# The exploitation of social tagging systems in libraries

**Constantia Kakali, Christos Papatheodorou**

Database & Information systems group

Laboratory on Digital Libraries and Electronic Publishing

Department of Archive and Library Sciences, Ionian University, Corfu, Greece

# Contents

- Social tagging
- Limitations of Folksonomies
- Social tagging systems in libraries
- Opacial system
- Preliminary experimental study
- Tag analysis and tag estimation
- Tag methodology proposal
- Policy issues

# Social tagging

- The social tagging has grown in popularity on the web-based services
  - Add bookmarks of sites they like to their personal collections of links
  - Organize and categorize these sites by adding their own terms, or tags
  - Share this collection with other people with the same interests
- Tags are metadata elements described resources and applied by users.
- The set of the tags introduced for a resource is called *folksonomy*, which it could be presented as a *tag cloud*

# Limitations of folksonomies

- Synonyms
- Homonyms
- Polysemy
- No defined relations (e.g., broader, narrower, and associative terms).
- Ampiquity
- Specificity
- Syntax

# Social tagging systems in libraries

- Penntaggs
- New opacs 2.0 (Scriblio, VuFind, Primo, Aquabrowser)
- Mtagger
- LibraryThing

# The OPACIAL System (1/2)


Developed by Panteion University Library, Athens, Greece, OPACIAL is an Online Public Catalog System which uses Web 2.0 technologies. It utilizes library catalogs using the Z39.50 protocol and it can also connect to DSPace digital repositories.

