LOOKING FOR HEALTH INFORMATION ON THE INTERNET: CAN SOCIAL BOOKMARKING SYSTEMS REPLACE EXPERT GATEWAYS?

V. Durieux[a], G. Kerdelhué[b]

[a] *ReSIC, Université Libre de Bruxelles, 1050 Brussels, Belgium*
[b] *CISMeF, Rouen University Hospital, 76031 Rouen, France*
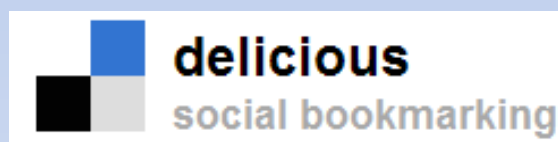
Introduction

- Internet = first information provider

- Weaknesses:
  – Organisation
  – Validation

- Alternatives to search engines:
  – Expert gateways (CISMeF)
  – Social bookmarking systems (*Delicious*)

Introduction - CISMeF

- French acronym for Catalogue and Index of French Language Health Resources on the Internet

- Quality controlled health gateway

- Standard tools to organize information:

  – MeSH Thesaurus (NLM) in French

  – Dublin Core Metadata format

- Currently : 67,000 indexed resources

Introduction - *Delicious*

- Social bookmarking service (2003) [Del.icio.us]

- To store, organize and share bookmarks using tags (natural language keywords)

