

Collaborative semantic structuring of folksonomies

Freddy LIMPENS, Fabien GANDON
Edelweiss, INRIA Sophia-Antipolis
firstname.name@sophia.inria.fr

Michel BUFFA
KEWI, I3S, CNRS-UNSA
buffa@unice.fr

Web Intelligence 2009 – Milano, Italia

Objectives

- *Organizing resources of organizations with structured voc & semantic tools*
 - *Combiningthe ease of use of tagging*

AND

...the power of ontologies

1. Bridging social tagging & ontologies

"Social Tagging" = *freely categorizing*

NOTES Le graphe ci-dessus donne une impression de la gamme de prix, en février 1998, des éoliennes modernes raccordées au réseau. 876 char

TAGS é

- écologie (14)
- énergie (6)
- économie (5)
- éco-construction (2)
- écoconstruction (2)
- étudiants (2)
- échange (1)
- école-doctoral (1)
- école-été (1)

is. E.g., hotels bargains newyork (not new york)
are this link with your network
for:iurics for:margooot

Save

...and sharing

What's New? | Signed in as JeanSaul | [Inbox](#) | [Se](#)

Recent eolienne Bookmarks

Recent | Popular

See all eolienne bookmarks in [Popular](#) or your [Bookmarks](#) or [Network](#).

Tags [eolienne](#) Type another tag

07 OCT 08 Combien coûte une éolienne ? [SAVED](#)
Le graphe ci-dessus donne une impression de la gamme de prix, en février 1998, des éoliennes modernes raccordées au réseau.
[jeansaul](#) [éolienne](#) [énergie](#) [renouvelable](#) [cout](#)

28 SEP 08 Le JT en plein écran | Antenne Centre Télévision [SAVE](#)
[zaccone](#) [windvision](#) [tv](#) [zaccone](#) [belgium](#) [eolienne](#)

26 SEP 08 Weole Energy - Le petit et moyen éolien [SAVE](#)
[fourretout](#) [eolien](#) [eoliennes](#) [energie](#) [credit](#) [impots](#) [renouvelable](#) [eolienne](#) 3

25 SEP 08 How I built an electricity producing wind turbine [SAVE](#)
[rodolpheg](#) [energy](#) [generator](#) [wind](#) [electronics](#) [diy](#) [turbine](#) [eolienne](#) 8

24 SEP 08 TIEOLE énergies éoliennes | Formations Auto-construction [SAVE](#)
autoconstruction d'une éolienne
[alain.hervais](#) [éolienne](#) 2

Loopwing Co.,Ltd. - Top [SAVE](#)
[dogezaemon](#) [power](#) [wind](#) [japon](#) [eolienne](#) 12

2008/09/24 > BE Espagne 74 > L'envolée de l'éolien espagnol [SAVE](#)
[lookatsciences](#) [espagne](#) [castille](#) [éolienne](#) [énergie](#) 2

23 SEP 08 Eolienne Skystream 3.7 Injection réseau | ENR [SAVE](#)
10 KE
[bedicou](#) [maison](#) [énergie](#) [électricité](#) [éolienne](#)

00:00:00 <http://delicious.com/environment/eolien.html> ... 89

limits of folksonomies

Spelling variations of tags:

newyork = new_york

ademe = ADEME

musique = music

Ambiguity of tags:

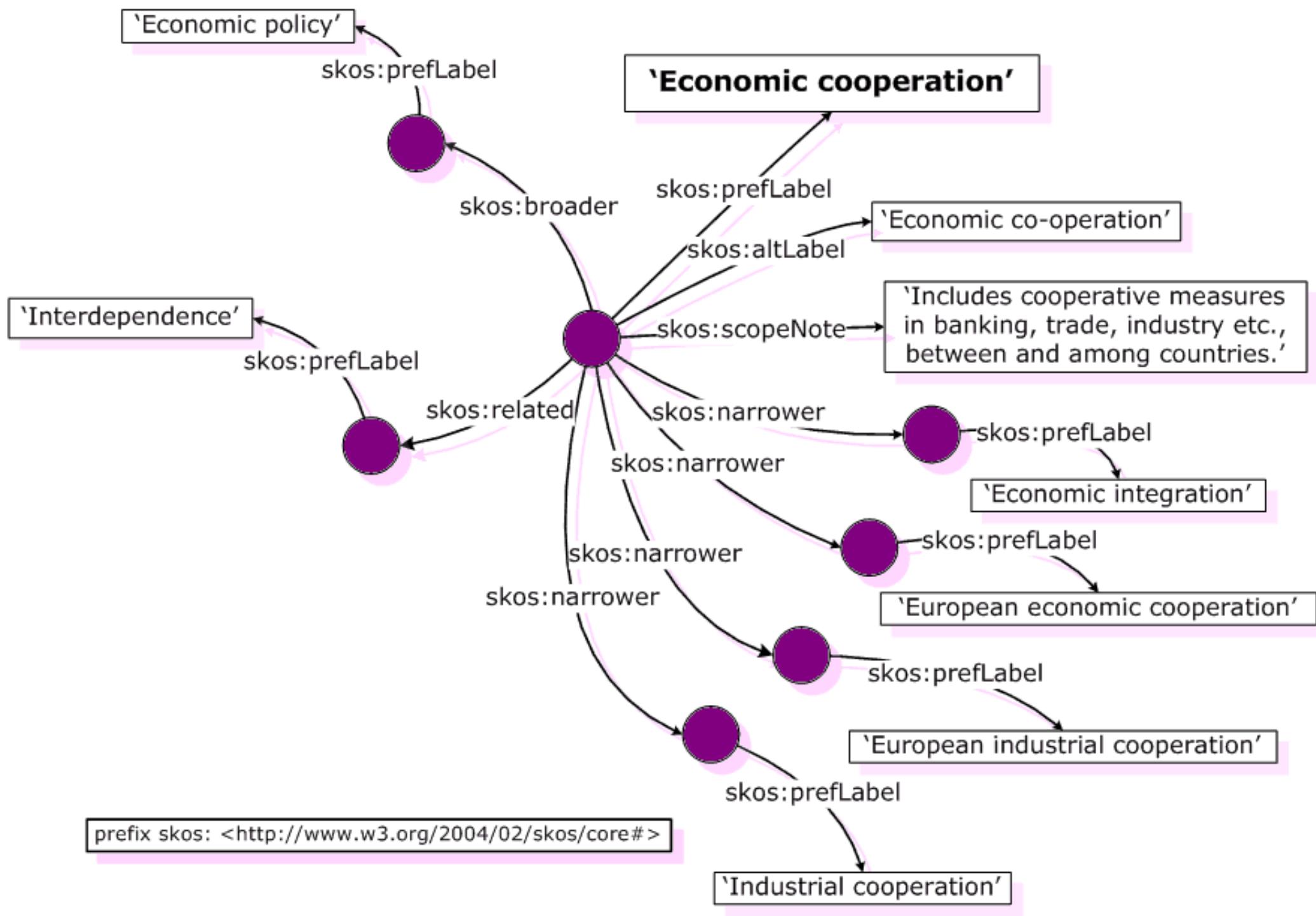
RDF =? Ressource Description Framework

RDF =? Rwanda Defense Force

Lack of semantic links between tags:

article artisan association atest avis
banque belgique bilancarbone bio biodiversite
bois bonusmalus bouteille bretagne buzz
calculateur canada cartouche catalogue CCD
CCS changement charity chimie chine citoyen
citroën climat co2 cochon collectif colloque
commerce commerceequitable communaute
communciation communication comparateur
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
europe evaluation evolution extensif faune
federation fertilisation filiere film finance flore
fondation forum **france** fruit GES
goodguide granddistribution graphique green
greenpeace greenwashing grenelle guide
habitat hitech HQE HTML hybride impact
import imprimante inde indice industrie
informaion **information** informatique
innovation institut interdiction **internet**

Ontologies = **structured views**
on the world



Limits of ontologies:

- **cost of elaboration**
- **cost of maintenance**
- integrating **community's expertise**

Our approach

- exploiting **automatic handling of tags** +
users expertise
- keeping tracks of **diverging point of views**

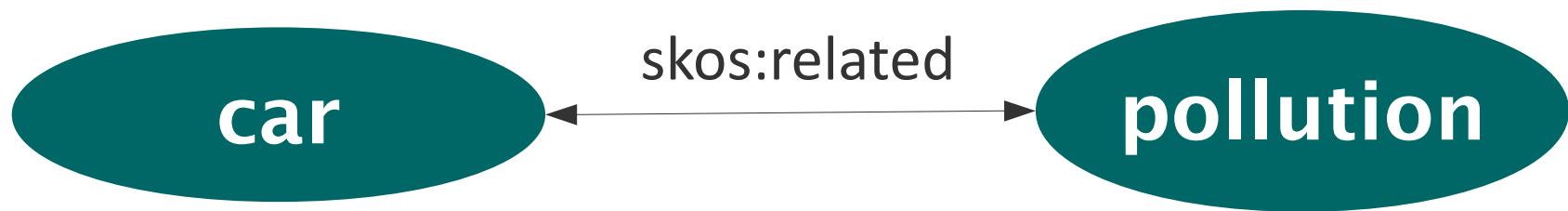
2. Model of folksonomy enrichment

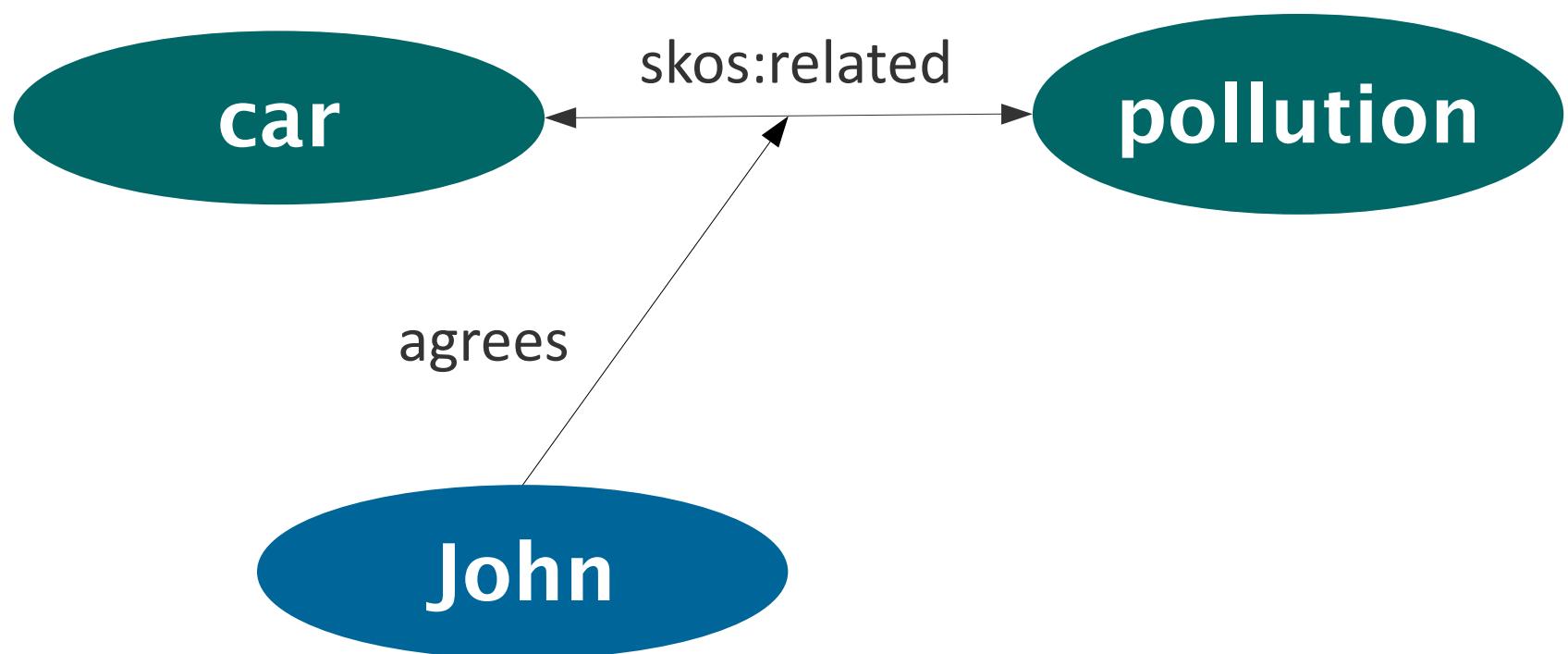
We propose a model to : ...