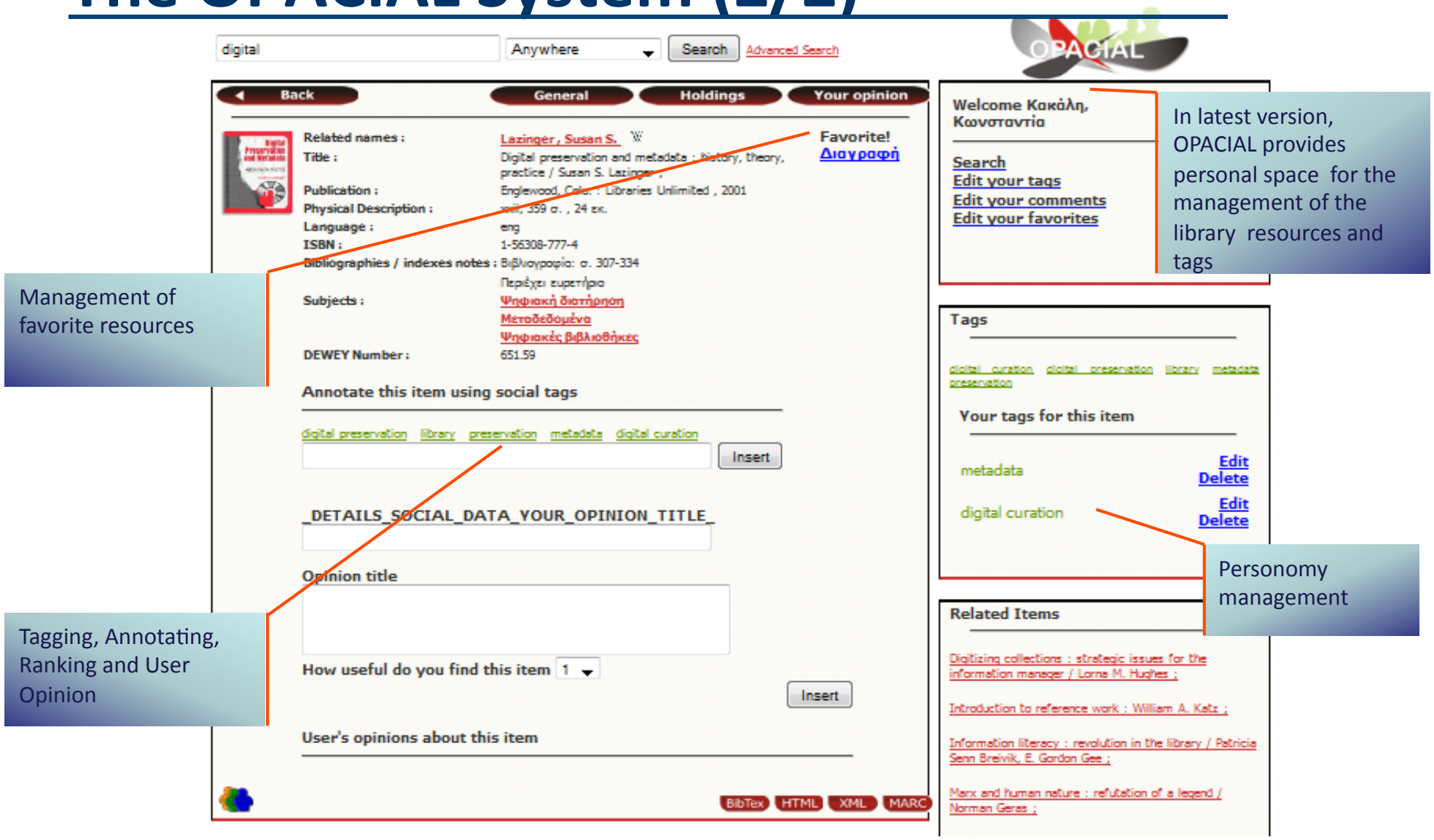
**http://opacial.sourceforge.net**

Opacial functionalities:

- Users insert tags, annotations
- Faceted navigation by social tags, subject headings, item type, year
- Integration with the University's digital repository, named Pandemos


DBIS database & information systems group ionian university

# The OPACIAL System (2/2)




Management of favorite resources

Tagging, Annotating, Ranking and User Opinion

In latest version, OPACIAL provides personal space for the management of the library resources and tags

Personomy management

# Preliminary experimental study

- Experimental study
  - Evaluation by 20 users
  - Scenario
  - Inserted more than 500 tags
  - Interviews
- Criteria
  - Usability (learn ability, navigation, information architecture, aesthitics
  - Usefulness (relevance, reliability, format, timeliness)
  - Satisfaction

