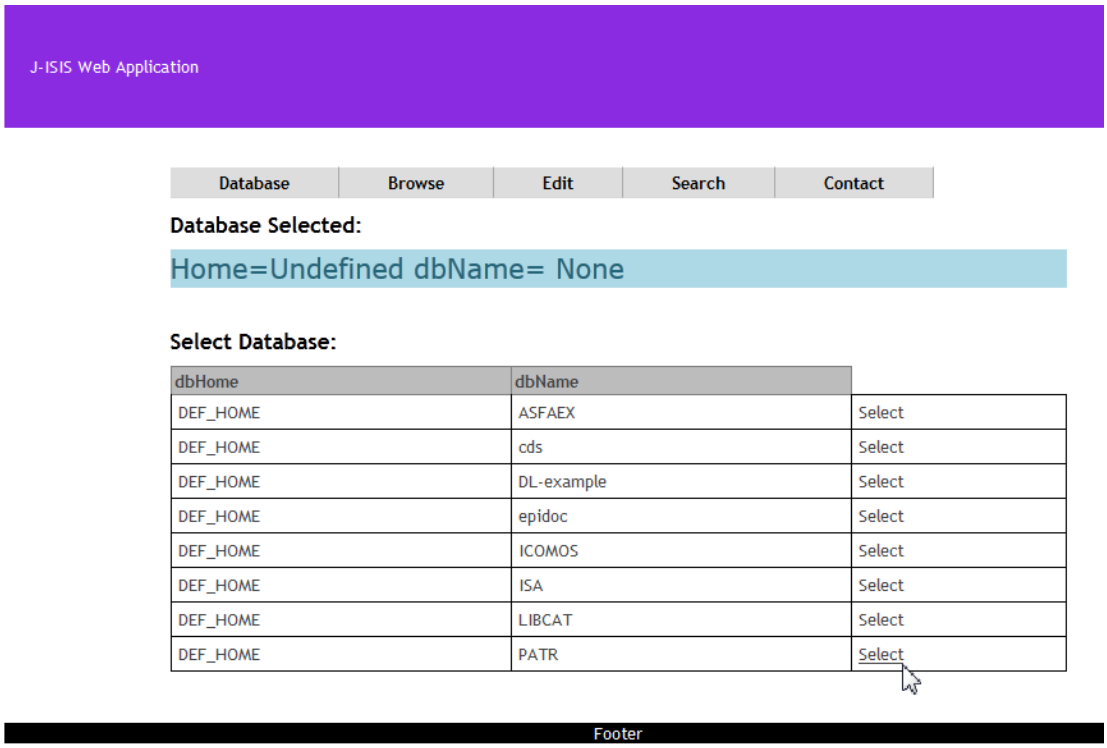
Common web file types and extensions		
File type	**Typical use**	**Extension**
HTML document (webpage)	HTML code / webpage	`.html`, `.htm`
JPEG	photos	`.jpg`

Common web file types and extensions		
File type	**Typical use**	**Extension**
GIF	graphical images, such as logos, graphs and charts	.gif
PNG	graphical images, such as logos, graphs and charts (similar, but superior to GIF)	.png
Portable Document Format (PDF)	documents, e-books, product catalogues (non-editable)	.pdf
Word / word-processing document **DOC, DOCX, ODF, XLS, PPT**	Documents, application forms, spreadsheets, Presentations, etc	.doc, .docx, .odf, .xls, .ppt etc