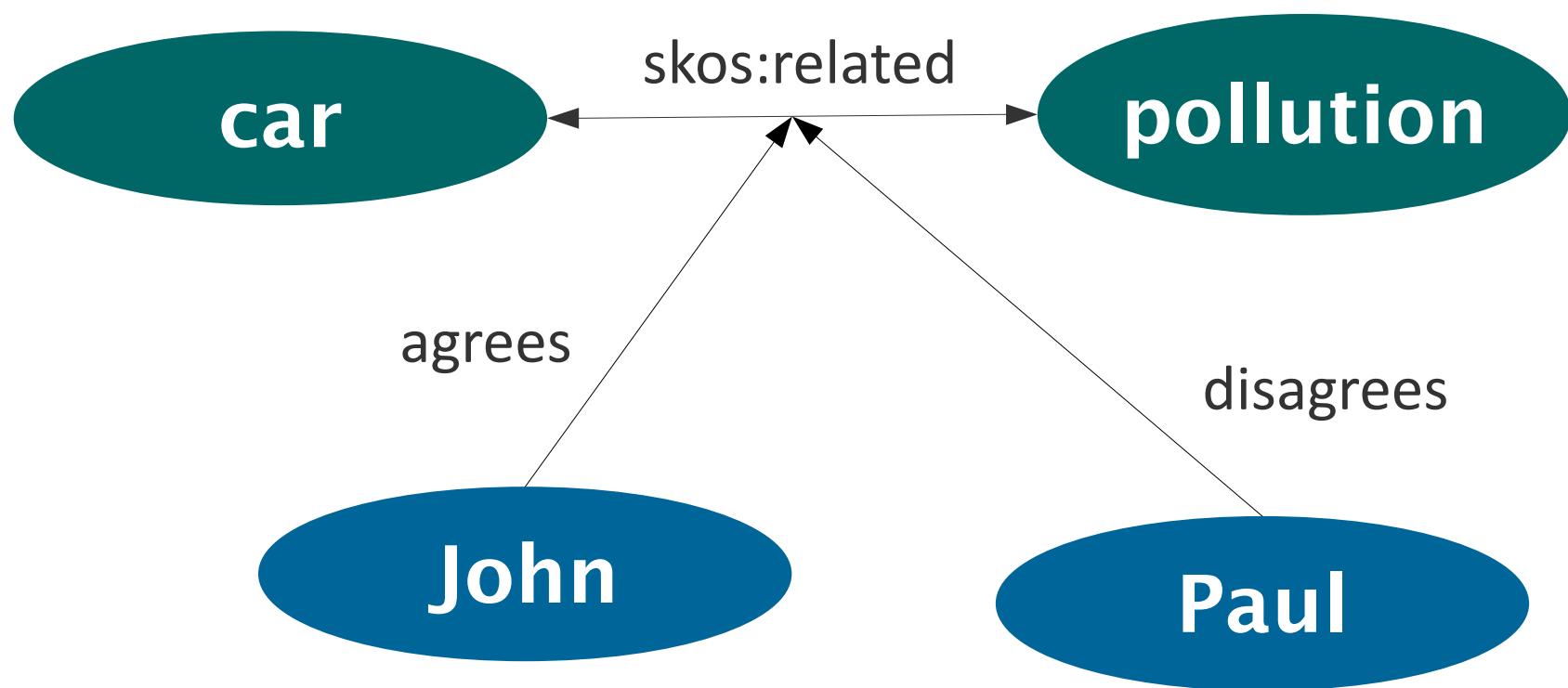
→ state something **about** the
relation between 2 tags