  = collaborative tagging

- Emergent list of tags = folksonomy

  ("folk taxonomy")
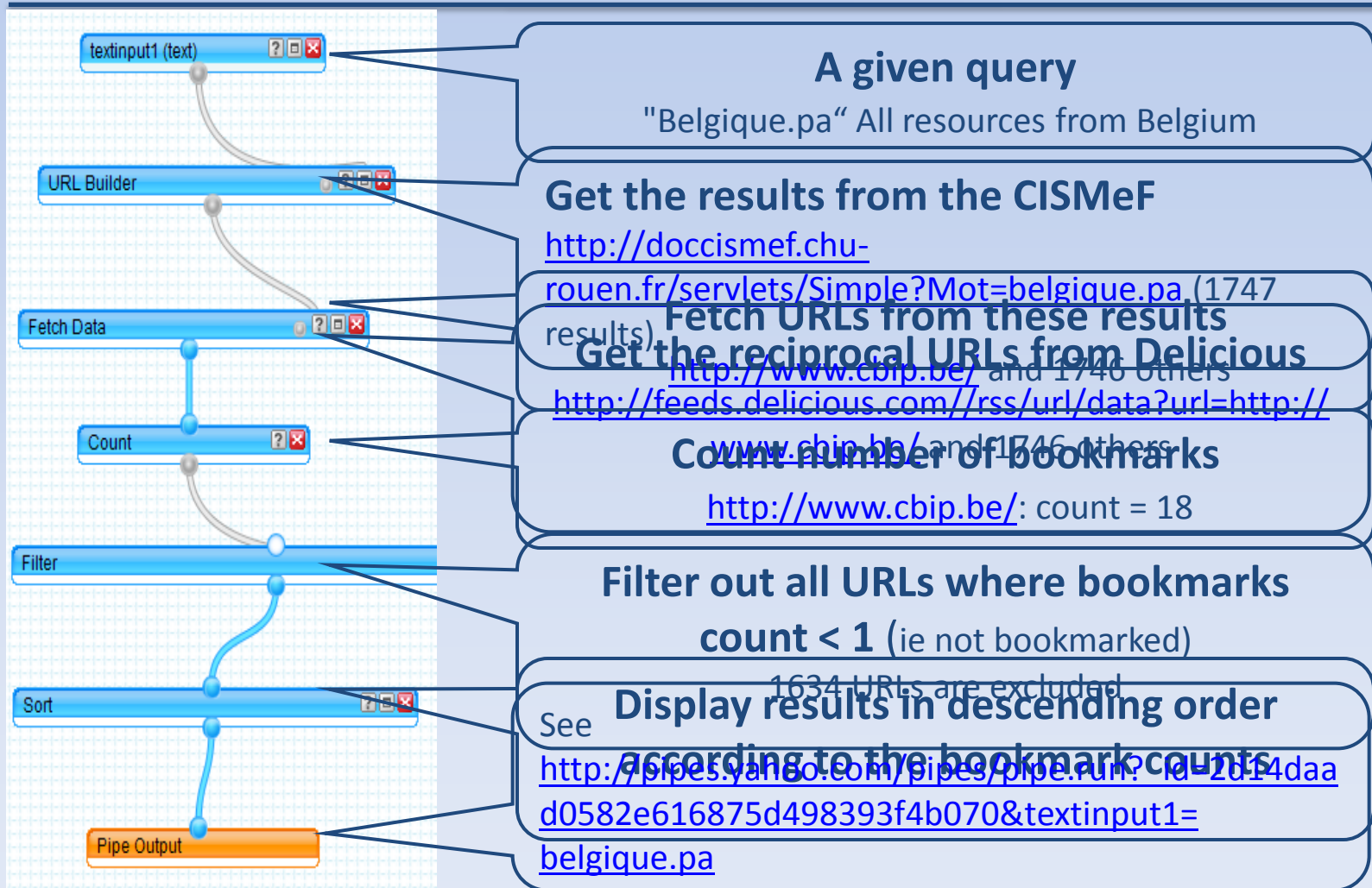
Research questions

- To what extent do healthcare resources listed in CISMeF overlap those available in Delicious?

- To what extent do tags in Delicious differ from descriptors in CISMeF for the same content?

- To what extent do users' tags provide added value to traditional descriptors in the information retrieval process?

CiSMeF

Materials & Methods

- Sample :
  - Healthcare resources from Belgium
  - Matching the query "belgique.pa" in CISMeF

- Identification of resources listed in both systems
  - Mashup: retrieves data from CISMeF and Delicious, and aggregated results

    - Using Yahoo! Pipes[TM] http://pipes.yahoo.com/
      (A graphical user interface for building data mashups from Web sources)

Materials & Methods - Mashup

A given query

"Belgique.pa" All resources from Belgium

Get the results from the CISMeF

http://doccismef.chu-rouen.fr/servlets/Simple?Mot=belgique.pa (1747 results)

Fetch URLs from these results

Get the reciprocal URLs from Delicious

http://www.cbip.be/ and 1746 others

http://feeds.delicious.com//rss/url/data?url=http://www.cbip.be/ and 1746 others

Count number of bookmarks

http://www.cbip.be/: count = 18

Filter out all URLs where bookmarks count < 1 (ie not bookmarked)

1634 URLs are excluded

Display results in descending order according to the bookmark counts

See http://pipes.yahoo.com/pipes/pipe.run?_id=2d14daad0582e616875d498393f4b070&textinput1=belgique.pa

CiSMeF

CiSMeF

+ belgique.pa [Recherche]

- **http://www.cfwb.be/**

  http://delicious.com/url/fda1da14ad13c90b5735820a8b696632
  Bookmarks = 19

- **http://www.itg.be/itg/**

  http://delicious.com/url/eb3c585a41de6ee4fb4dd57d76593122
  Bookmarks = 19

- **http://www.cbip.be/**

  http://delicious.com/url/5541715e7b5b4d42f07b76410c27bfa0
  Bookmarks = 18

- **http://allserv.rug.ac.be/~rvdstich/eugloss/welcome.html**

  http://delicious.com/url/3f2b3b7fd9a62442a7fe460dd9785b6f
  Bookmarks = 16

- **http://www.ulb.ac.be/**

  http://delicious.com/url/3edb99a2da785a500c99c96e37cd3ba1

Materials & Methods
Resources overlap

- Required data:
  - Number of healthcare resources from Belgium listed in Delicious
    - To be determined
  - Number of healthcare resources from Belgium coded in CISMeF
    - Provided by the Mashup
  - Number of these resources listed in both systems
    - Provided by the Mashup

