Upload an html content in Faculty Deposit

Scenario:

The User Services and Applications Working Group have requested that faculty be able to upload html content in the Faculty Deposit service. An example is this article http://www.istl.org/09-winter/viewpoint.html sent in through the Faculty Deposit service. This journal article is available only in html format and not in PDF.

If Article, Refereed or Article, Non-refereed is selected from the Faculty Deposit field "category", an option of File Type "HTML" is provided (in addition to the ones already provided) on the File Upload page. When HTML format is selected, it will allow the user to enter only one URL for depositing. We will consider allowing multiple URLs (e.g., for a multiple-page article) if there is user need in the future.

After the HTML URL is submitted in Faculty Deposit, the WMS will create PDF-1, XML-1, and JPEGTHUMB-1 datastreams. The WMS will create ARCH1 to include the PDF format of the file only. The repository will not archive the file in HTML format.

Implementation:

- **PDF server:** need to provide a method to convert from HTML URL to PDF file without the HTML tags.
- **WMS:** will use the new PDF server method to create a PDF-1 datastream and configure the file policy to accept HTML content in faculty deposit and in "native" WMS for the content models Document and Manuscript.
- **Faculty Deposit:** On the File Upload page, provide an option to accept an HTML URL and check if the URL is valid before sending it to the WMS. If the URL is not valid, open up a popup window to notify the submitter: "This is not a valid URL. Please check the link and re-submit it. If the problem persists, contact the RUcore staff<contact link>."