

Helping online communities to semantically enrich folksonomies

Freddy Limpens, Fabien Gandon

Edelweiss, INRIA Sophia Antipolis
{freddy.limpens, fabien.gandon}@inria.fr

Michel Buffa

Kewi, I3S, Université Nice-Sophia Antipolis



Edelweiss



WebScience 2010, Raleigh, NC, USA, 26-27 april 2010

Social tagging for bookmarking

The screenshot shows the Delicious website interface. At the top, there is a navigation bar with the Delicious logo and menu items: Home, Bookmarks, People, and Tags. A search bar is located on the right side of the navigation bar. Below the navigation bar, there is a "Home" section with a sub-header "See what's hot on Delicious, or go to your Bookmarks" and a right-hand sidebar with links: "Save a new bookmark", "Go to a user", "Explore a tag", and "Look up a URL". The main content area has three tabs: "Fresh Bookmarks", "Hotlist", and "Explore Tags". Below the tabs, there is a text box explaining that tags are keywords used to describe bookmarks, followed by a "Learn more about tags" link and a "Type a Tag" input field with a search button. The main content area displays a list of bookmarks under the "java" tag. The first bookmark is "How to Write Google Android Apps" with 154 saves. The second bookmark is in Japanese: "『JavaRa』はJavaを更新し、古いバージョンを削除してくれるポータブルツール! : ライフハッカー [日本版]" with 15 saves. On the right side, there is a "Popular Tags" section listing "design", "blog", and "video".

delicious Home Bookmarks People Tags Search Delicious Search

Home
See what's hot on Delicious, or go to [your Bookmarks](#)

Save a new bookmark
Go to a user
Explore a tag
Look up a URL

Fresh Bookmarks Hotlist Explore Tags

Tags are keywords people use to describe their bookmarks
[Learn more about tags](#)

Type a Tag

java
See more

How to Write Google Android Apps SAVE 154

『JavaRa』はJavaを更新し、古いバージョンを削除してくれるポータブルツール! : ライフハッカー [日本版] SAVE 15

Popular Tags
design
blog
video

<http://delicious.com>

... and the resulting



FOLKS - ONOMIES

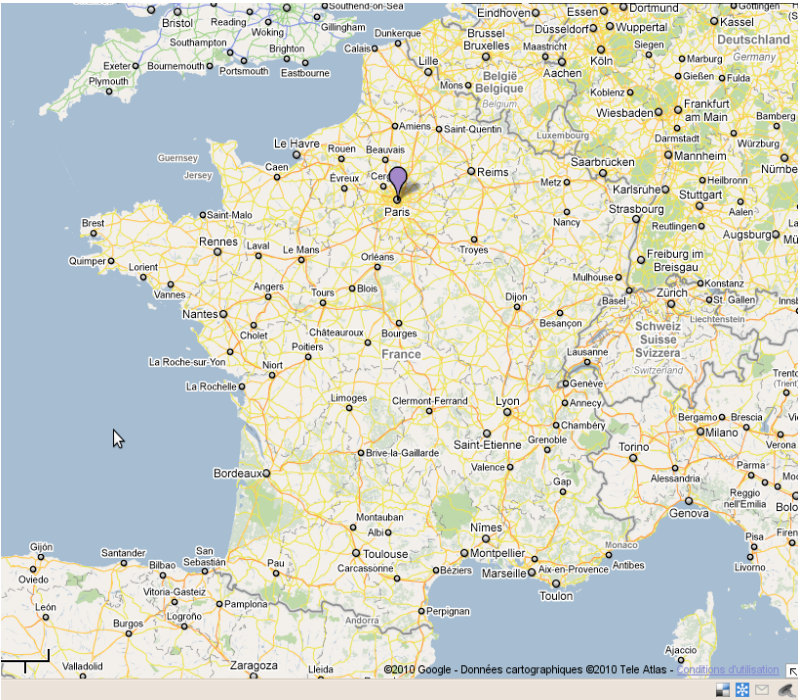
Spelling variations of tags

newyork = new_york = nyc

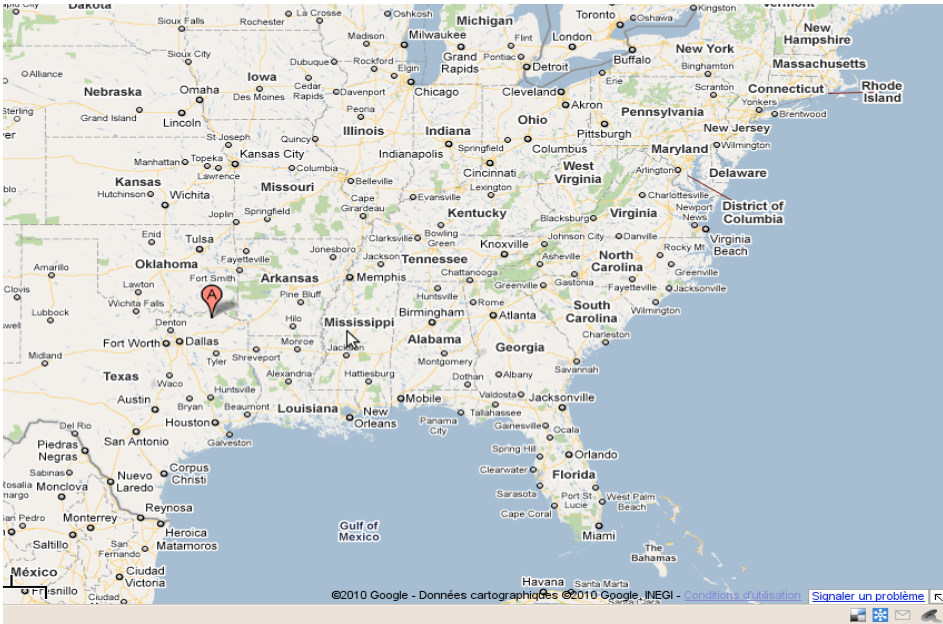
Amibiguity of tags



.. in
France ?

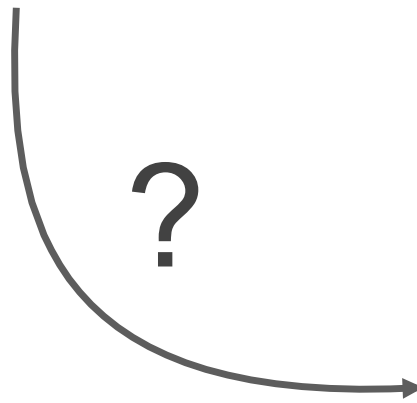


... or in
Texas ?

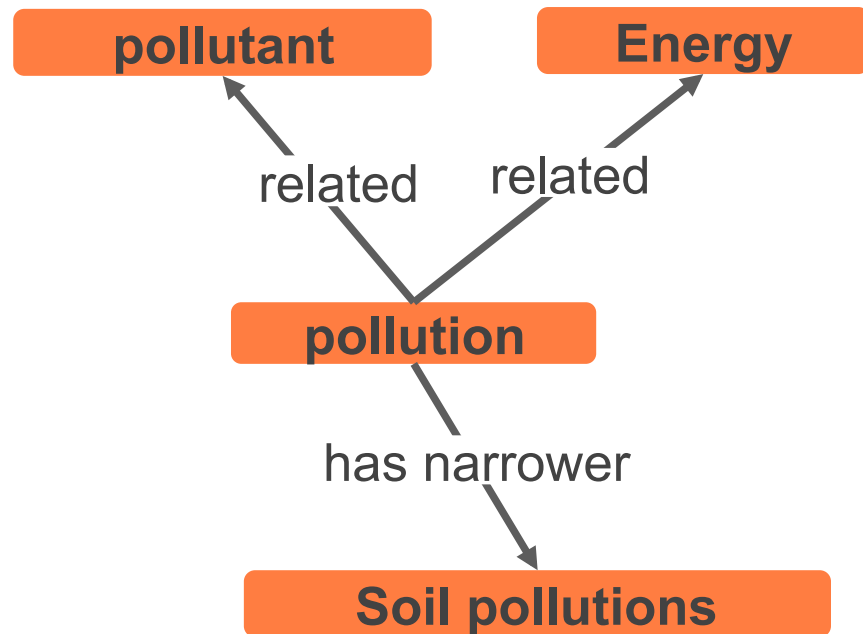


communciation communication compareteur
 compensation comportement condition
 conference confiance **consommation**
 contact convention cosmetique coton
 covoiturage creation critique CSS cuisine
 dechet dell demarche developpement document
 donnee durable dynamique eau ecoconception
 ecolife **ecologie** ecommerce economie
 elevage emballage **emission** empreinte encre
 encyclopedie **energie** enquete enseigne
 entreprise **environnement** eolien epargne
 ericsson espece ethique etiquette etude
europa evaluation evolution extensif faune
 federation fertilisation filiere film finance flore
 fondation forum **france** fruit GES
 goodguide grandedistribution graphique green
 greenpeace greenwashing grenelle guide
 habitat hitech HQE HTML hybride impact
 import imprimante inde indice industrie

