

Advanced Data Modeling

Summer Semester 2009

- Exercises VIII -

*To be handed in before **2009-07-14, 10:00** via e-mail to bercovici@uni-koblenz.de And dividino@uni-koblenz.de, subject line: [ADM] ...*

Please send your eclipse project and your output file.

1. De.li.ci.ous Dataset

In the HDFS you can find some files about de.li.ci.ous users. The file describe user and their bookmarks web sites(url). For each bookmark you can find a timestand representing the date when the user has bookmarket it and the respectives tags he has associated to it.

The files are organized as following:

[userid] | [url] | [timestamp] | [tag]

2. Trying Out

- a. Write a Java code for Haddop that returns a list of timestamp and urls that were bookmarked at the same time. (5 points)
- b. Write a Java code for Hadoop that return a list of tags and the url that we tagged with them. (5 points)