ULB

CiSMeF

Materials & Methods
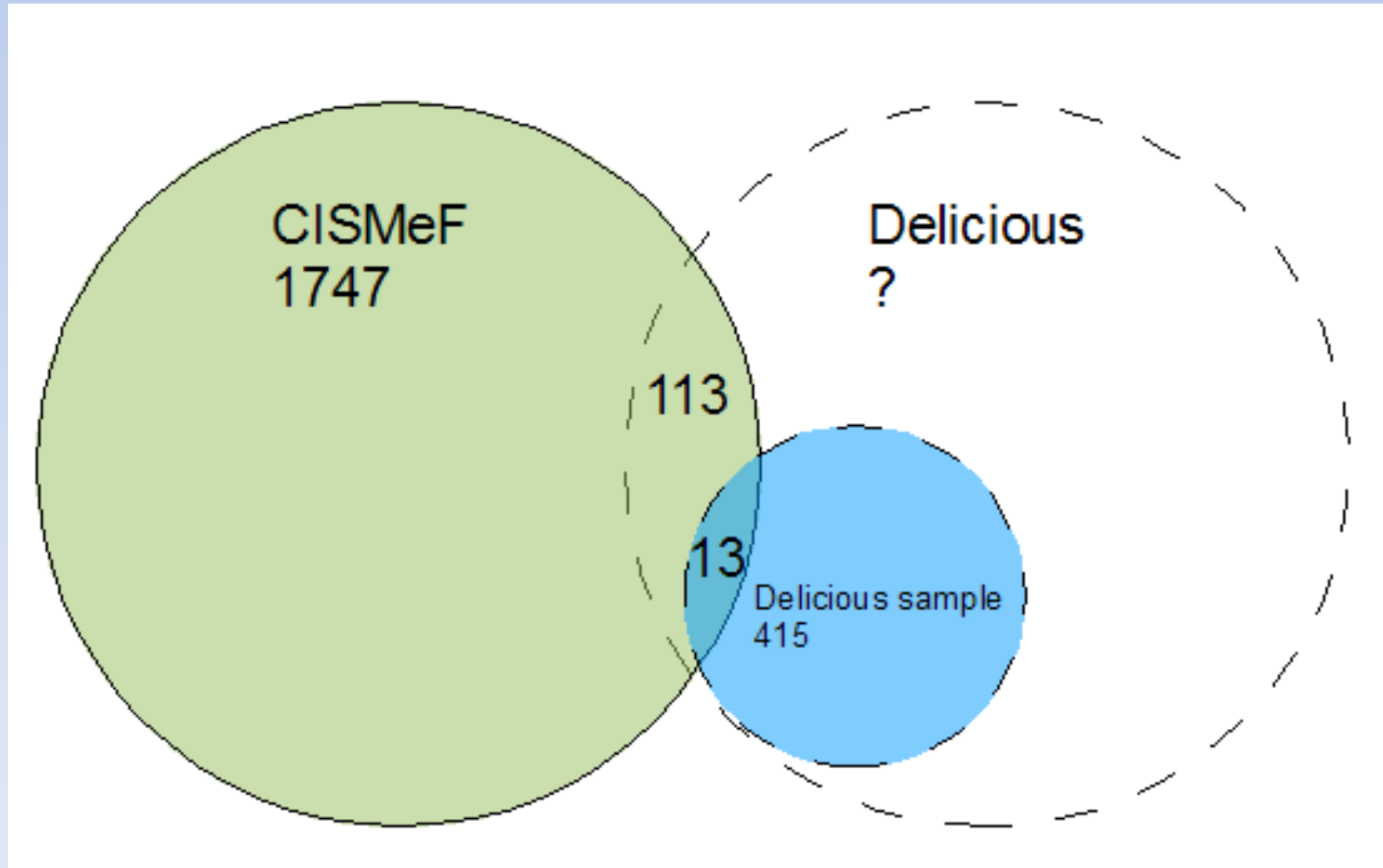## Resources overlap

- Number of healthcare resources from Belgium listed in Delicious

  – can neither be precisely determined, nor even estimated

- Identification of a sample:

  (belgium OR belgique OR brussels) AND

  (health OR santé OR medicine OR médecine OR medical)

  ➢ 415 resources founded

ULB

CiSMeF

Materials & Methods
Term comparison

- Kipp's seven-point scale:

    1. Same
    2. Synonym
    3. Broader Term
    4. Narrower Term
    5. Related Term
    6. Related
    7. Not Related

ULB

CiSMeF

Results - Resources overlap

Results - Resources overlap

- Very poor resources overlap (6.5%)

- 85% of common resources = website homepages

- Website homepages = only 18.5% of all CISMeF URLs

- ➢ CISMeF librarians tend to index deeper links and thus more specific documents than the Delicious users

ULB

CiSMeF

Results - Resources overlap

- only 4 articles listed in Delicious out of the 1228 indexed in CISMeF

➤ Delicious users tend to focus on more general and popularized resources, while CISMeF on scholar content

Conclusion:

➤ Complementary as providers of healthcare resources from Belgium

Results - Descriptive statistics

- 113 resources listed in both systems

- 7 resources excluded (not tagged)

- 288 posts

- 230 unique Delicious users

- 86% of users tagged only 1 resource

- 49% of resources tagged by only 1 user

- 561 librarian's descriptors (231 unique)

- 747 users' tags (431 unique)

Results - Term comparison

Same	Synonym	Broader Term	Narrower Term	Related Term	Related	**Not Related**
201 (27%)	20 (3%)	118 (16%)	16 (2%)	21 (3%)	114 (15%)	**257 (34%)**

⚠️ "NR" does not mean useless or meaningless

ULB   CiSMeF

Results - Term comparison

Same	Synonym	Broader Term	Narrower Term	Related Term	Related	Not Related
201 (27%)	20 (3%)	118 (16%)	16 (2%)	21 (3%)	114 (15%)	257 (34%)

Thesaural relations (51%)

- 85% of "TR" = "Same" and "Broader Term"
- Most common "BT" = "health" and "medicine"
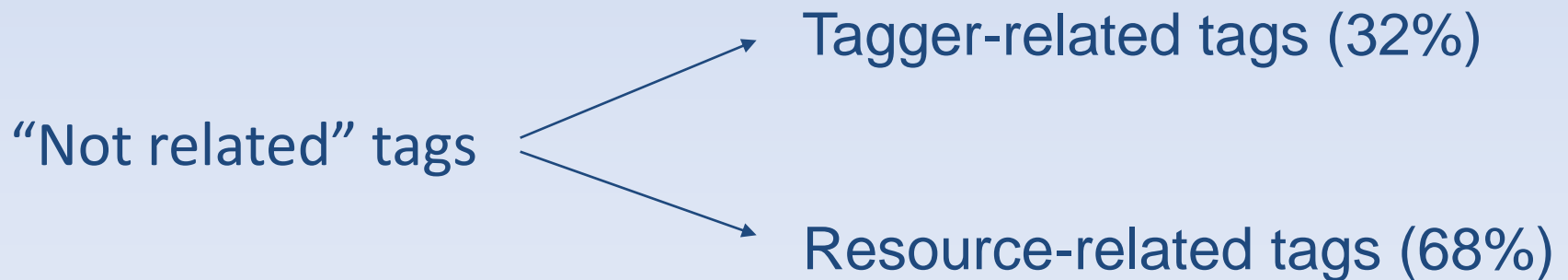
Results - Term comparison

Same	Synonym	Broader Term	Narrower Term	Related Term	**Related**	Not Related
201 (27%)	20 (3%)	118 (16%)	16 (2%)	21 (3%)	**114 (15%)**	257 (34%)

- Additional access points to a resource
- Ex.: "government" assigned to governmental web site of the Social Security

