From Bibliography Files to Open Archives: the Sequel

Jean-Michel HUFFLEN
LIFC — University of Franche-Comté
BachoTEX, 1st May 2011
Contents

Populating HAL Open Archives
Additional information within .bib files
Additional files
Ready to go!
Other programs to do that?
Targetting other Open Archives sites
Conclusion & appendix
Remember... At BachoTEX 2010...

French research institutions have to put articles’ texts and metadata into the HAL Open Archives site.
Remember...

At BachoTEX 2010...

French research institutions have to put articles’ texts and metadata into the HAL Open Archives site.

Institutions are graded regarding these publications.
Starting from .bib files

Since 1996 ➔ publications registered into .bib files.
Starting from .bib files

Since 1996 ➞ publications registered into .bib files.

Open Archives ← advanced search features ➞ some information to be supplied is not included into .bib files.
May 2010

\texttt{MiBIB\TeX}'s extension + additional XML files for:

- our institution's people (affiliation),
- conferences and journals.
Intermezzo

Activity report $\rightarrow$ classifying:

- *referenced* conferences, non-referenced ones;
Intermezzo

Activity report $\implies$ classifying:

- referenced conferences, non-referenced ones;
- the same for journals.
Intermezzo

Activity report $\Rightarrow$ classifying:

• *referenced* conferences, non-referenced ones;

• the same for journals.

Being referenced depends on the conference—for sure!—
Activity report $\rightarrow$ classifying:

- *referenced* conferences, non-referenced ones;
- the same for journals.

Being referenced depends on the conference—for sure!—but *may depend on the year*.
Activity report $\implies$ classifying:

- *referenced* conferences, non-referenced ones;
- the same for journals.

Being referenced depends on the conference—for sure!—but *may depend on the year*. The same for journals.
Tagging .bib files

INHAL yes, no;
Tagging .bib files

INHAL yes, no;

EQUIPE authors' research team;
Tagging .bib files

INHAL yes, no;

EQUIPE authors' research team;

CLASSEMENT mnemonic, such as ACLI, ACLNI, ACLF, ACLNF, etc.;
Tagging .bib files

INHAL yes, no;

EQUIPE authors’ research team;

CLASSEMENT mnemonic, such as ACLI, ACLNI, ACLF, ACLNF, etc.;

LANGUE the text’s language;

DOMAINEHAL identifies topics.
More fields

ABSTRACT KEYWORDS PDF

But how to fill them in?
More fields

ABSTRACT       KEYWORDS       PDF

But how to fill them in? And [quite] quicky?
Graphical interface

Designed and programmed by Jean-Michel Caricand, an engineer at the LIFC.
Graphical interface

Designed and programmed by Jean-Michel Caricand, an engineer at the LIFC.

$key \implies key\text{.abstract}, key\text{.keywords}, key\text{.pdf}.$
Information about people

LIFC’s people ← periods.
Information about people

LIFC’s people ↔ periods.

External people ↔ affiliations:
Information about people

LIFC’s people ← periods.

External people ← affiliations:

• as numbers registered by HAL,

• or given *in extenso*.
Steps

- .bib file
  - parser
  - SXML
  - pretty printer
  - XSLT 2.0
  - ...-hal.xml
  - XSLT 2.0
  - ...-mlbiblio.xml

string list
Allegro molto nervoso

Very big difference between \textsc{BibTeX} and \textsc{M\textbackslash BibTeX}: 
Allegro molto nervoso

Very big difference between \texttt{BIBTEX} and \texttt{M\textsc{\small l}BIBTEX}:

the latter is \textit{less permissive}!
Allegro molto nervoso

Very big difference between \BIBTEX\ and \MIBIBTEX: the latter is \textit{less permissive}!

More check about \texttt{ADDRESS} fields.
Allegro molto nervoso

Very big difference between \texttt{\LaTeX} and \texttt{M\LaTeX}: the latter is \emph{less permissive}!

More check about \textsc{Address} fields.

\Longrightarrow A kind of compilation of \texttt{.bib} files.
Allegro molto nervoso

Very big difference between \textsc{BibTeX} and \textsc{M\textbackslash BibTeX}:

the latter is \textit{less permissive}!

More check about \texttt{ADDRESS} fields.

\implies A kind of compilation of .bib files.

\texttt{EDITORS}, \texttt{PUBLISHER} in an \texttt{@ARTICLE}.
Another program to do that?

Yes!
Another program to do that?

Yes! (bib2hal)
Another program to do that?

Yes! (bib2hal) provided that you belong to the INRIA!!
Another program to do that?

Yes! (bib2hal) provided that you belong to the INRIA!!

Quite similar: additional fields, extra-information in XML files.
Another program to do that?

Yes! (bib2hal) provided that you belong to the INRIA!!

Quite similar: additional fields, extra-information in XML files.

Suitable for one person, difficult use at the level of an institution.
Installation and use

http://lifc.univ-fcomte.fr/home/~jmhufflen/texts/superreport/
Installation and use

http://lifc.univ-fcomte.fr/home/~jmhufflen/texts/superreport/

File smlbibtex-1.3.tar.gz.
Installation and use

http://lifc.univ-fcomte.fr/home/~jmhufflen/texts/superreport/

File smlbibtex-1.3.tar.gz.

Tested on Linux and Mac OS X.
Installation and use

http://lifc.univ-fcomte.fr/home/~jmhufflen/texts/superreport/

File smlbibtex-1.3.tar.gz.

Tested on Linux and Mac OS X.

Install a Scheme interpreter: preferably bigloo.
Installation and use

http://lifc.univ-fcomte.fr/home/~jmhufflen/texts/superreport/

File smlbibtex-1.3.tar.gz.

Tested on Linux and Mac OS X.

Install a Scheme interpreter: preferably bigloo.

Make sure that GNU make works.
Update w.r.t. the article

Single quotes are *no longer* required for the values related to the XSLT stylesheet, that is, for the options:

--with-hallogin  --with-external
--with-halpassword  --with-domain-table
--with-people
Usable at other places?

France: use configure's options.
Usable at other places?

France: use configure’s options.

Other countries: just adapt the second step, that is, the XSLT stylesheet. See also how to use the CLASSEMENT field.
Appendix

Programs:

ar-style, hal, mlbibtex2xml, mlbibtex
Appendix

Programs:

ar-style, hal, mlbibtex2xml, mlbibtex

Now ML\BIBTEX works in compatibility mode, because some configuration points are not fixed yet. (Forthcoming!)
Appendix

Programs:

ar-style, hal, mlbibtex2xml, mlbibtex

Now ML\BIBTEX works in compatibility mode, because some configuration points are not fixed yet. (Forthcoming!)

Latin1 encoding (input and output). Other encodings are planned.
Finally

Long but successful!
Finally

Long but successful!

\textsc{MisBibTeX} is winning again!