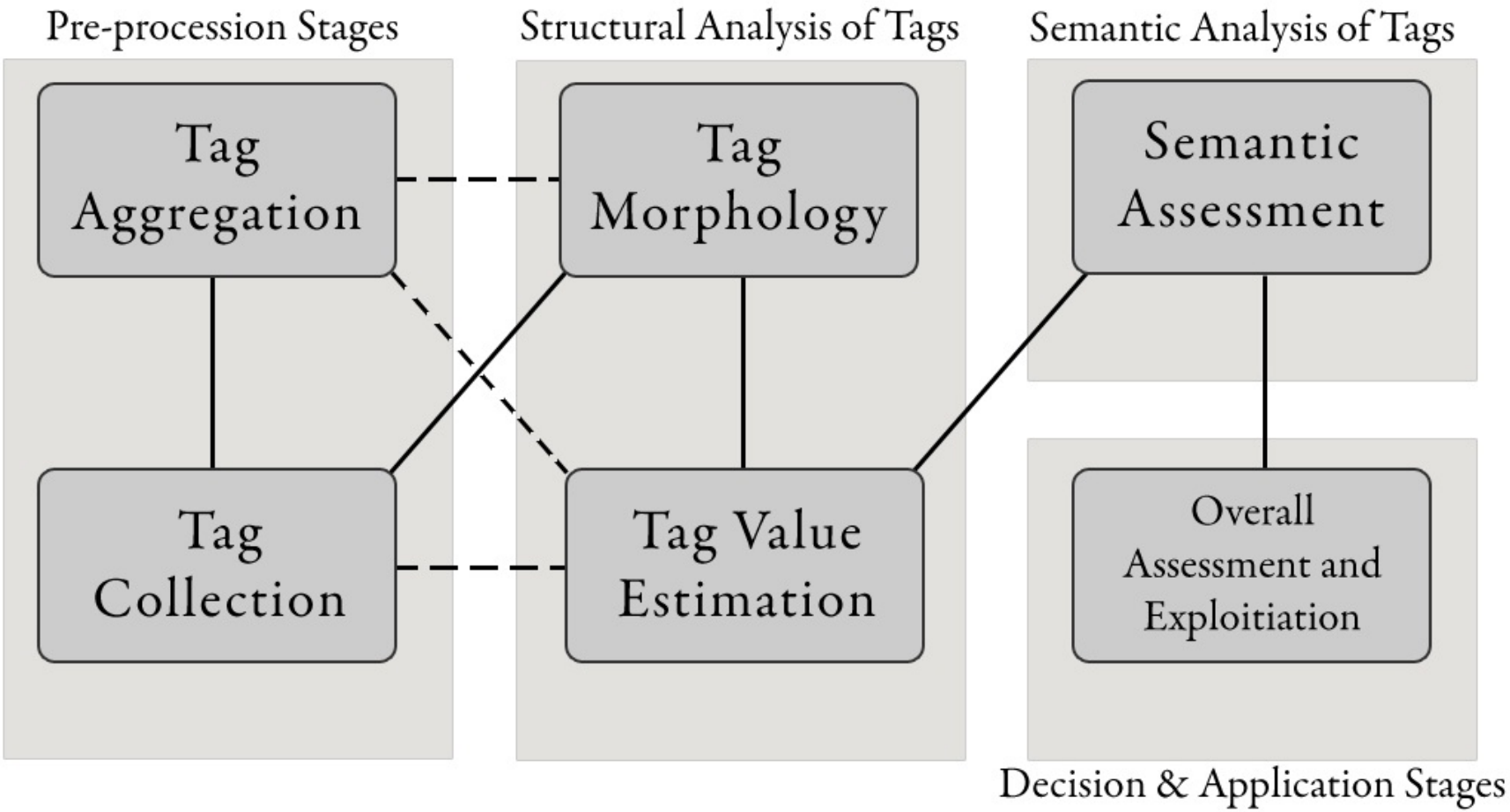
# Results

- Social tagging is found as a usable and useful functionality
  - Users use it either to describe more precisely the OPAC records, or to correct a wrong subject term
- The co-existence of subject index and folksonomy improve the information seeking effectiveness
- Users suggested to add tags in order to create folksonomies and created bibliography lists
- Finally found OPACIAL quite satisfactory and the usability criteria (navigation & aesthetics) are highly rated.

# Tag analysis research

- The stages of the proposed methodology decompose the general problem to the following research questions:
  - Is it possible to categorize the social tags in order to investigate the users' subject indexing trends?
  - Do the social tags contribute to the improvement of the subject index of an information organization and how?
  - How expert-based and user-based subject indexing processes could be integrated in a symbiotic, cooperative environment?

# Tag analysis methodology



# Collecting Tags and Studying Their Morphology

- Morphology as lexical units (single words or phrases, language, abbreviations, plural or singular, orthographical errors, names or abstracts entities).
- Table 1 shows that as the number of subject headings per record increases, the number of the tags decreases. This result confirms the assumption that tagging plays an enhancing role to weak subject descriptions.

Table 1: Cross-reference of the number of subject headings over the tagged records

| | | Number of Subject Headings | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | | **0** | **1** | **2** | **3** | **4** | **5** | **6** | **8** | **Total** |
| **Number of tags** | **1** | 2 | 43 | 60 | 57 | 45 | 3 | 4 | 1 | 215 |
| | **2** | 1 | 13 | 19 | 8 | 8 | 5 | 3 | 0 | 57 |
| | **3** | 0 | 18 | 22 | 10 | 7 | 4 | 5 | 0 | 66 |
| | **4** | 0 | 12 | 17 | 20 | 2 | 0 | 0 | 1 | 52 |
| | **5** | 0 | 5 | 20 | 8 | 0 | 2 | 0 | 0 | 35 |
| | **6** | 0 | 5 | 12 | 8 | 4 | 1 | 0 | 0 | 30 |
| | **7** | 0 | 8 | 5 | 4 | 4 | 0 | 0 | 0 | 21 |
| | **8** | 0 | 2 | 5 | 7 | 2 | 0 | 0 | 0 | 16 |
| | **9** | 0 | 6 | 4 | 6 | 2 | 0 | 0 | 0 | 18 |
| | **10** | 0 | 3 | 3 | 4 | 0 | 0 | 0 | 0 | 10 |
| | **11** | 0 | 10 | 10 | 6 | 7 | 0 | 0 | 0 | 33 |
| | **14** | 0 | 1 | 5 | 4 | 4 | 0 | 0 | 0 | 14 |
| | **15** | 0 | 4 | 7 | 2 | 2 | 0 | 0 | 0 | 15 |
| | **Total** | 3 | 130 | 189 | 144 | 87 | 15 | 12 | 2 | 582 |