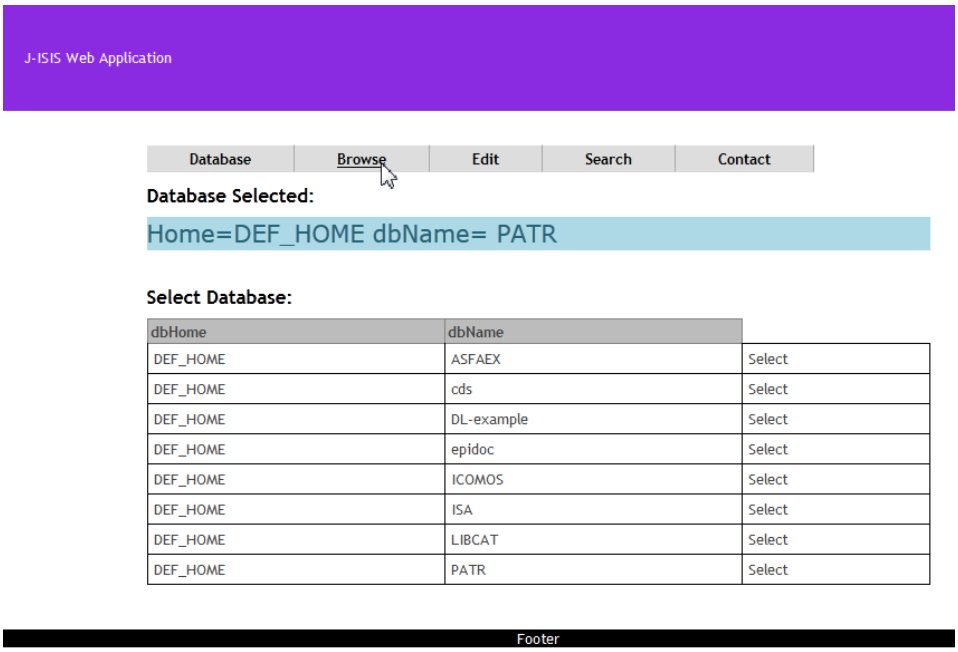
If you want to index the document text, it is better to use the Digital Library Data Entry module. In that case, the document will be converted in Plain text format and indexed, the document will be copied on the server side and the document URL will be generated for you.