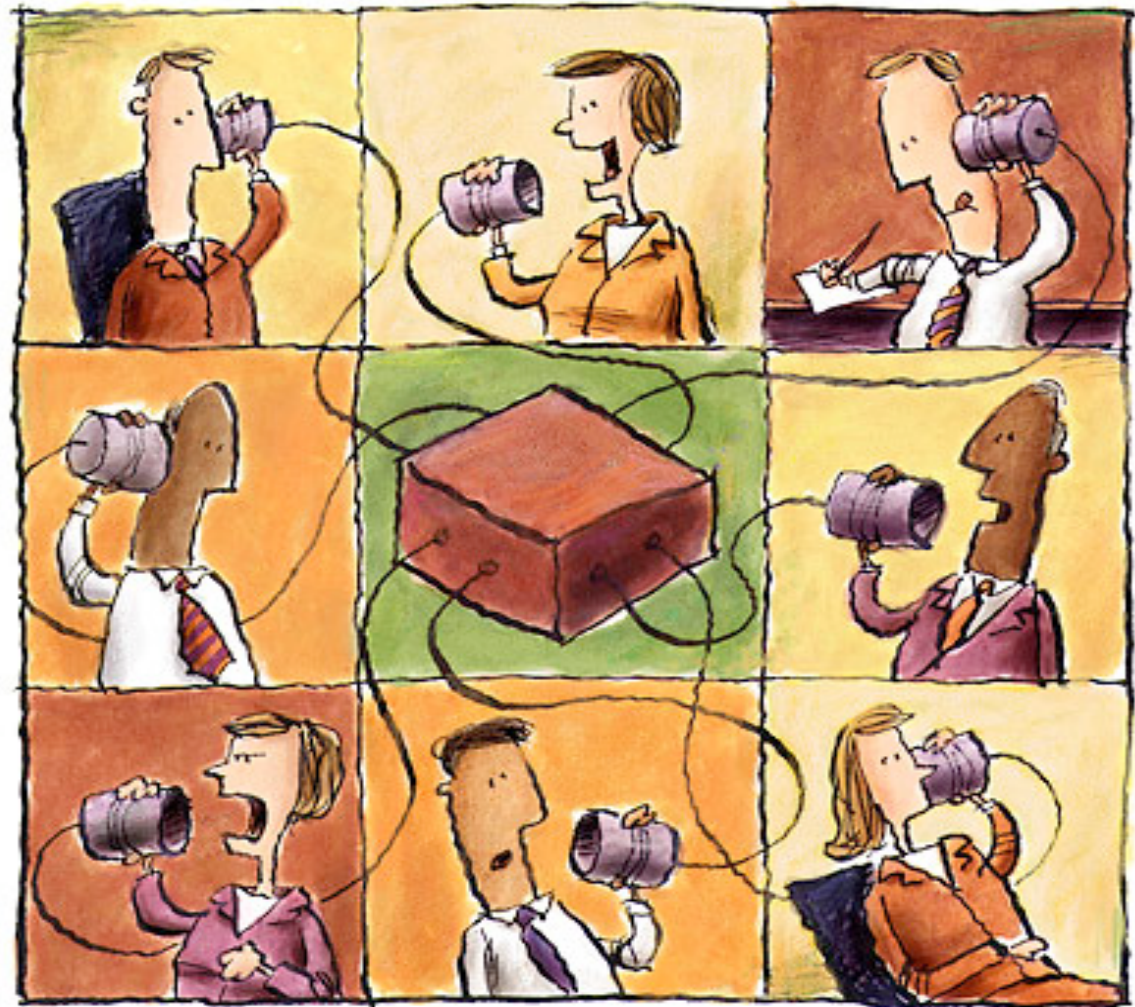
How to turn folksonomies ...



...into **comprehensible topic structures** ?



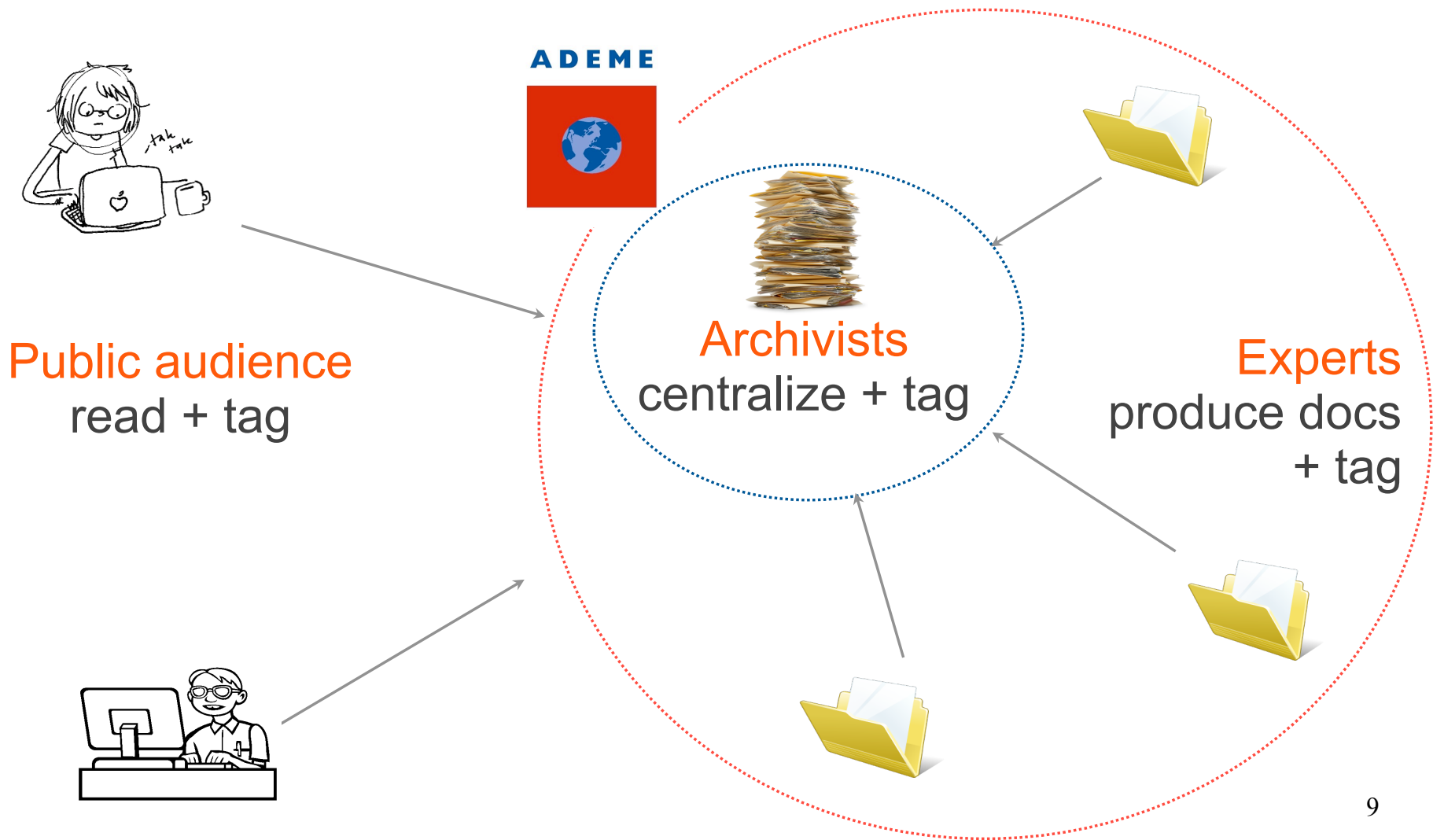
... by collecting
all user's
expertise
into the
process



Our approach

- Integrate usage-analysis for a tailored solution
- Supporting diverging points of view
- Automatic processings
+
- Human expertise through user-friendly interfaces