AND

→ support diverging **points of view**

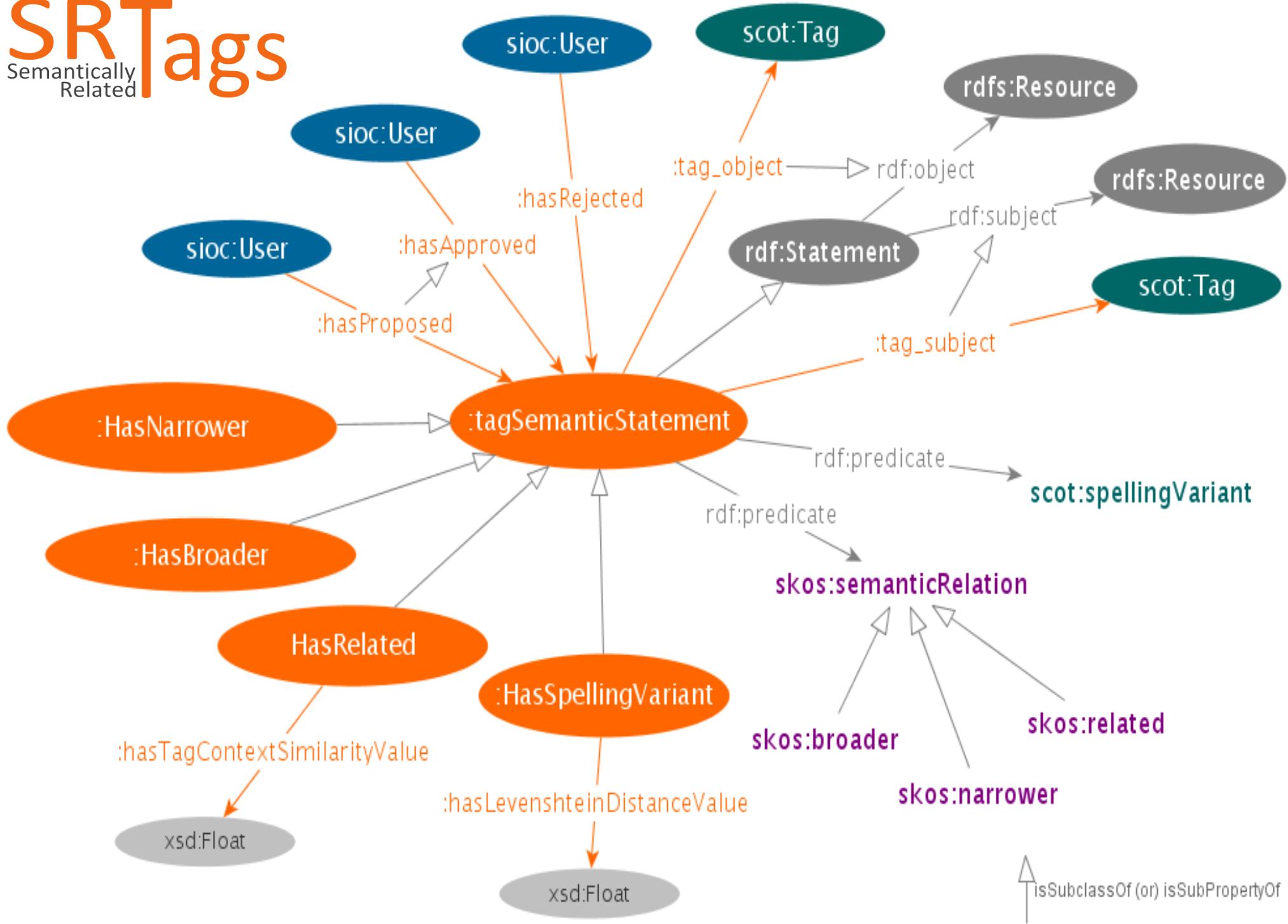






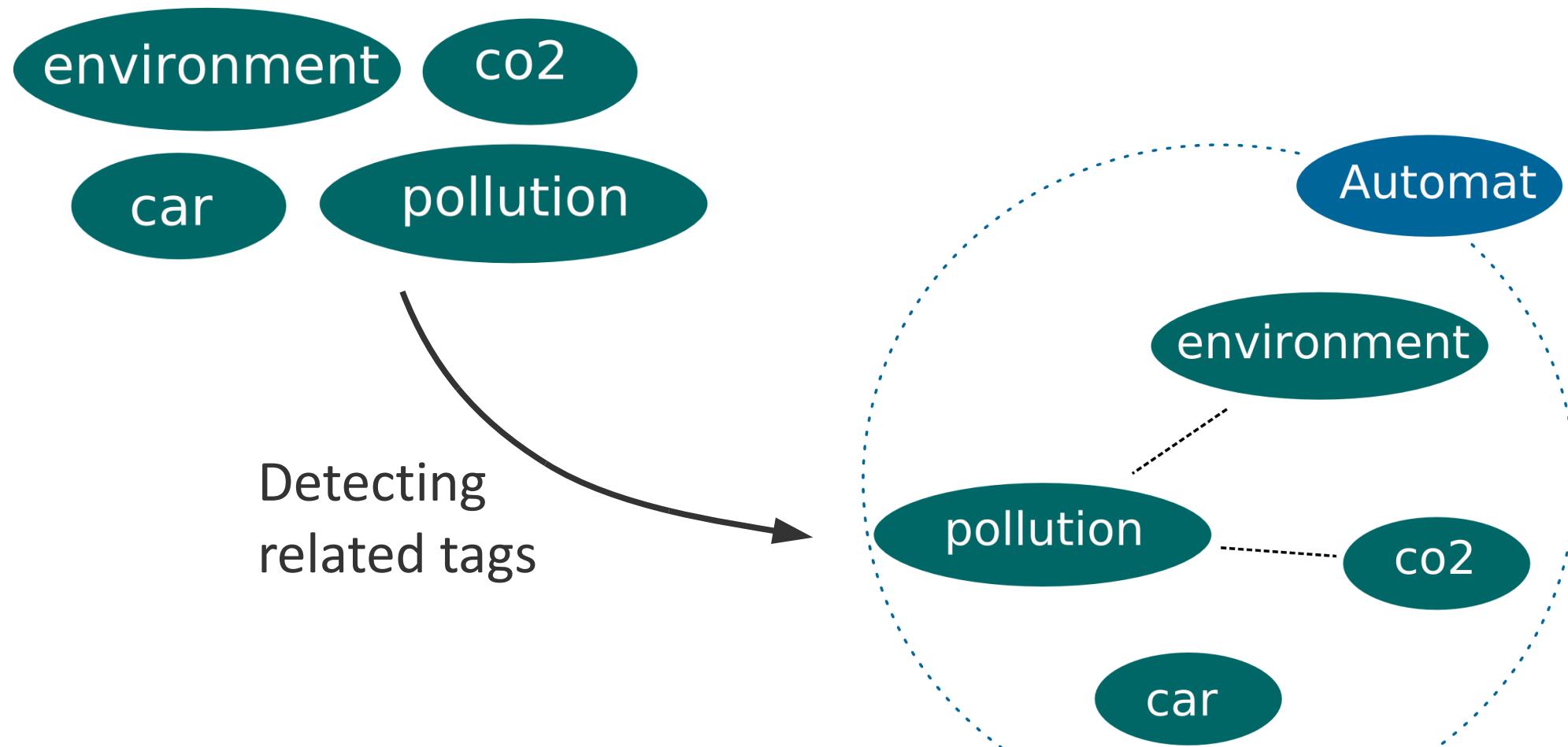
SRTags

Semantically Related



3. Collaborative and assisted semantic organization

Automatic handling of folksonomies



Detecting related tags

→ Computing tag similarity

(Mika, 2005; Cattuto et al., 2008)