V. Using Web-JISIS

i) Select the PATR demo database



ii) Click on Browse



iii) The 1st record is displayed with the images

J-ISIS Web Application

Database | Browse | Edit | Search | Contact

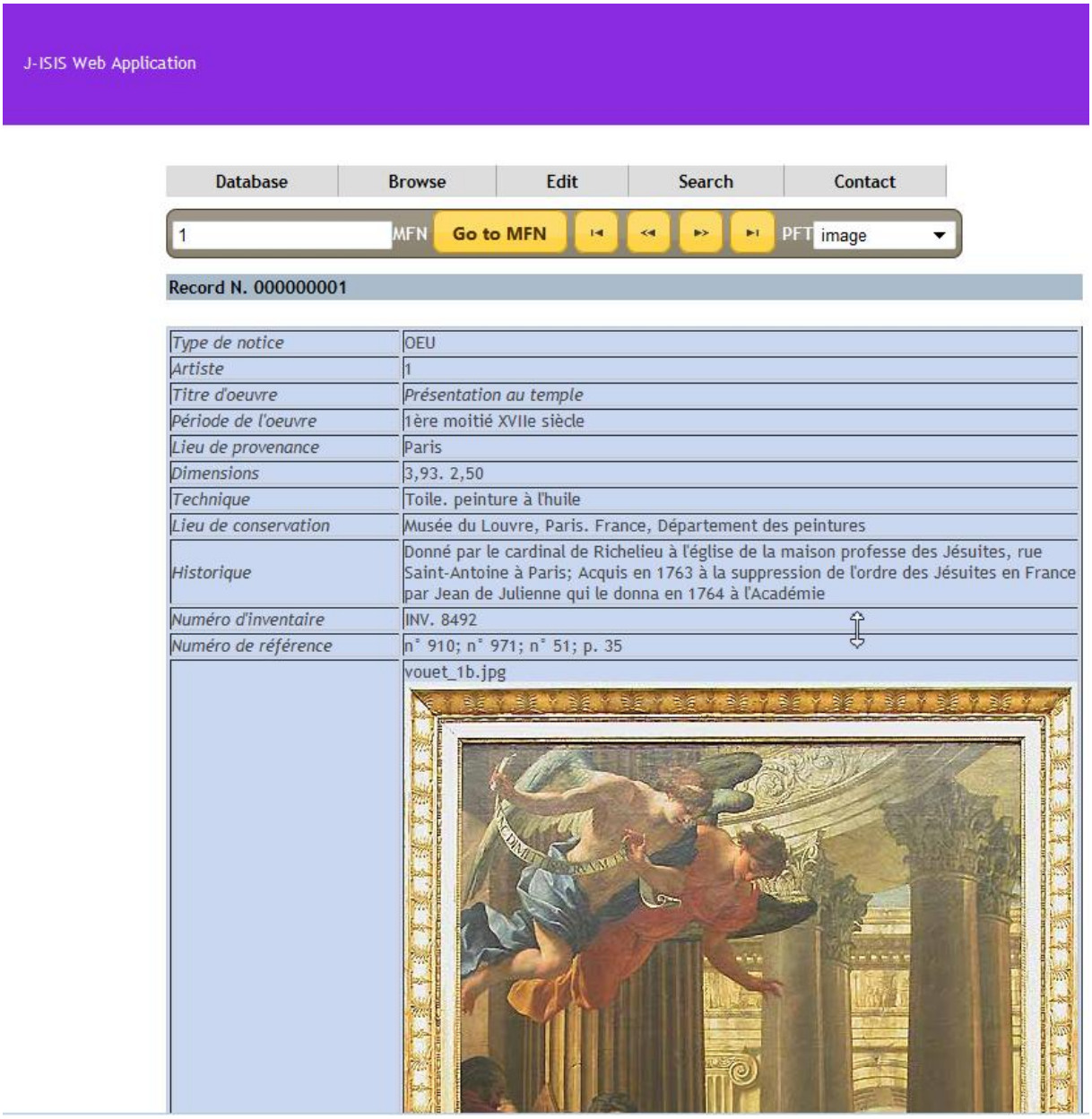
1 MFN Go to MFN PFT RAW

RECORD(1)

Tag	Field/Occurrence
1:	*<<OEU>>*
3:	*<<Peinture>>*
10:	*<<1>>*
20:	*<<Présentation au temple>>*
30:	*<<1ère moitié XVIIe siècle>>*
42:	*<<Paris>>*
50:	*<<^h3,93^l2,50>>*
60:	*<<^mToile^tpeinture à l'huile>>*
95:	*<<OEU>>*
96:	*<<1>>*
97:	*<<1>>*
97:	*<<2>>*
97:	*<<2>>*
97:	*<<4>>*
100:	*<<^mMusée du Louvre^cParis^pFrance^dDépartement des peintures>>*
110:	*<<Donné par le cardinal de Richelieu à l'église de la maison professe des Jésuites, rue Saint-Antoine à Paris>>*
110:	*<<Acquis en 1763 à la suppression de l'ordre des Jésuites en France par Jean de Julienne qui le donna en 1764 à l'Académie>>*
130:	*<<INV. 8492>>*
131:	*<<n° 910>>*
131:	*<<n° 971>>*
131:	*<<n° 51>>*
131:	*<<p. 35>>*
170:	*<<^vvouet_1a.jpg^dvouet_1b.jpg>>*
171:	<  >>

iv) Clicking on vouet_1b.jpg loads the image in the browser

v) Selecting the PFT image displays the following screen:



vi) And selecting the PFT notice will display the following screen:

J-ISIS Web Application

Database | Browse | Edit | Search | Contact

1 MFN Go to MFN PFT notice

Record N. 000000001

Artiste	VOUET, Simon (1590-1649) école française
Titre d'oeuvre	*Présentation au temple*
Période de l'oeuvre	1ère moitié XVIIe siècle
Lieu de provenance	Paris
Dimensions	3,93. 2,50
Technique	Toile. peinture à l'huile
Lieu de conservation	Musée du Louvre, Paris. France, Département des peintures
Historique	Donné par le cardinal de Richelieu à l'église de la maison professe des Jésuites, rue Saint-Antoine à Paris; Acquis en 1763 à la suppression de l'ordre des Jésuites en France par Jean de Julienne qui le donna en 1764 à l'Académie
Numéro d'inventaire	INV. 8492
Image (lien)	vouet_1a.jpg vouet_1a.jpg
Bibliographie	MUSÉE DU LOUVRE, Département des peintures. Paris. ROSENBERG, Pierre. REYNAUD, Nicole. COMPIN, Isabelle. *Catalogue illustré des peintures : école française XVIe et XVIIe siècle*. 2. Paris : Ed. des Musées nationaux, 1974 [réf. n° 910] MUSÉE DU LOUVRE, Département des peintures. Paris. BRIÈRE, Gaston. *Catalogue des peintures exposées dans les galeries. 1, Ecole française*. Paris : Musées nationaux, 1924 [réf. n° 971] MUSÉE DU LOUVRE, Département des peintures. Paris. BRIÈRE, Gaston. *Catalogue des peintures exposées dans les galeries. 1, Ecole française*. Paris : Musées nationaux, 1924 [réf. n° 51] LACLOTTE, Michel. CUZIN, Jean-Pierre. *Louvre : la peinture européenne*. Paris : Scala, 1987 [réf. p. 35]

Footer