

#226 - Replace CVS with SVN on www.inf

- test server is skelp.inf, with aliases svntest.inf & pubtest.inf
- in the web-publishing world, users will need to authenticate via CoSign before attempting to publish anything - use a <site>/login trigger for this

...browser edit

- login via [https://pubtest\[.INST\].inf.ed.ac.uk/login](https://pubtest[.INST].inf.ed.ac.uk/login)
- browse to
[https://www2\[.INST\].inf.ed.ac.uk/PAGE](https://www2[.INST].inf.ed.ac.uk/PAGE)
- File: Edit Page to open SeaMonkey Editor
- publish to [https://pubtest\[.INST\].inf.ed.ac.uk](https://pubtest[.INST].inf.ed.ac.uk)

...browser edit (cont'd)

- user edits page in SeaMonkey and hits "publish"
- apache uses page URL to select configuration options (eg pubtest.cstr.conf or pubtest.cstr.ssl.conf)
- configuration options trigger publish script:
 ScriptAlias /cgi-bin /liveroot/pcgi
 Script PUT /cgi-bin/publish.pl
- publish script calls `"/usr/bin/svn -q commit ..."`
 [access via /disk/data/infweb/svn/webdav_authz,
 created at startup]

...browser edit (cont'd)

- SVN uses repo-wide hook scripts (/disk/data/infweb/svn/infweb/hooks) to trigger commit actions
- repo-wide pre-commit script runs /disk/data/infweb/svn/scripts/sgmlcheck.pl
- repo-wide post-commit script runs /disk/data/infweb/svn/scripts/auto-export.pl
- confirmed /disk/data/infweb/checkout/pcgi/editmeta uses new meta-data location (\$metabase = "/disk/data/html.meta/")
- creation of new page by browsing to non-existent URL and editing & publishing that page

...CLI edit

- users checks out file(s) to local space
- user edits page & commits file
- publish script (publish.pl) is NOT called
- SVN uses repo-wide hook scripts (/disk/data/infweb/svn/infweb/hooks) to trigger commit actions
- repo-wide pre-commit script runs /disk/data/infweb/svn/scripts/sgmlcheck.pl
- repo-wide post-commit script runs /disk/data/infweb/svn/scripts/auto-export.pl

...DB pages sync

- in the CVS world, the DB file was stored in CVS and exported to /liveroot (checkout), but we don't want to perpetuate the CVS/SVN use (just integrate into /liveroot)
- create separate auto-pages tree, /web-auto
- implement the cvs commit of DB files to cvs.inf (aka knox), so that we DON'T need a CVS *server* on skelp. The DB will commit to cvs.inf (which should put files somewhere local)

...DB pages sync (cont'd)

- add Alias for each location, for example:

```
Alias /research/hcrc      /liveroot/web-auto/research/hcrc
Alias /people             /liveroot/web-auto/people
```

- DB conduit commits a checked-in version to somewhere like `cvs.inf:/disk/cvs/DBci`
- CVS file gets checked-out to somewhere like `www.inf:/disk/DBco`
- use "cvs export" to update contents of `/disk/data/infweb/checkout/web-auto (/liveroot/web-auto)` using contents from `/disk/DBco`

...DB pages sync (cont'd)

- need to co-ordinate the run of "cvs export", so that it's not run until the DB-dump has finished
- need to check we retain .meta file functionality
- only \$CVSROOT/CVSROOT/logininfo needs to be triggered by DB conduit "cvs commit", which should be used to signal www.inf to pull over changed files to /liveroot/web-auto

To Do

Complete the implementation of the auto-generated update mechanism, including:

- scripting of transfer mechanism for CVS-committed files from DB conduits (on knox) to new www.inf
- co-ordinating the run of "cvs export" so that it's not run until the DB-dump has finished (either time-constrained via cron, or triggered by a flag file).
- preventing DB-generated files from being over-written by web-published files, possibly by removing relevant write-access via /svn/webdav_authz)
- Review the various elements already completed - making sure that they are still consistent with each other and provide the necessary functionality. A dry-run installation from scratch would be desirable.

Notes...

- have kept notes of journey so far...
- ...currently around 170,000 words
- American universities limit PhD dissertations to 100,000 words
- Median length for all books is about 64,000 words
- Novels are typically between 100,000 and 175,000 words
 - typical mystery novel might be in the 60,000 to 80,000 word range
 - Huxley's "Brave New World" has 64,531 words
 - "War and Peace" has 544,406 words