→ Different similarity contexts

- resources
- users
- tags

Detecting related tags

tags	tag1	tag2	tag3
tag1	freq	cooc	cooc
tag2	cooc	freq	cooc
tag3	cooc	cooc	freq

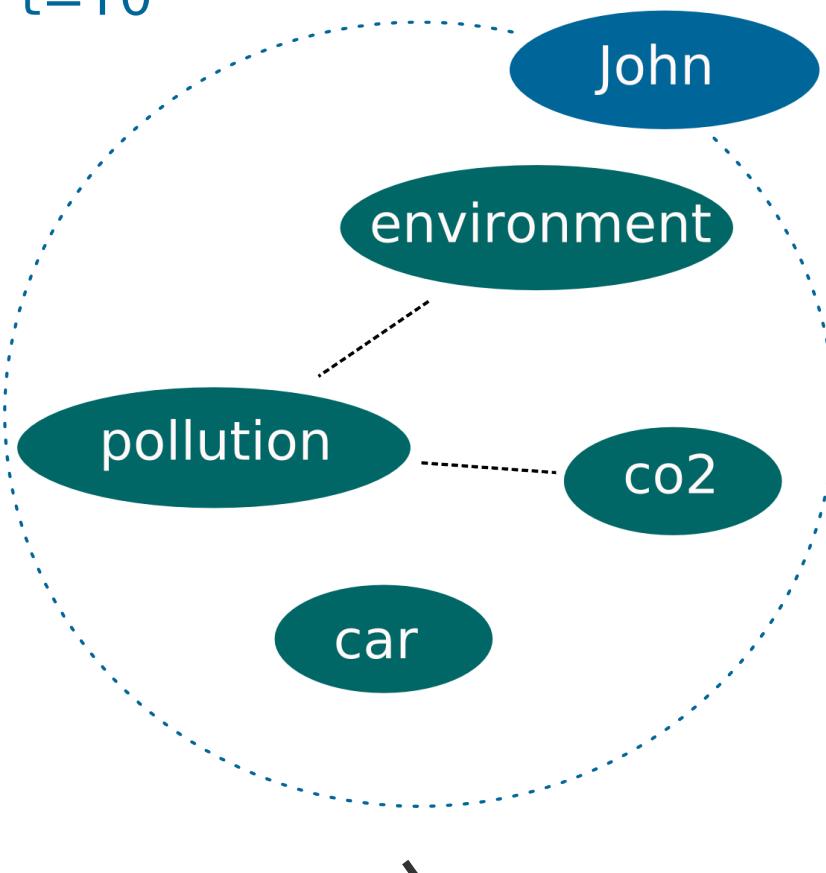
Detecting related tags

Tag	"Most related" tags (Tag-context)
environnement	développement_durable, énergie
environment	france, green, ecology, sustainable, energy
déchets	administration, EEDD, renouvelable
développementdurable	ecologie , EEDD, france, climat ,imported

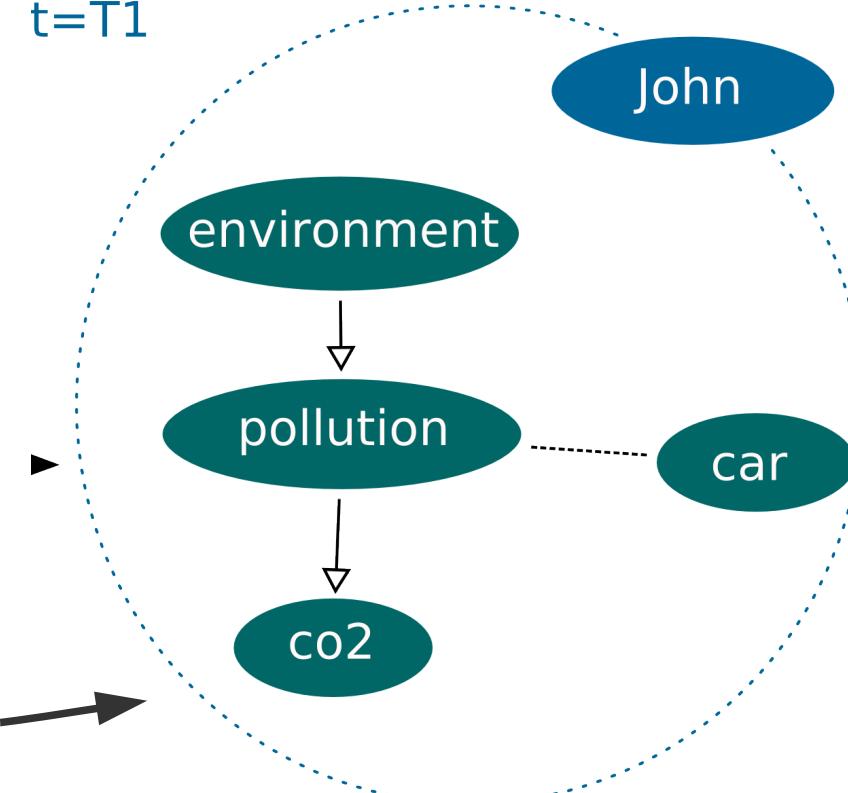
(delicious.com data)

Exploiting users' expertise

$t=T_0$



$t=T_1$



Micro-editing
(SRTags Editor)