# Estimation of the Tags' Value

- 582 tags revealed the following types of tagging behaviors:
  - enhance / refine the thematic description of a record, manifested by two partial behaviors (80%):
    - uses terms that belong in the authority file as tags but not for the particular record
    - adds new terms, disjoint from the authority file descriptors.
  - correct the thematic description of a record, to propose more accurate terms for the particular bibliographic records (2,1 %).
  - both of them, the insertion of new terms, disjoint to the local authority terms, expressing new concepts or synonyms (3,8 % )
  - without significant value (13,4 %)
- 46.2% (269 tags) of the total amount of tags is not present in the existing authority file

# Folksonomy and KOS : overlap matching

| Number of tags that exist in other KOS (percentages inside parentheses) | | | | | |
|---|---|---|---|---|---|
| | **LCSH Authority** | **NDC Thesaurus** | **SSIT Thesaurus** | **Wikipedia** | **WordNet** |
| **Exist exactly** | 76 (28.3) | 26 (9.7) | 35(13.0) | 166 (61.7) | 26 (9.7) |
| **Not exist** | 176 (65.4) | 229 (85.1) | 234 (87.0) | 66 (24.5) | 243 (90.3) |
| **Exist by Reference** | 17 (6.3) | 14 (5.2) | - | 37 (13.8) | - |
| **Total** | 269 (100.0) | 269 (100.0) | 269 (100.0) | 269 (100.0) | 269 (100.0) |

- 5 systems were selected with criteria: coverage, language and relevance
- LC authorities cover the 34.6% of the 269 “missing tags” (28.3% in main entries and 6.3% as non-preferred terms)
- 166 of the tags (61.7%) existed as articles in Wikipedia, which follows closely the vocabulary evolution of the scientific communities

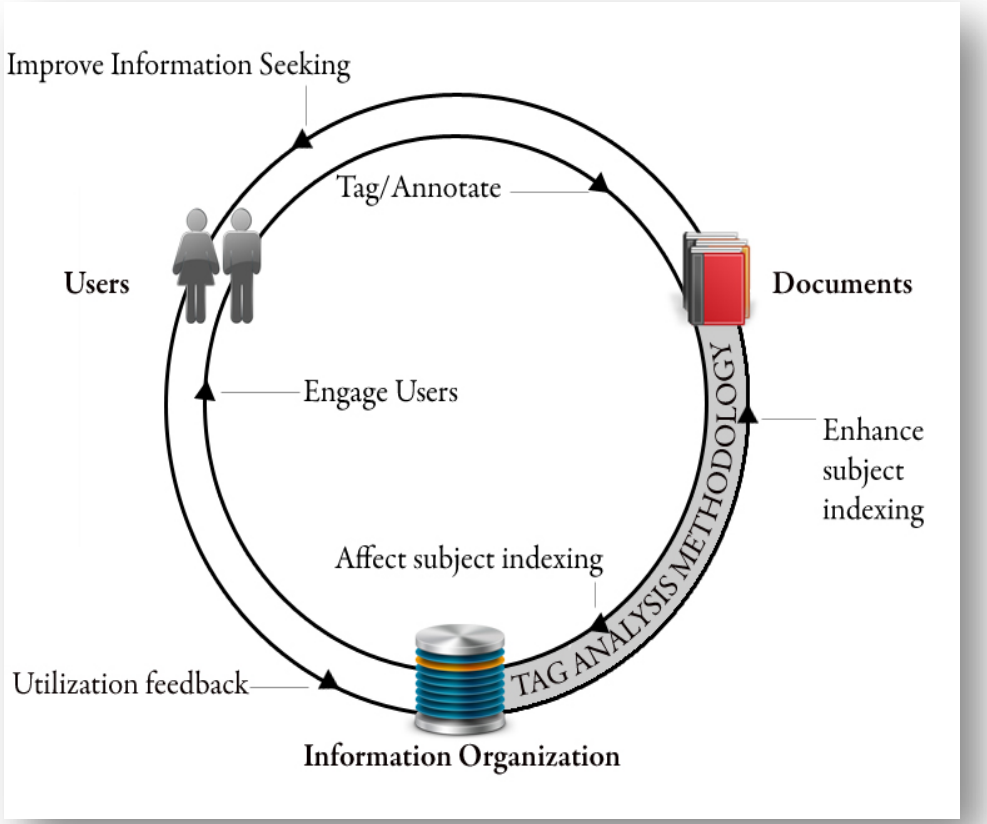
# Folksonomy and KOS : semantic assesment