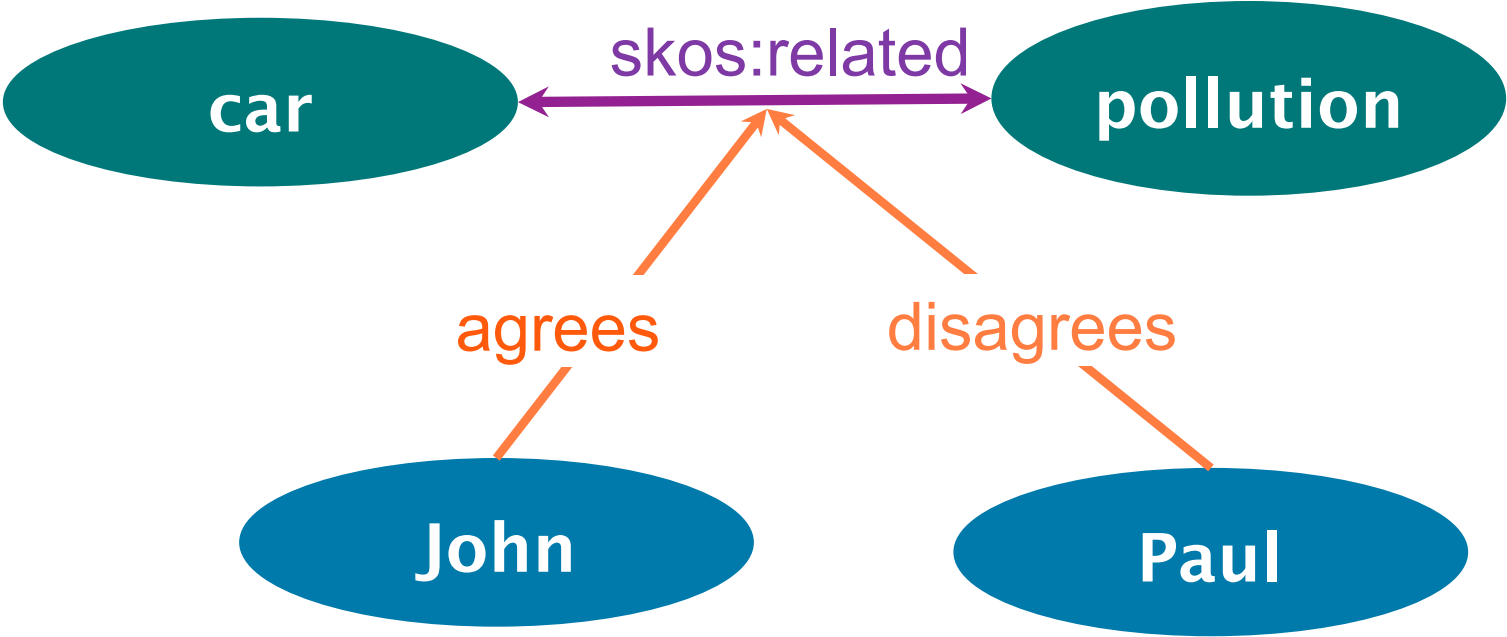
Concrete scenario



Supporting diverging points of view

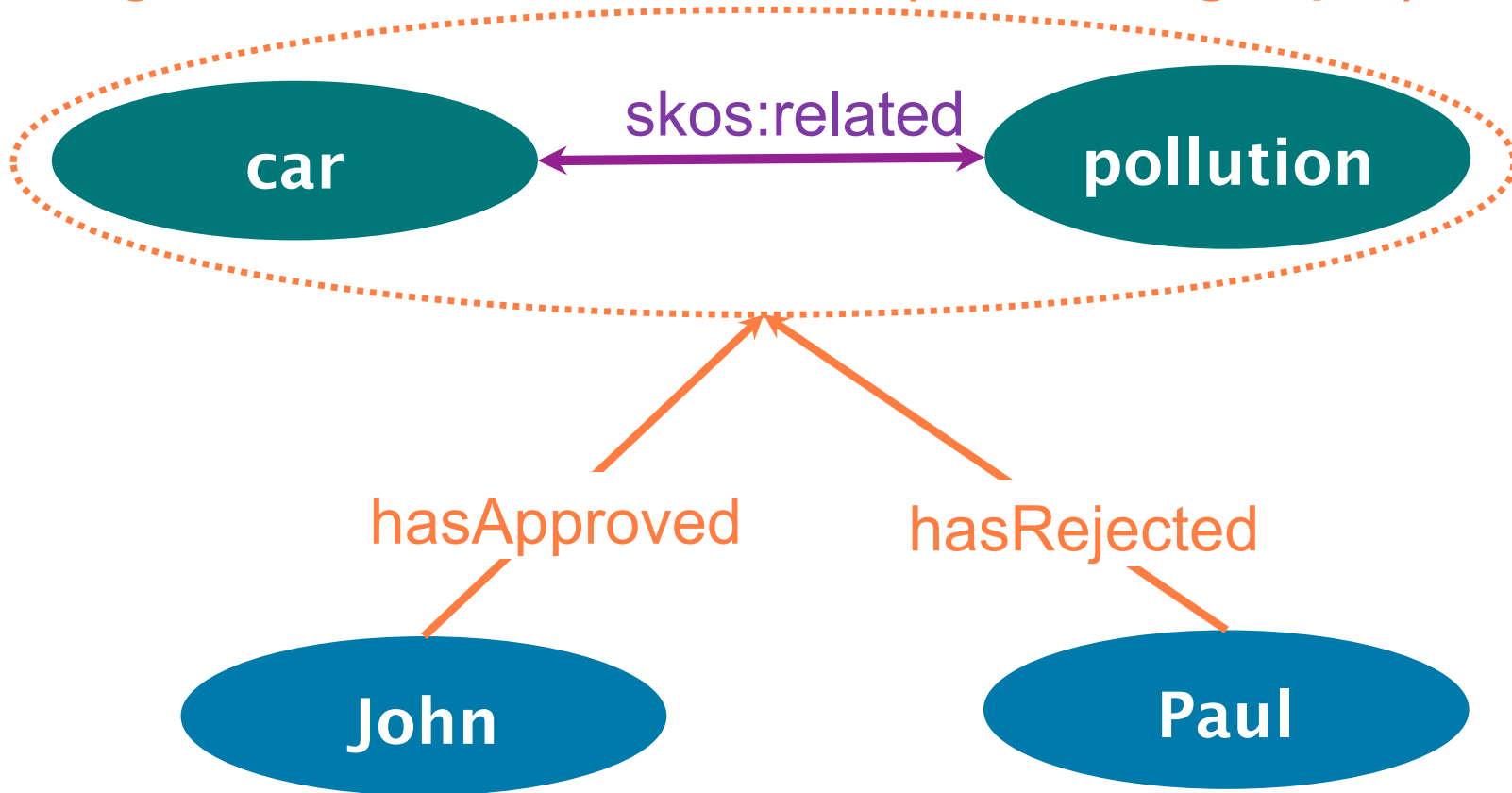


Supporting diverging points of view

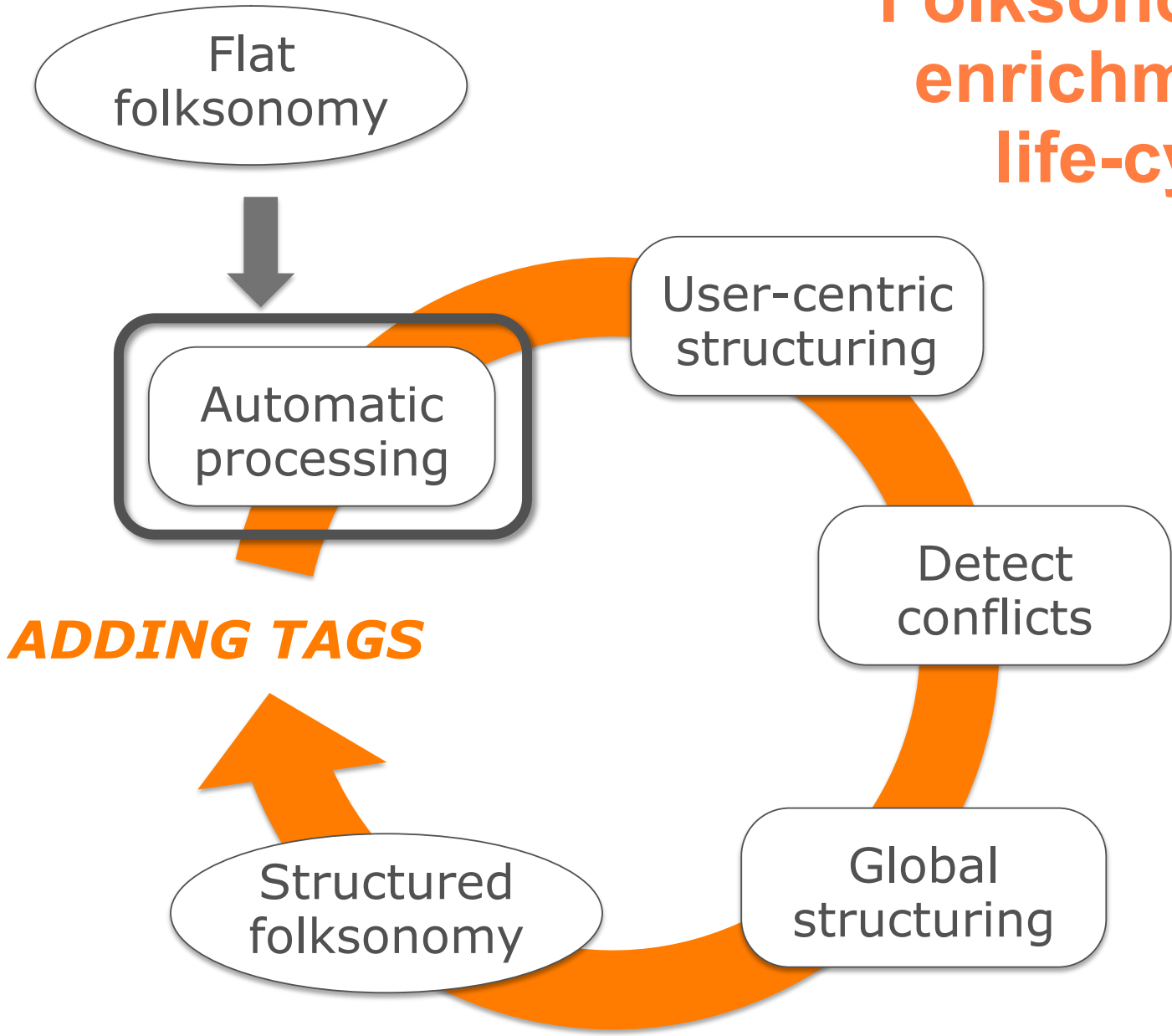


Supporting diverging points of view

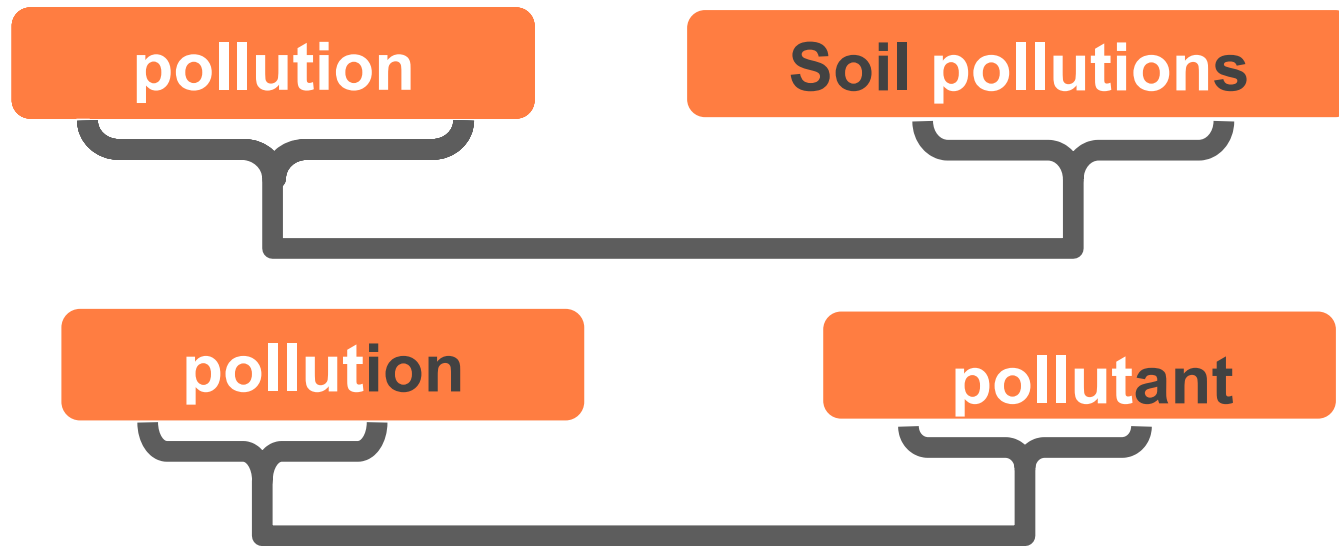
tagSemanticStatement (named graph)



Folksonomy enrichment life-cycle