Integrating semantic functionnalities within everyday tasks

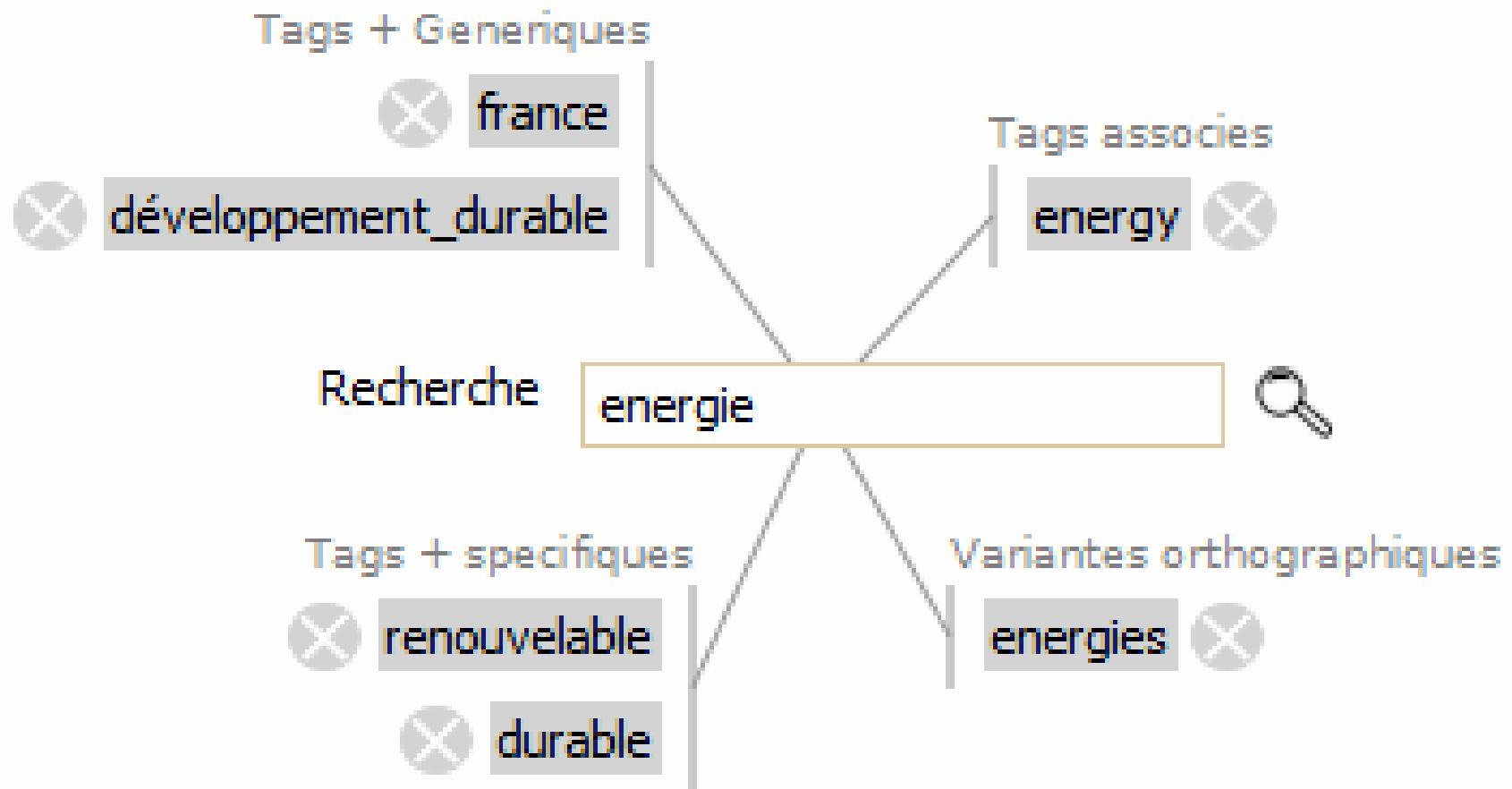
Screenshot of Mozilla Firefox showing the Semantic Related Tag Editor and Bookmarks Browser.

The left panel, "Semantic Related Tag Editor", displays a search interface for the tag "energie". It includes sections for "Tags + Generiques" (france, développement_durable), "Tags associés" (energy), "Tags + spécifiques" (renouvelable, durable), and "Variantes orthographiques" (energies). A search bar contains the term "energie".

The right panel, "Bookmarks Browser", shows a list of resources associated with the tag "energie". The title is "Liste des ressources associées au tag \"energie\"". A note states: "NB: sont inclus les ressources associées aux variantes orthographiques (energies)".

