OOo-Sitemap

Many OpenOffice.org pages are published in a multi sub-domain structure. See http://wiki.services.openoffice.org/wiki/Infrastructure_Overview for details. With the exception that the main site is now hosted on kenai, this is probably accurate.

Partial active address list:

Project name	URL	Hosted at
About	http://about.openoffice.org	Kenai
API	http://api.openoffice.org	Kenai
Bugzilla	http://openoffice.org/bugzilla/	Kenai
Development	http://development.openoffice.org	Kenai
Distribution	http://distribution.openoffice.org	Kenai
Documentation	http://documentation.openoffice.org	Kenai
Download	http://download.openoffice.org	Kenai
Main page	http://www.openoffice.org	Kenai
Marketing	http://marketing.openoffice.org	Kenai
Native pages list	http://l10n.openoffice.org	Kenai
Projects list and individual addresses (146 projects)	http://projects.openoffice.org	Kenai
Support	http://support.openoffice.org	Kenai
Extensions	http://extensions.services.openoffice.org	OSUOSL
Forums	http://user.services.openoffice.org	Oracle
Templates	http://templates.services.openoffice.org	OSUOSL
Wiki	http://wiki.services.openoffice.org	Oracle

also see OpenOffice Domains for a more complete list. Generally, any project has its own URL.

A sitemap of the webpages located on kenai.com is add'ed above. Same NLC projects are missed cause tecnical issues. (e.g. es.oo.o)

Archive create

Possible:

1. Web content checkout via SVN URL.
   In the AOOo project in https://svn.apache.org/viewvc/incubator/ooo/trunk/tools/dev/ is a script and web project list that automates checkout and update.
   Look for fetch-all-web.sh and web-list.txt. The text file needs to be edited. The script performs svn update on existing project directories to save time.
   Here is the how to do it individually.
   Syntax:
   svn co https://svn.openoffice.org/svn/<$projectname>~webcontent your_local_dir
   Example:
   svn co https://svn.openoffice.org/svn/download~webcontent download --> to get all website content from the download project
   Do it analog with the other projects.
2. Wiki: database dump ~~(Clayton Cornell is able to help with this)~~. Clayton is no longer an available resource. TerryE has dumps and full VM copies of both the wiki and forums. Use me (TerryE) as a source (subject to access approvals).
3. Bugzilla: I hope, ORACLE will provide a database dump if not, we can use XML export. Bugzilla can import this XML's.
4. Forums: As I know we have admins of the OOo user forums in our group, they can make a dump of the database via the PHPbb admin interface.
5. Extensions and Templates: We really need to backup this. AFIAK the servers of this services are not hosted by ORACLE, they are hosted at OSUOSL.
6. Use wget

Note: I (rbircher) have allready a script to make a serie checkout of all projects, the only thing that I need is a .txt file who lists all project names (line break separated)

Todo plan

1. Create full sub-domains list (Substantial progress)

2. Create archive (can do it in a people.apache.org account. (development, documentation, download, projects, and www take 2.7GB.)
3. Determine how to deal with current "projects" (many!): what to do with per project mailing list; combine or eliminate projects? (See OpenOffice Domains for suggestions on web sites.) We especially need to discuss/migrate mailing lists if still active (i.e. activity since Jan, 2011).
4. Selecting needed content
5. Move contents to new pages or directly import SVN checkouts from OO.o site.
6. Contact project heads via e-mail to ascertain continued interest if needed.