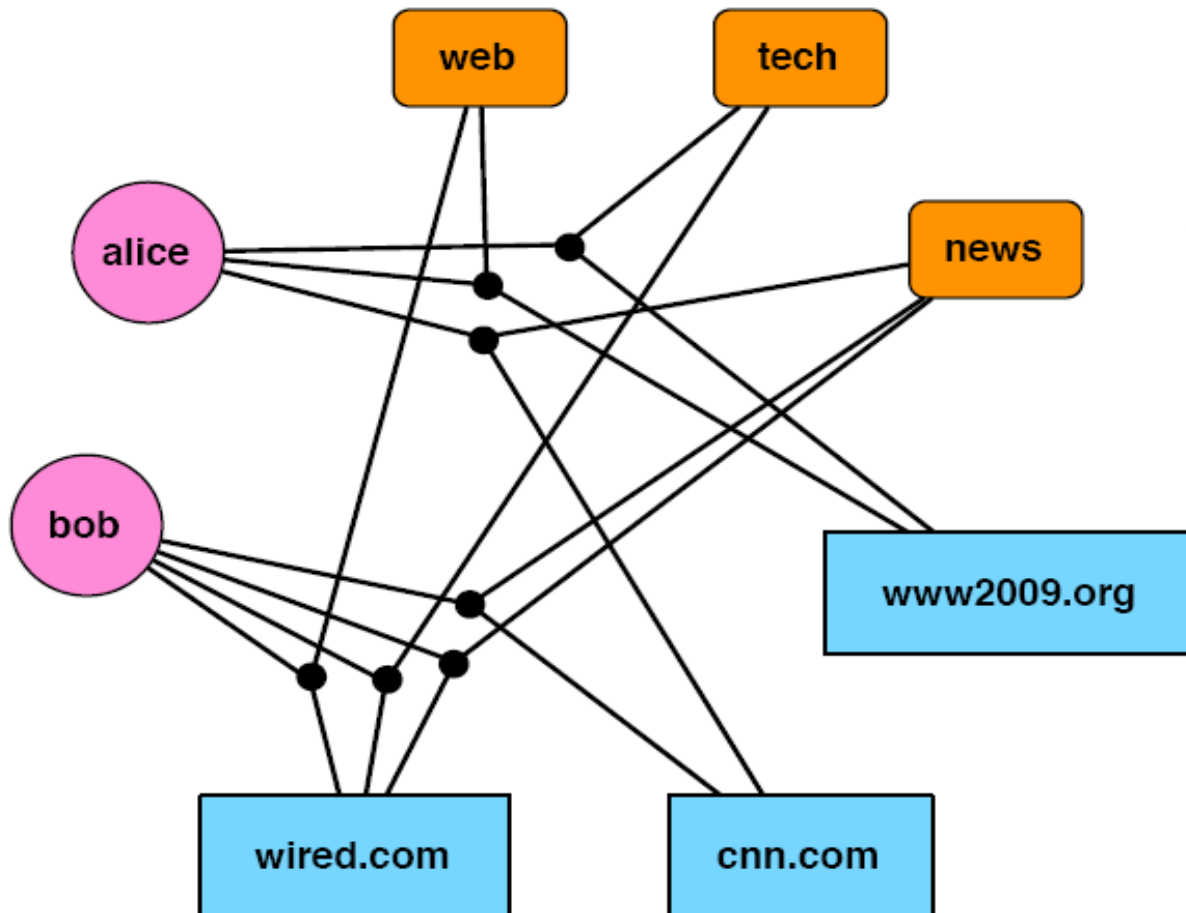
1. Comparing Tag labels with string edit distances



1. Comparing Tag labels with string edit distances

- Evaluation of 30 edit distances
- Combining the best metrics
- Needs complement !

2. Analyzing the tri-partite structure of folksonomies

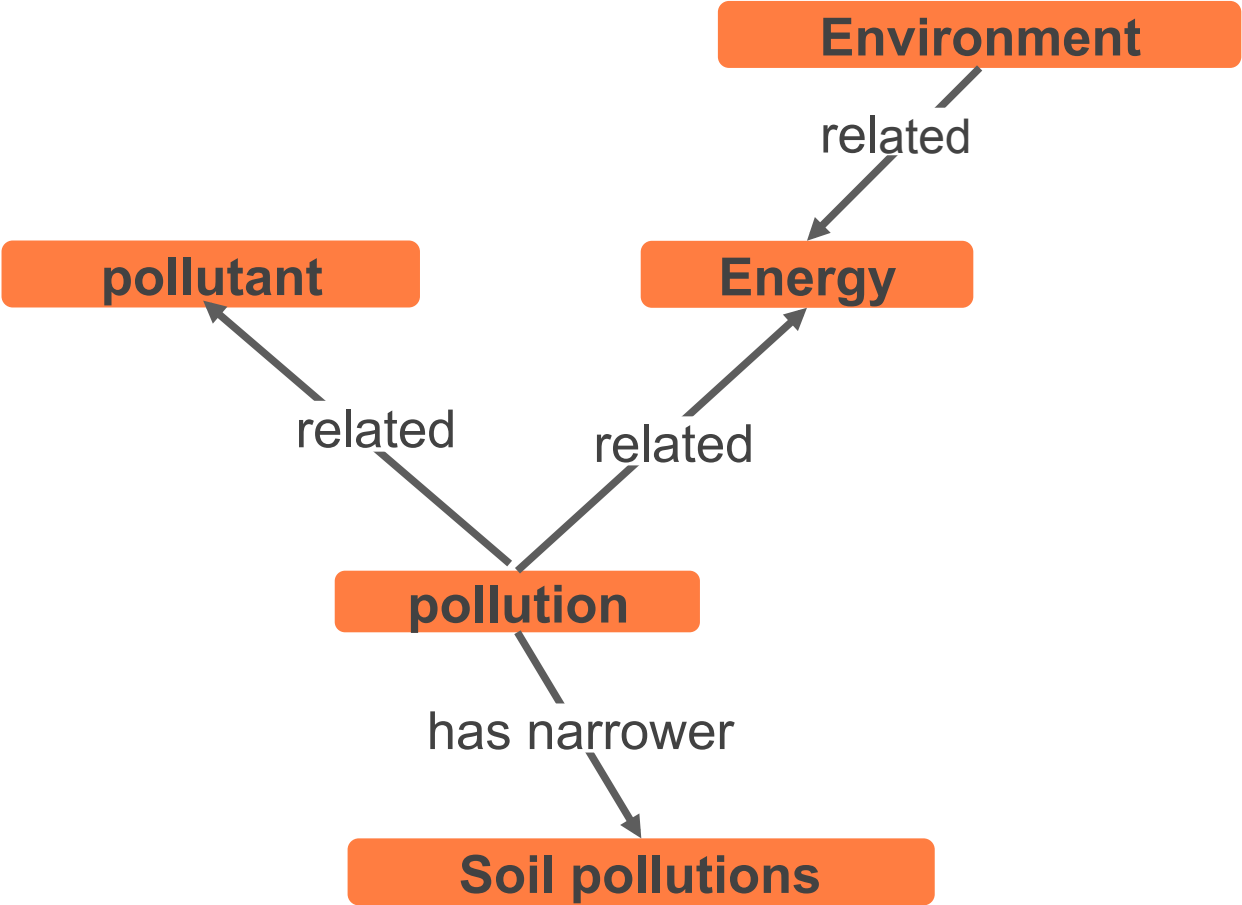


Association via:

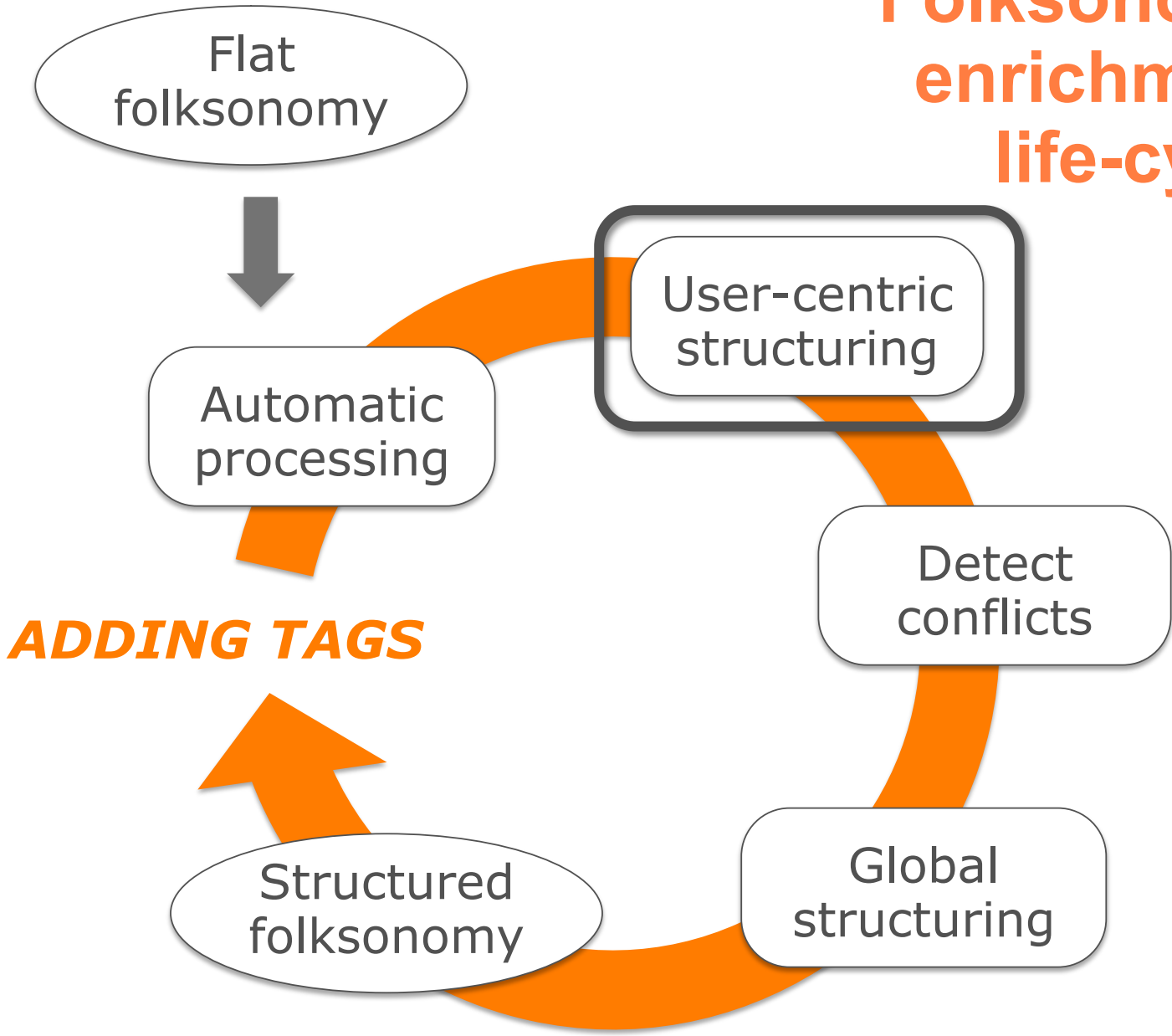
- Users
- tags

Fig. Markines *et al.* (2009)

Result: automatically suggested semantics



Folksonomy enrichment life-cycle



Embedding structuring tasks within everyday activity (searching e.g)

The screenshot shows a web browser window with the URL `chrome://srtag/content/taggingsBrowser/taggings-browser.xul`. The page title is "Semantic Related Tag Editor" and the user is logged in as "anonym2".

The interface features a search bar with the text "pollution". Below the search bar, there are four categories of related tags:

- BROADER:** A list of tags including "pollution transfrontaliere", "pollution visuelle", "pollution organique", "pollution harmonique", "pollution photooxydante", and "pollution episodique".
- RELATED:** A list of tags including "environnement- pollution", "pollution atmospherique impact sante", "3 traceur de pollution", "halo de pollution", and "pollution developpement instrumental".
- NARROWER:** A list of tags including "pollution transfrontaliere", "pollution visuelle", "pollution organique", "pollution harmonique", "pollution photooxydante", and "pollution episodique".
- CLOSE MATCH:** A list of tags including "pollution".

On the right side of the browser window, there is a panel titled "List of resources tagged with 'pollution'". It includes a note: "NB: Includes results for spelling variant tags (pollution)". Below this, there is a list of resources:

- [Comptes \(Les\) économiques de l'environnement : Données économiques de l'environnement, rapport de la commission des comptes et de l'économie de l'environnement. Editions 2001, 2002, 2003, 2005](#)
Tagged by **svic_referent** with : administration publique cout pollution politique de l'environnement ...
- [Stratégie nationale de développement durable. Enraciner l'avenir dans l'action.](#)
Tagged by **svic_referent** with : territoire peche climat sante ...
- [air des beffrois. ATMO Nord/Pas-de-Calais prend son envol \(L'\)](#)
Tagged by **svic_referent** with : polluant reseau mesure surveillance no2 ...
- [pollutions olfactives des installations classées : de l'évaluation de la gêne aux techniques de réduction. salon de l'aveyron le 8 et 9 février 2005](#)
Tagged by **svic_referent** with : emission odeur air agriculture ...
- [eurodeur biorodeur 2005, les solutions de dépollution extérieures et intérieures les impacts sanitaires de la pollution atmosphérique. Actes des conférences 22 et 23 juin 2005](#)
Tagged by **svic_referent** with : qualite air epidemiologie traitement odeur ...
- [gestion de l'environnement pour les PME-PMI \(La\)](#)
Tagged by **svic_referent** with : dechet industriel management environnemental legislation pollution

The browser window also shows a "Done" status at the bottom left.