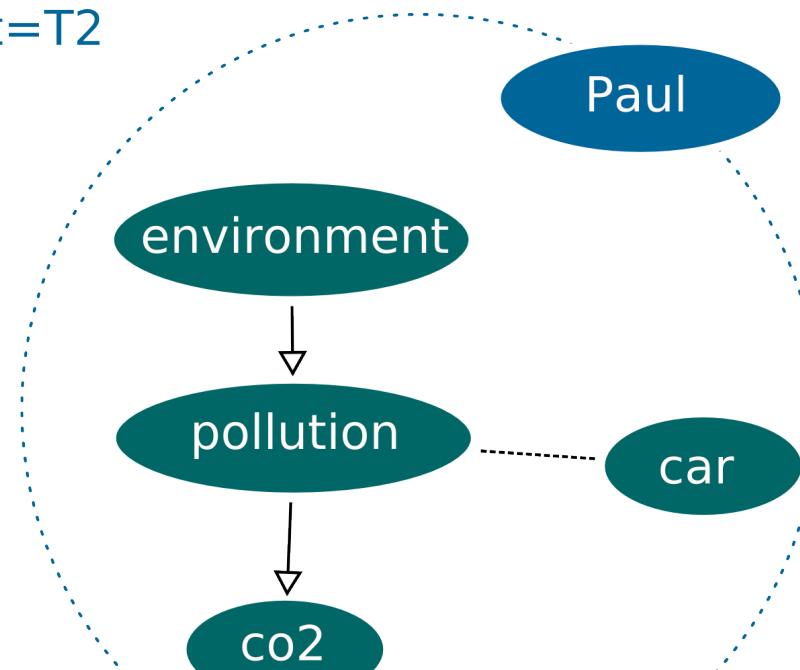
- [Bilan Carbone™ Personnel](#)
Tagged by concorde1415 with: propre energie durable developpement ...
- [ADEME](#)
Tagged by 7clem with: energie ...
- [ADEME - Agence de l'Environnement et de la Maîtrise de l'Energie](#)
Tagged by frederi.mandin with: environnement energy energie ...
- [ADEME - Agence De l'Environnement et de la Maîtrise de l'Energie](#)
Tagged by janjac with: technologie solaire eolien environnement ...
- [ADEME](#)
Tagged by lrichard with: Société Pratique environnement energie ...
- [ADEME](#)
Tagged by pur92 with: legislation immobilier france environnement ...
- [Indicateurs transports énergie environnement](#)
Tagged by meltedbloc with: transports environnement energie ...
- [Eco-prêts, le comparateur des prêts dédiés aux économies d'énergie](#)
Tagged by lemano with: pret immobilier energie ...
- [Bienvenue sur le site Agrice](#)
Tagged by bibobongo with: sciences science environnement energy ...
- [Guide pratique de l'ADEME, Les éoliennes.](#)
Tagged by Auclert with: renouvelable eoliennes energie ademe ...
- [Guide pratique de l'ADEME, « Pour maîtriser vos dépenses d'énergie, les aides financières \(habitat\) 2006 ».](#)
Tagged by hodac with: energie conseils ademe ...

