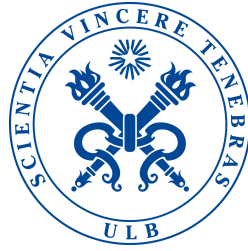


UNIVERSITÉ LIBRE DE BRUXELLES



INFO-H-509 : XML AND WEB TECHNOLOGIES

---

## Project 2 : XSLT

---

*Authors:*

Yasin ARSLAN

Jacky TRINH

*Professor:*

Stijn VANSUMMEREN

*Assistant:*

Dhananjay IPPARTHI

# Contents

<b>1</b>	<b>Project 2</b>	<b>2</b>
1.1	Presentation . . . . .	2
1.2	Hypotheses . . . . .	2
1.3	Conclusion . . . . .	3

# Chapter 1

## Project 2

### 1.1 Presentation

As part of the course INFO-H-509: XML and Web Technologies, we were asked to write a single XSLT 2.0 stylesheet that generates a HTML file for each author and editor from a small excerpt of the DBLP website.

### 1.2 Hypotheses

- We did not take into account the information that was not present in the excerpt such as the *www* type of record,... Reasons are that it was either something irrelevant or something that we did not know how we should use it (for example, the chapter field was not in the excerpt and we did not find if it should be used or not in some dblp example).
- So far from what we saw, the url from a record can either be a valid one or a local one. A valid url can direct us in an online website whereas a local url will only work if we possess the local files (see the dblp pdf for further information). To remediate this problem, when we detect that an url is a local one (it does not contain the "http" part), we add the "*http://dblp.uni-trier.de/*" prefix to the local url to make it works.
- For the publication reference, we tried to do the same format as in the dblp website. For each type of record, we checked on their website for a specific author from our excerpt and tried to mimic their format with our limited information. Indeed, for example they show the pages number from a phd thesis but there wasn't any information about pages in a phd thesis in our excerpt, so we did not know if all phd thesis were like this or not and therefore we did not take into account the pages information in this case. The format are thus different depending on the type

of record we are facing and the available information.

- We assume that the electronic edition is always in the *ee* field, therefore if there is an *url* field, it is for additional information used for linking to a journal, for linking to extra content in a book or proceedings with the [content] tag or for linking to extra information with the school information (mastersthesis and phdthesis).

## 1.3 Conclusion

With this project, we learned how to implement a XSLT stylesheet that could easily parse a XML file and create many HTML file with all the information parsed. We were somewhat surprised at how fast it was actually done (parsing through many line of data and creating a thousand of files).