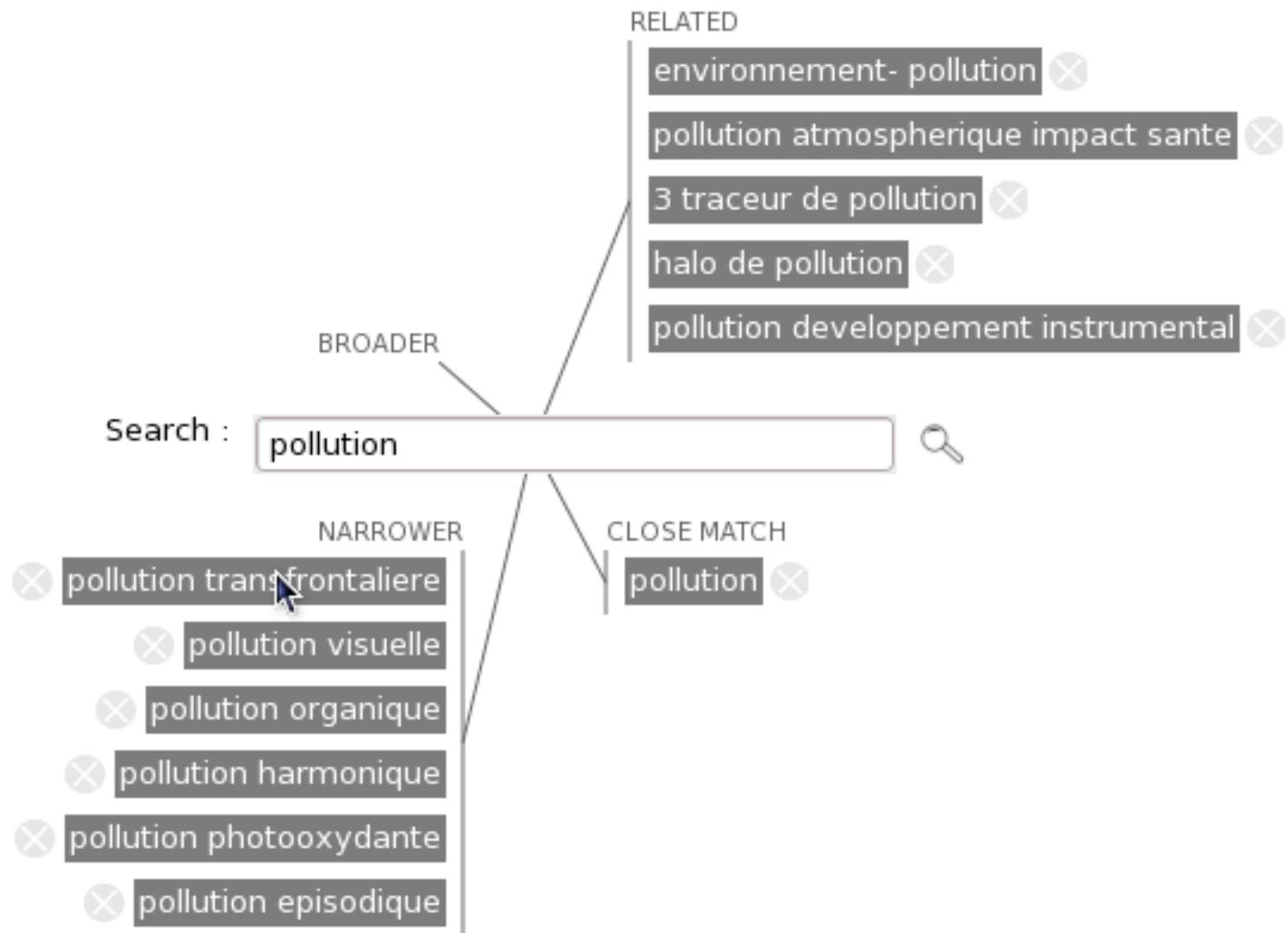
Embedding structuring tasks within everyday activity (searching e.g)

List of resources tagged with "pollution"

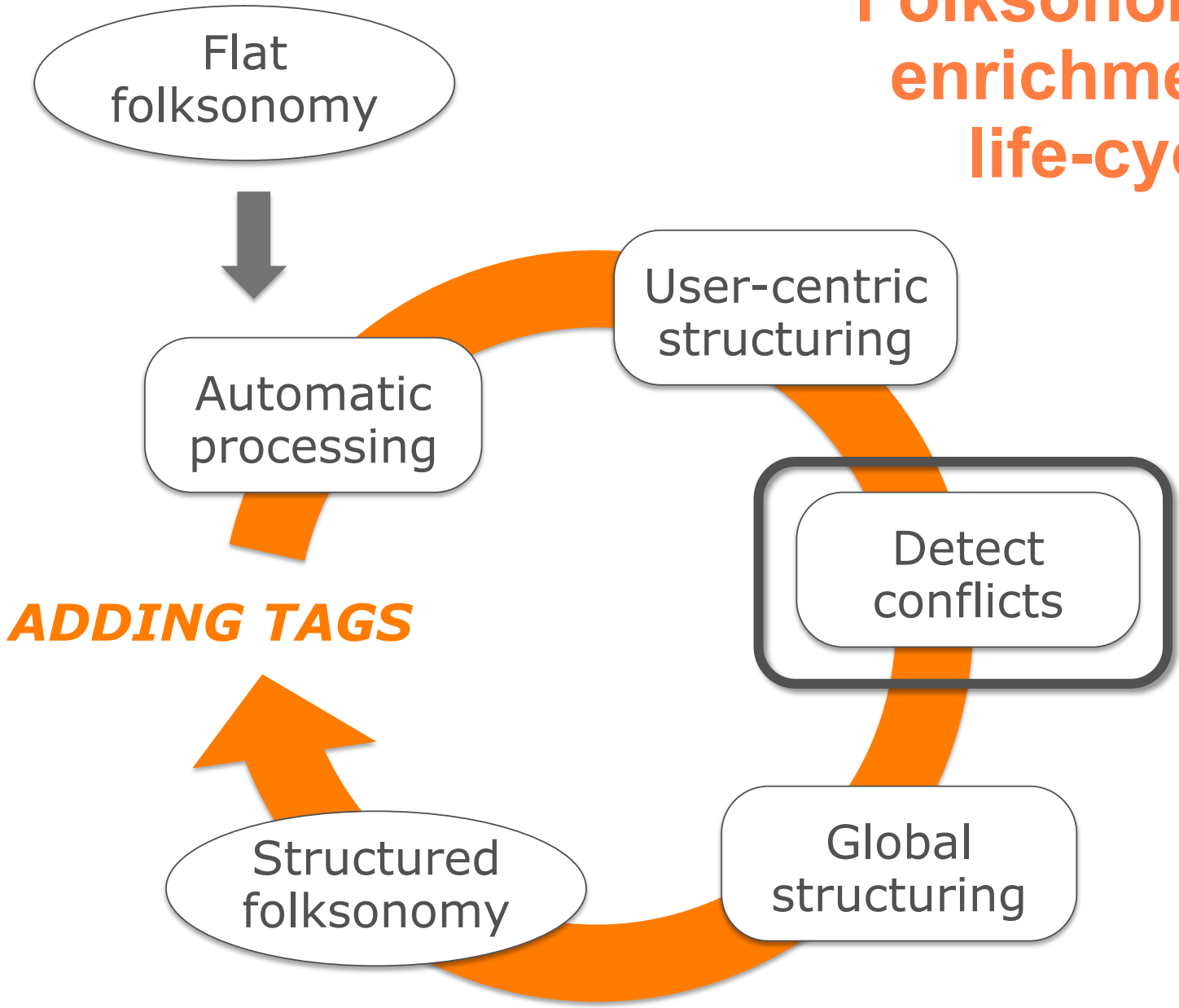
NB: Includes results for spelling variant tags (pollution)

- [Comptes \(Les\) économiques de l'environnement : Données économiques de l'environnement, rapport de la commission des comptes et de l'économie de l'environnement. Editions 2001, 2002, 2003, 2005](#)
Tagged by **svic_referent** with : administration publique cout pollution politique de l'environnement ...
- [Stratégie nationale de développement durable. Enraciner l'avenir dans l'action.](#)
Tagged by **svic_referent** with : territoire peche climat sante ...
- [air des beffrois. ATMO Nord/Pas-de-Calais prend son envol \(L'\)](#)
Tagged by **svic_referent** with : polluant reseau mesure surveillance no2 ...
- [pollutions olfactives des installations classées : de l'évaluation de la gêne aux techniques de réduction. salon de l'aveyron le 8 et 9 février 2005](#)
Tagged by **svic_referent** with : emission odeur air agriculture ...
- [eurodeur biorodeur 2005, les solutions de dépollution extérieures et intérieures les impacts sanitaires de la pollution atmosphérique. Actes des conférences 22 et 23 juin 2005](#)
Tagged by **svic_referent** with : qualite air epidemiologie traitement odeur ...