- Investigation of the semantic relation between the folksonomy tags and the local authority file terms
  - For each record $e_k$ we observe the semantic relation $r_{ij}$ between the pairs ($t_i$ and $s_j$ ) that correspond to the tags ($t_i$) and the subject headings ($s_j$) used for the thematic description of $e_k$
  - 1420 pairs ($t_i$, $s_j$) were generated, 1125 are unique.
- *Example*: the pair 'archetype' – 'Symbolism (Psychology)',

  in LC authorities : subject heading 'Archetype (Psychology)'
  associative relation 'Symbolism (Psychology)'.
- The majority of the pairs are not correlated in any KOS (60.6%)
- Once more, Wikipedia includes the majority of the correlated pairs.

# Resuming tag exploitation

- (a) the macro decision-making level




- (b) the micro decision making level:
  - particular actions and tasks regarding the inter-relations of tags and headings
  - Three steps proposed:
    - Creating new subject headings (new terminology)
    - Correction the thematic description
    - Creation of rich references

# Continuing the research

- Does social tagging upgrade the library services?
- Should the library develop a policy for encouraging the insertion of social tags?
- How we could incorporate tags in authorities?
- Investigate the acceptance of social tagging by the librarians ?
  - Could be exploited to improve the subject indexing of an academic library collection?
  - Assess the semantic value of the tags with respect to the subject headings.

# New experimental setup

- A sample of 30 bibliographic records holdings:
  - 72 subject headings, 66 being unique
  - 18 records in English, 1 in French,  11 in Greek
- 540 tags
  - 120 tags OPACIAL
  - 420 tags LibraryThing
- 9 subject cataloguers

# General Observations

- OPACIAL has more representative and accurate tags than LibraryThing (60% vs 40%).
- OPACIAL hosts material focusing on social sciences and serves a scholar community that uses a specialized vocabulary.
- LibraryThing is a general-purpose collaborative cataloguing service, has too general terms.
- Opacial tags usually are narrower terms and Library Thing usually are broader terms.

# Analysing the results

- Only 21 tags are identical to the subjects
- 355 of tags exist in the library authorities
- *Example*: In the record below, 34 tags, 28 of them exist in the library's authority file, 11 could be used in the specific thematic description , 2 new terms

| Bibliographic Record | Subject Headings | Tags |
|---|---|---|
| **Author**: Weber, Max (1864-1920), Roth, Guenther (Editor), Wittich, Claus (Editor). **Title**: Economy and society : an outline of interpretive sociology / Max Weber; edited by Guenther Roth and Claus Wittich **Publication**: Berkeley, Calif. : University of California Press, c1978 | Sociology<br><br>Economics | 19th century 20th century Europe Germany Verstehen Weber bureaucracy class structure economic sociology economics economy german history interpretation knowledge philosophy political economy political science political theory politics religion social theory society sociological theory sociology state the state theory world history Αξιολογική Ελευθερία Γερμανοί Φιλόσοφοι Κατανόηση Κοινωνιολογία |

# Recommendations for Policy

- Activate user participation through the information literacy program.
- Create groups of faculty members and post graduate.
- Create a wiki for the communication between specialized groups - subject librarians
- Use tags as non-preferred terms and add relations to the corresponding preferred.
- Create authority records “out of the LCSH frontiers”, especially in cases the tags are identical to Wikipedia entries.

# Further research issues

- Comparative analysis of folksonomies with other knowledge organizations systems
- Analyzing user seeking behavior in a opac enhanced with a tagging functionality.
- And finally, why users insert tags in resources
  - Information retrieval aspects
  - User communication
  - Scholarly terminology evolution

- Thank you for your attention !
- Questions