ULB

CiSMeF

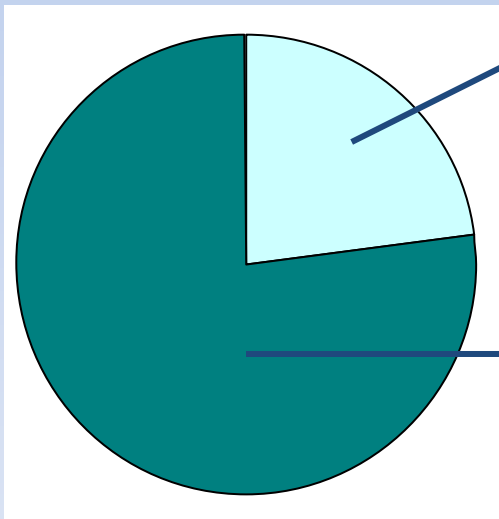# Results - Term comparison

Same	Synonym	Broader Term	Narrower Term	Related Term	Related	**Not Related**
201 (27%)	20 (3%)	118 (16%)	16 (2%)	21 (3%)	114 (15%)	**257 (34%)**

"Not related" tags → Tagger-related tags (32%)

"Not related" tags → Resource-related tags (68%)

ULB

CiSMeF

Results - Term comparison

Tagger-related tags
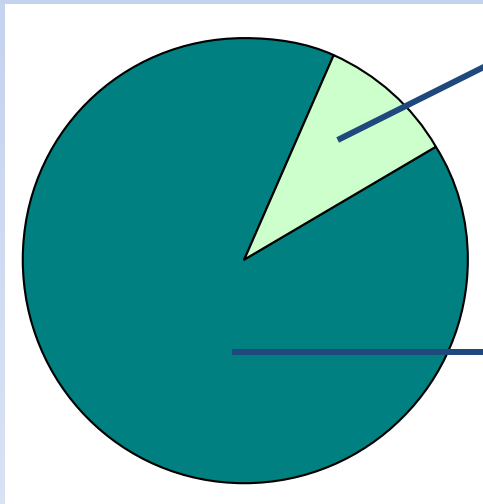

Time & Task Management (23%)
- Ex.: "work", "thesis", etc.

Self Signification (77%)
- Ex.: "question_21_incitervisites"

➢ Meaningless and useless for any user other than the tagger

ULB

CiSMeF

Results - Term comparison

Resource-related tags

Qualities & characteristics (10%)

- Type ("guide")
- Language ("FrenchTerm")
- Characteristic ("interesting")

Subject-related (90%)

- Acronyms
- Proper nouns

➢ Additional access points to a resource

ULB

CiSMeF

Results - Usefulness of tags

Same	Synonym	Broader Term	Narrower Term	Related Term	Related	Not Related
201 (27%)	20 (3%)	118 (16%)	16 (2%)	21 (3%)	114 (15%)	257 (34%)

Useful tags (88%)

Tagger-rel.	Resource-rel.
87 (32%)	170 (68%)

ULB

CiSMeF

Results - Usefulness of tags

Same	Synonym	Broader Term	Narrower Term	Related Term	Related	Not Related
201 (27%)	20 (3%)	118 (16%)	16 (2%)	21 (3%)	114 (15%)	257 (34%)

Additional access points

Tagger-rel.	Resource-rel.
87 (32%)	170 (68%)

Results - Usefulness of tags

- ± 25% of descriptors not represented by tags

➢ Numerous resources incompletely described by their assigned tags BUT still findable in Delicious

- Only 6 resources (0.5%) not findable by their assigned tags

  (Ex.: a resource with the only "mam" tag assigned)

Conclusions

- Low resources overlap between CISMeF and Delicious

➢ **Complementary as providers of healthcare resources from Belgium**

- Users tags identical to librarians descriptors, or more general

- Numerous additional access points provided by users tags

- ± 25% descriptors not represented by any tags

➢ **Users tags complement and even compete with librarian's descriptors but in no way could substitute for them**