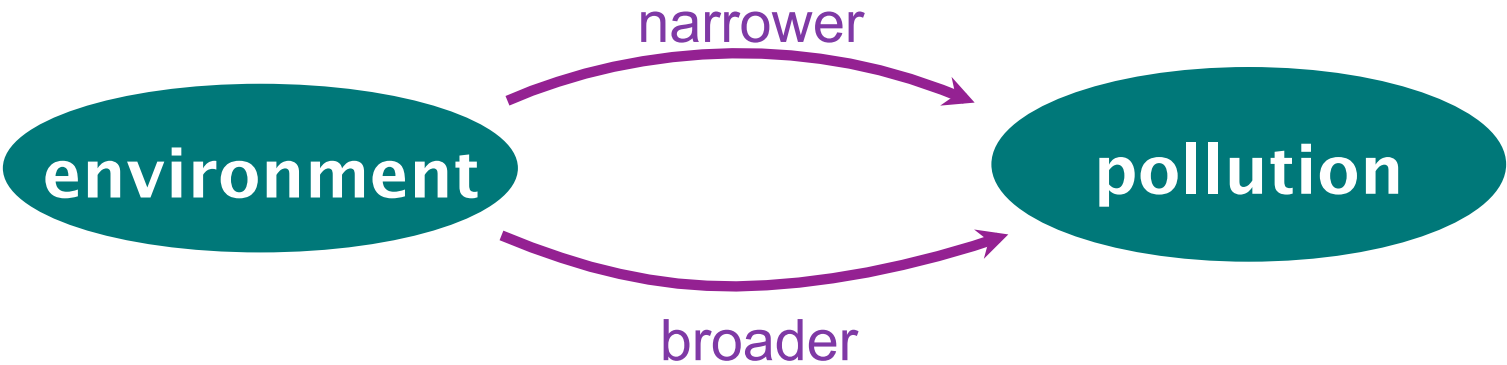
Capturing user's point of view



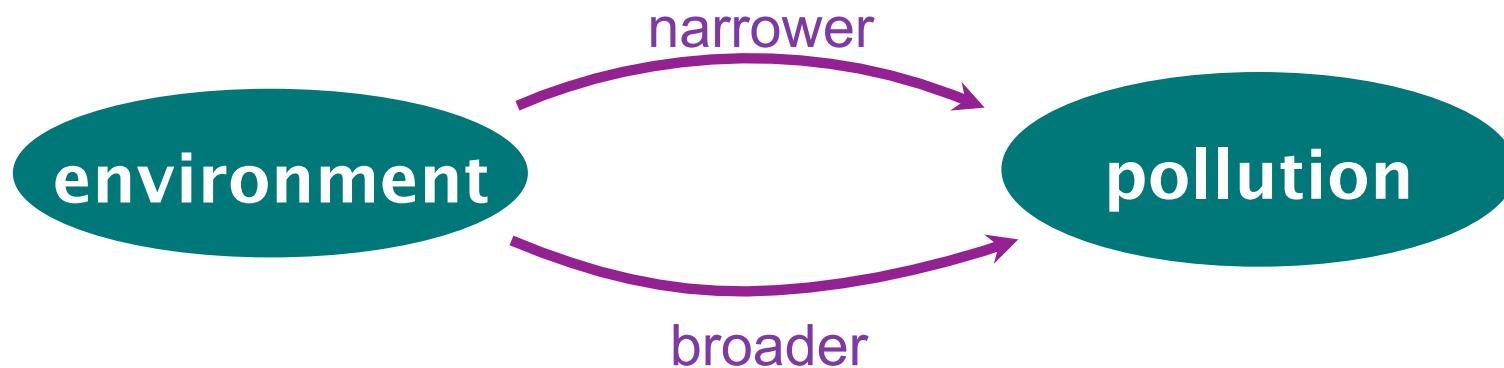
Folksonomy enrichment life-cycle



Conflict detection



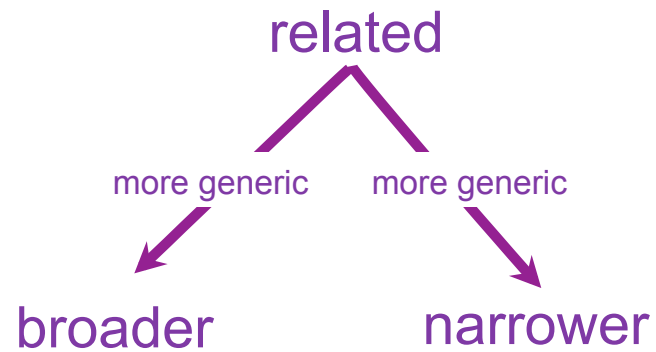
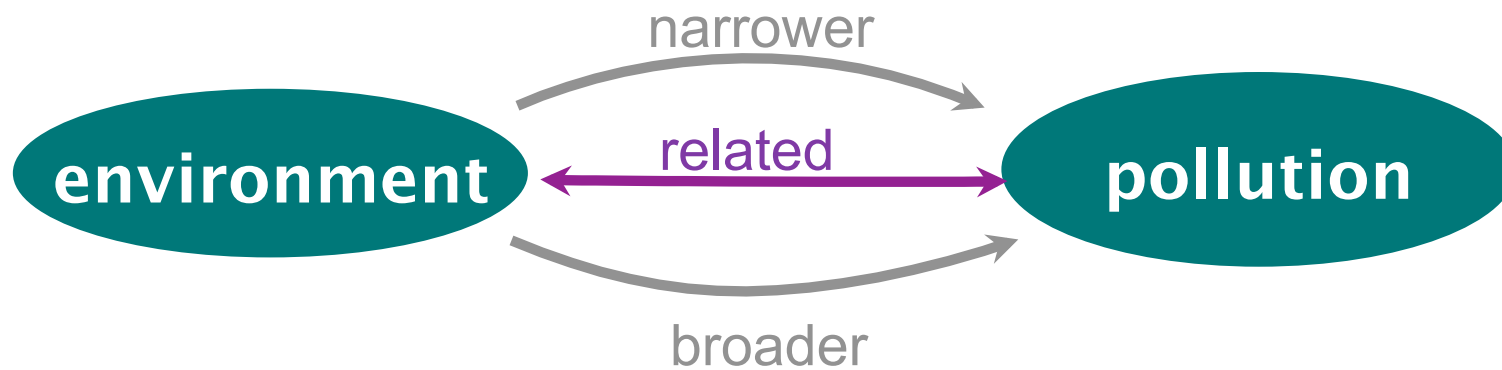
Conflict detection



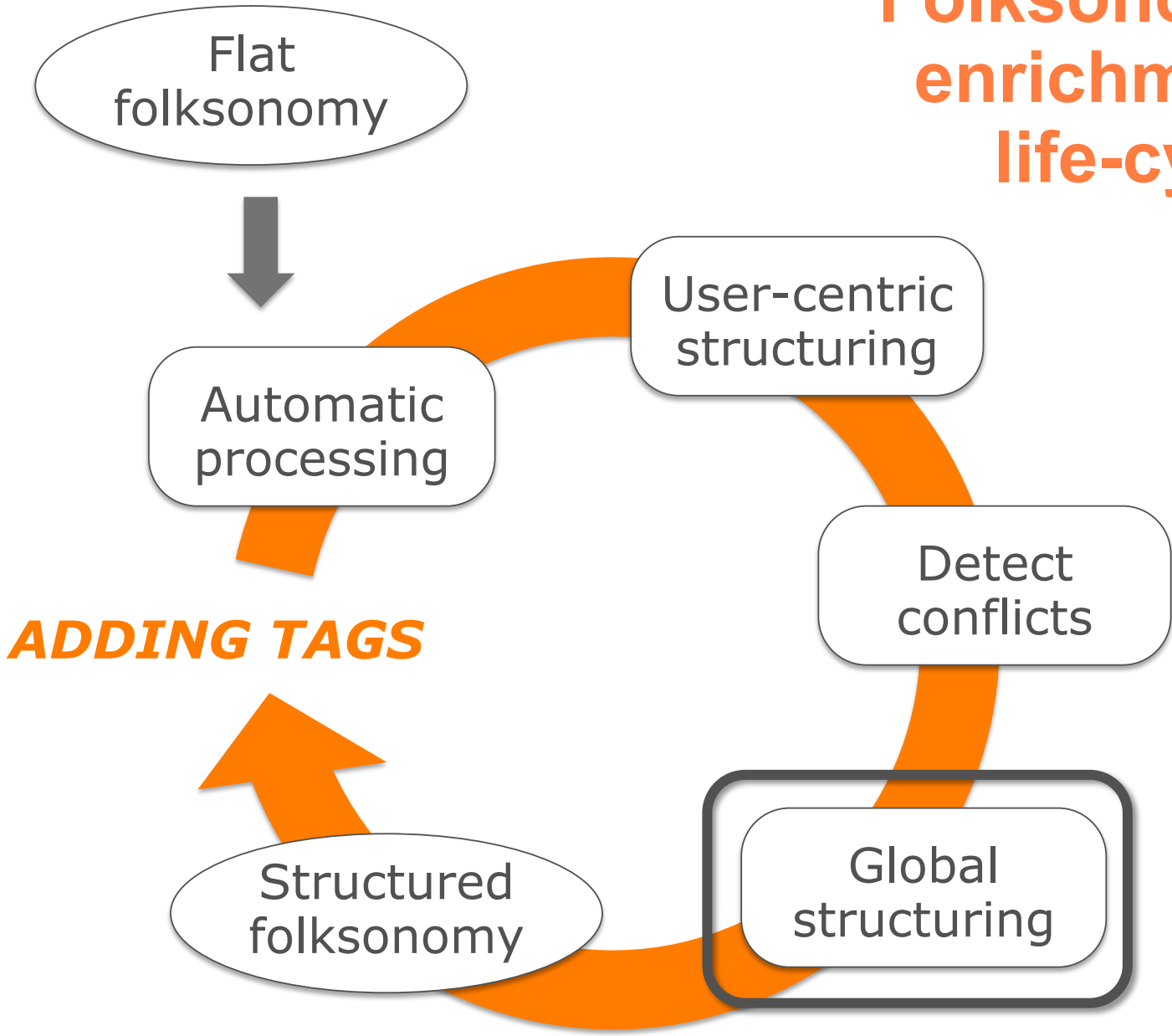
Using rules e.g:

```
IF num(narrower) / num(broader) ≥ c  
THEN narrower wins  
ELSE 'more generic' wins
```

