

BibTeX Bibliography Index Maker: Notes

Ramon Xuriguera

1 BibTeX

Aspectes del format BibTeX a tenir en compte:

- Com podem distingir entre diferents tipus d'entrada (article, book, inproceedings, etc.) a partir del fitxer?
- Format dels noms. Un nom consisteix de diferents parts: First, von, Last, Jr. El token *von* o *de la* cal posar-los en minúscules. Per tal que el nom es reconegui, cal que tingui el format: von Last, Jr, First. D'aquesta manera, si hi ha més d'un cognom no passa res.
- Caràcters Unicode entre claus per poder ser utilitzats correctament per l'estil *alpha*. Per exemple `Jos{\'}{e}`
- Per prevenir que BibTeX canviï un text a minúscules, cal posar el text entre claus.
- Si hi ha massa autors, truncar la llista amb *et al.*
- Utilitzar abreviatures de tres lletres per als mesos
- Utilitzar el camp **key** per a organitzacions amb un nom llarg, de manera que s'utilitzin les inicials de l'organització al fer una cita

2 PDF

Si un fitxer PDF ha estat generat utilitzant el paquet **hyperref** o similar, podem obtenir les següents dades del fitxer sense haver-ne d'extreure el text: número de pàgines, títol, autor, assumpte i paraules clau. Si aquests camps no s'han omplert al generar el fitxer, estaran en blanc.

2.1 Aspectes a considerar

- Caràcters Unicode
- Glyphs com ara *fi* corresponen a més d'una lletra

2.2 Text extraction

2.2.1 PDFBox

2.2.2 pyPDF

2.2.3 PDFMiner

2.2.4 jPod

3 Articles Db

3.1 DBLP

DBLP proporciona un

3.2 Portal ACM

3.3 Altres

- CiteSeerX
- Arxiv.org <http://arxiv.org/help/api/index>

4 Software existent

http://en.wikipedia.org/wiki/Comparison_of_reference_management_software

4.1 JabRef

- Java
- Llicència: GPL
- És possible crear plug-ins amb *Java Plug-in Framework*
- <http://jabref.sourceforge.net/>

4.2 Mendeley

- Propietari
- No accepta plug-ins
- <http://www.mendeley.com/>

4.3 BibDesk

- C, Objective C
- Llicència: BSD
- <http://bibdesk.sourceforge.net/>