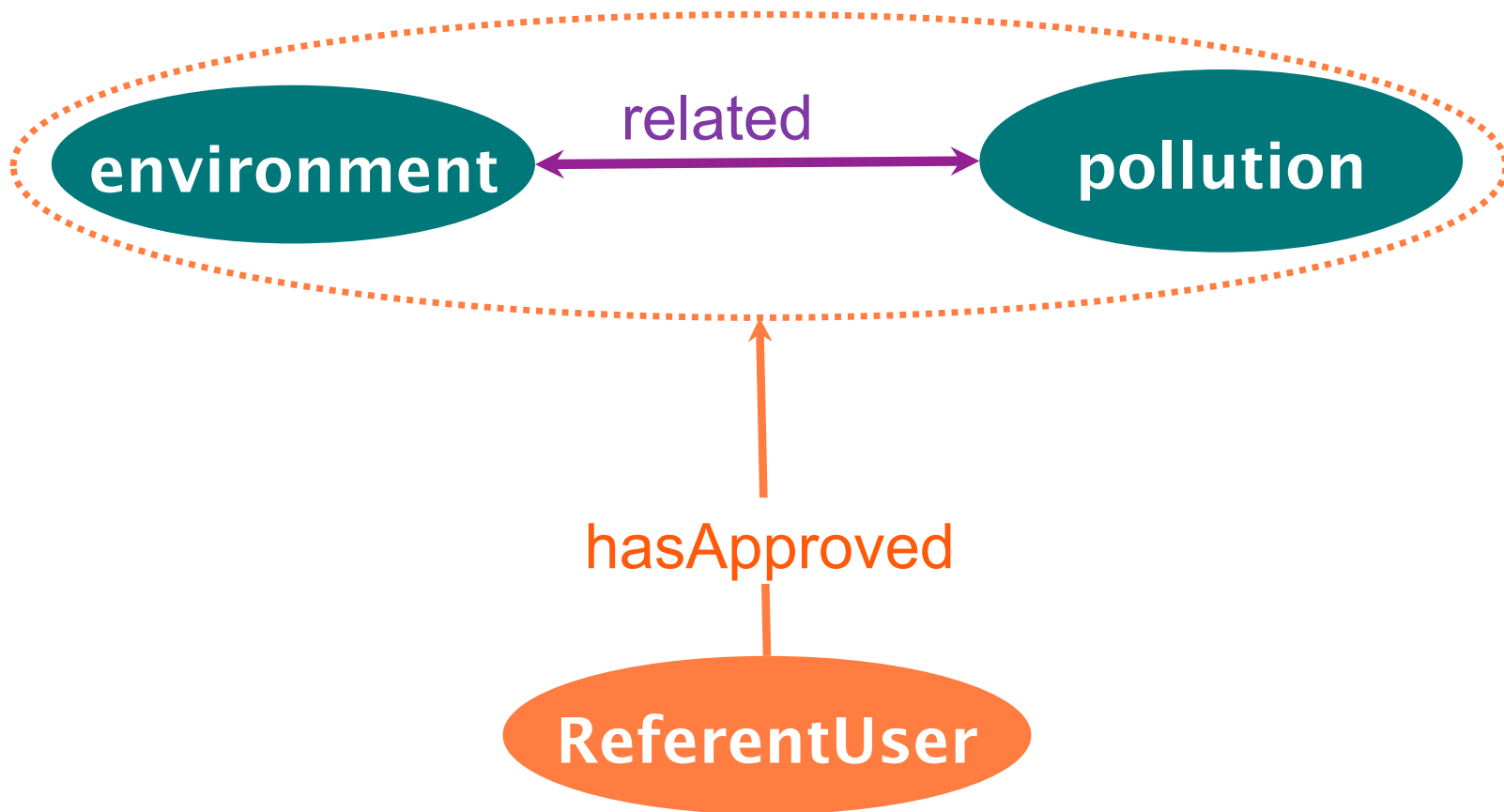

Conflict detection



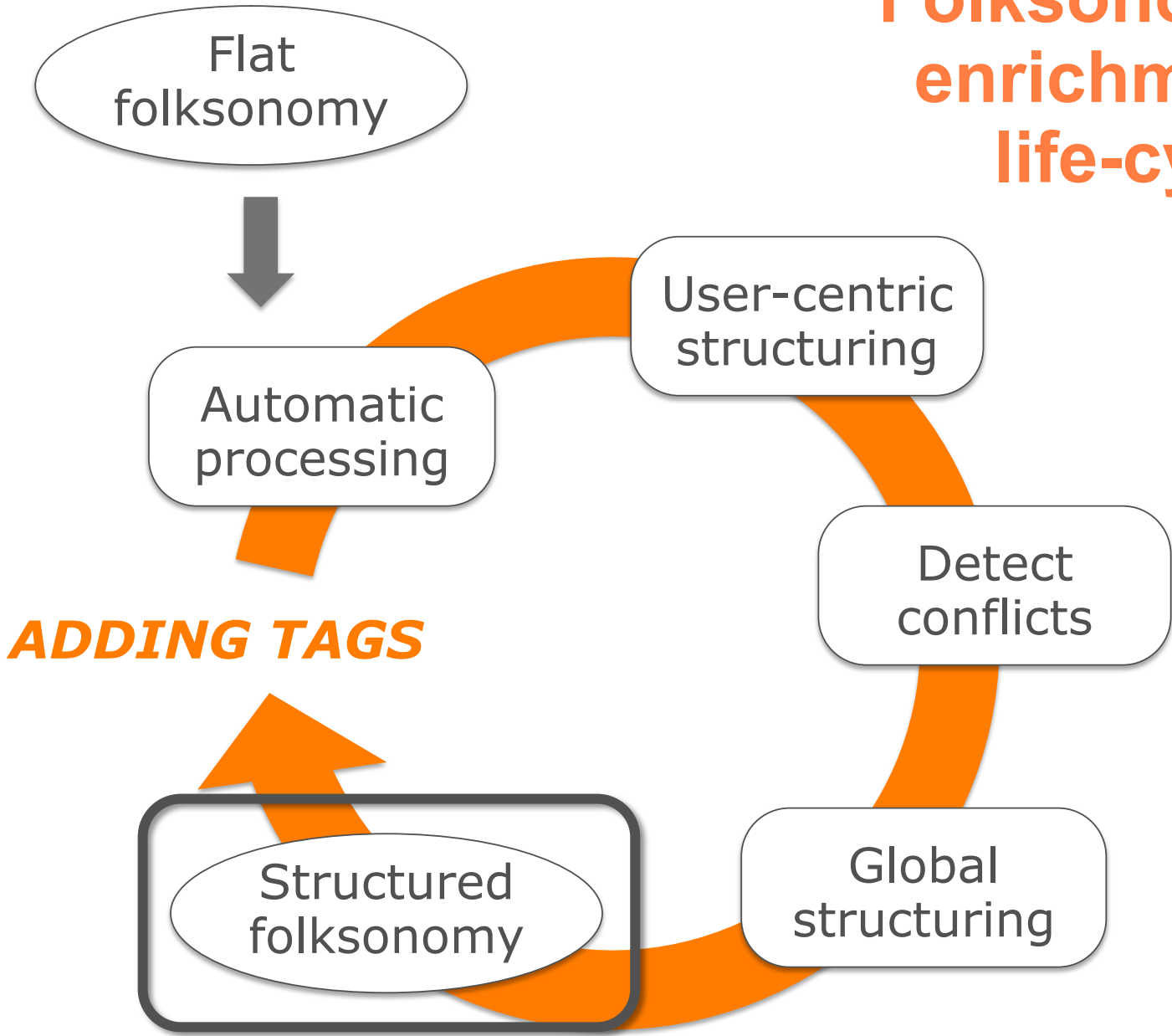
Folksonomy enrichment life-cycle



Global structuring by Referent



Folksonomy enrichment life-cycle



**Take away message
(conclusion)**

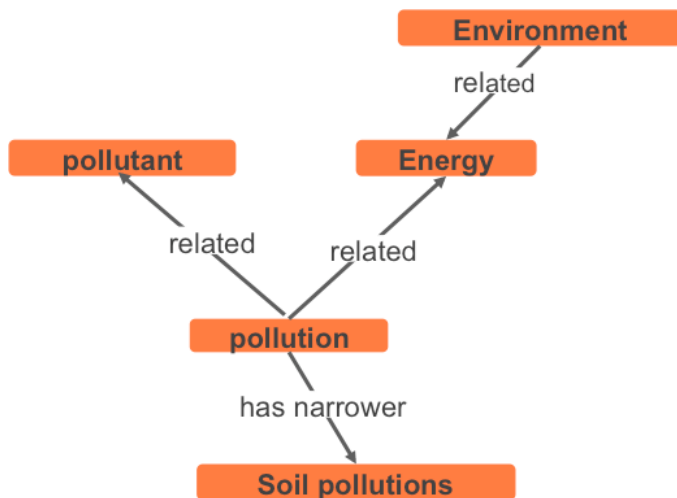


What we do :

Help communities



structure their tags



Our contributions:

- **Usages** analysis
- **Automatic** processing of tags
- Tag structuring embedded in **every-day tools**
- Supporting **multi-points of view**

Implementation & tests

- ADEME dataset (~10000 tags)
- Tag server
- Tag searching interface

Perspectives

- More automatic methods
- More ontological resources
- Other interfaces (tagging, global structuring)
- Test + Evaluation @ Ademe & Orange Labs

Thank you !

`freddy.limpens@inria.fr`

`http://www-sop.inria.fr/members/Freddy.Limpens/`

`http://isicil.inria.fr`