Integrating semantic functionnalities within everyday tasks

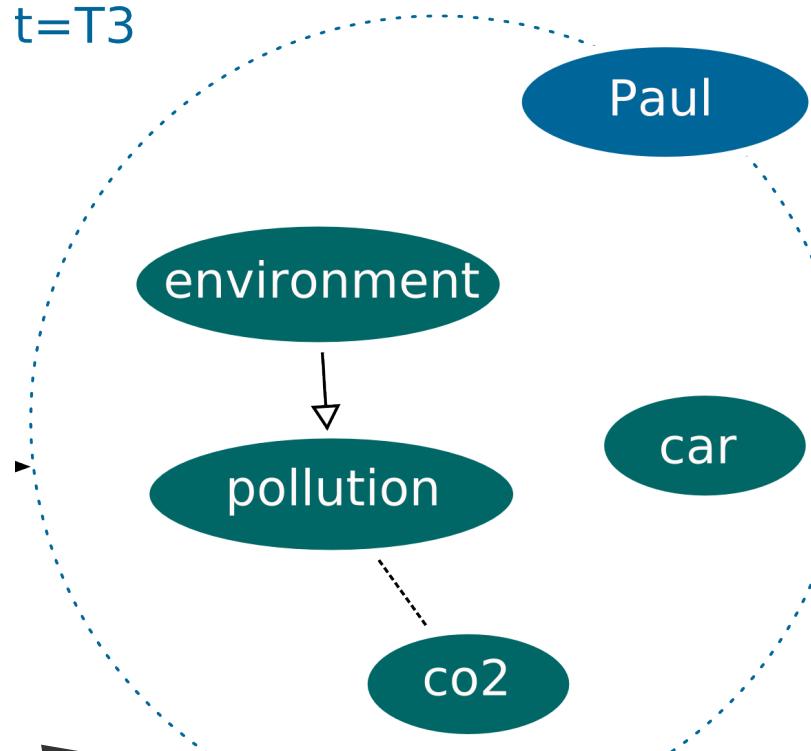


Supporting diverging points of view

$t=T2$

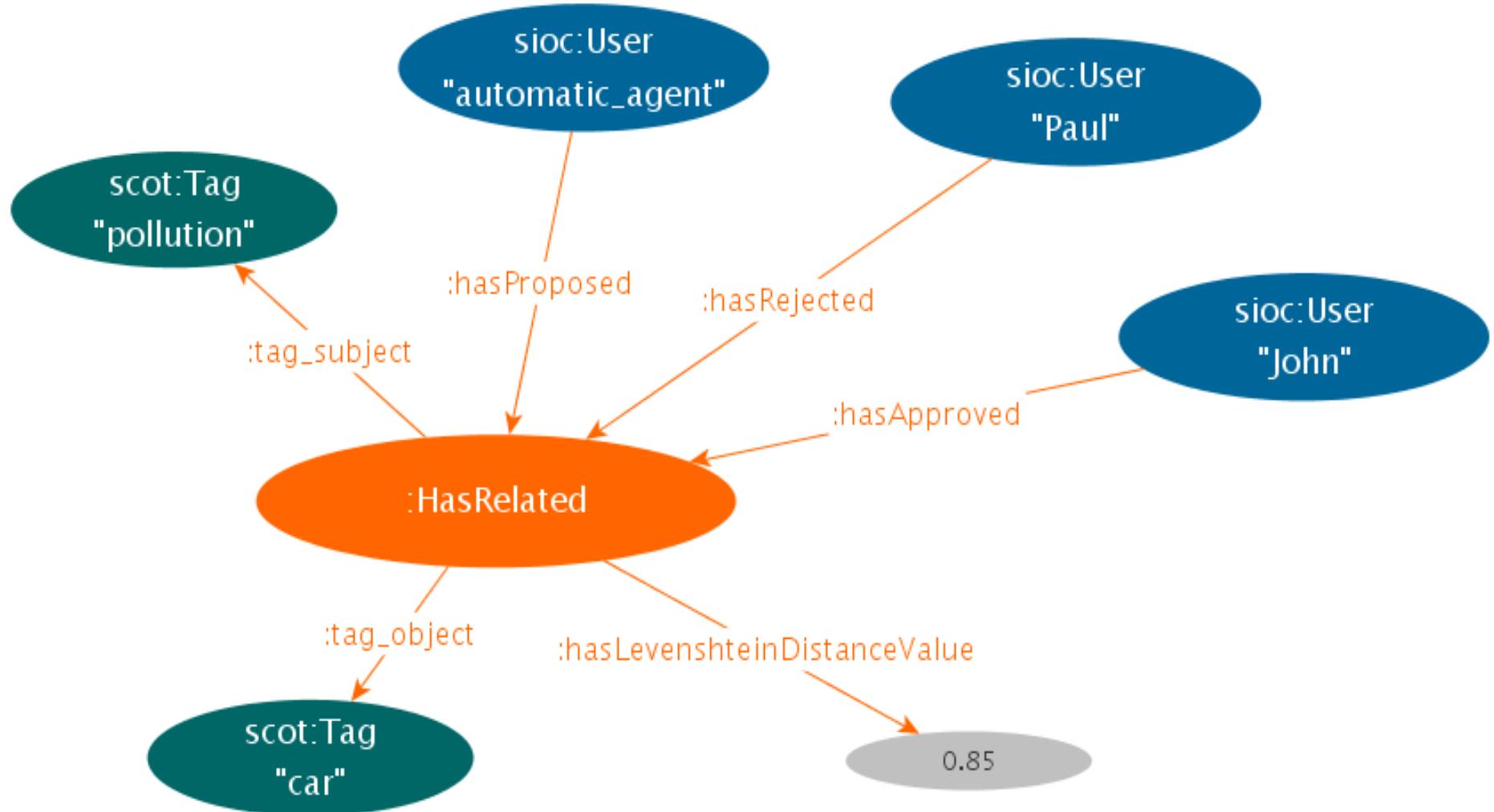


$t=T3$



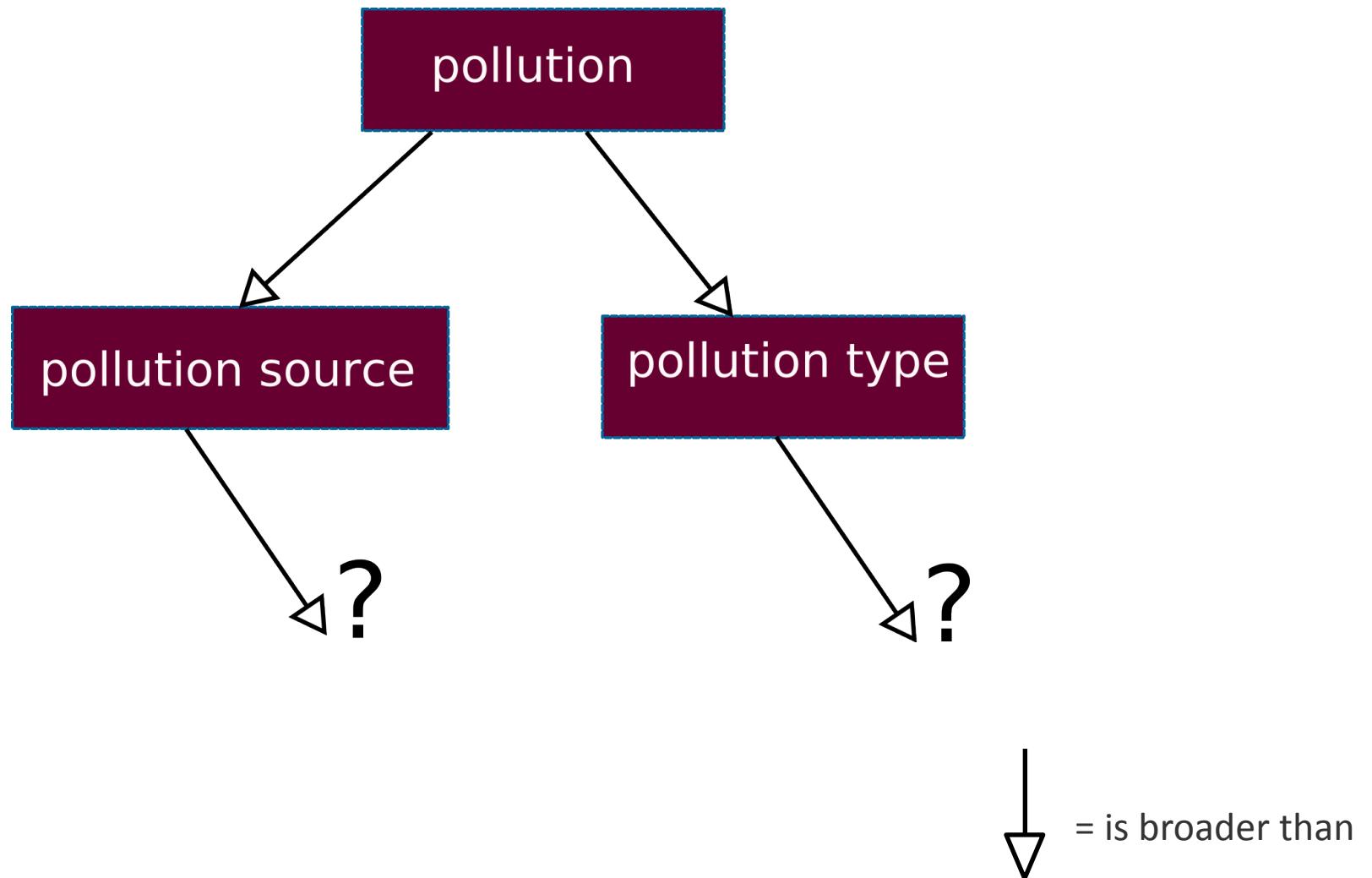
Micro-editing
(SRTags Editor)

Supporting diverging points of view



4. Applications

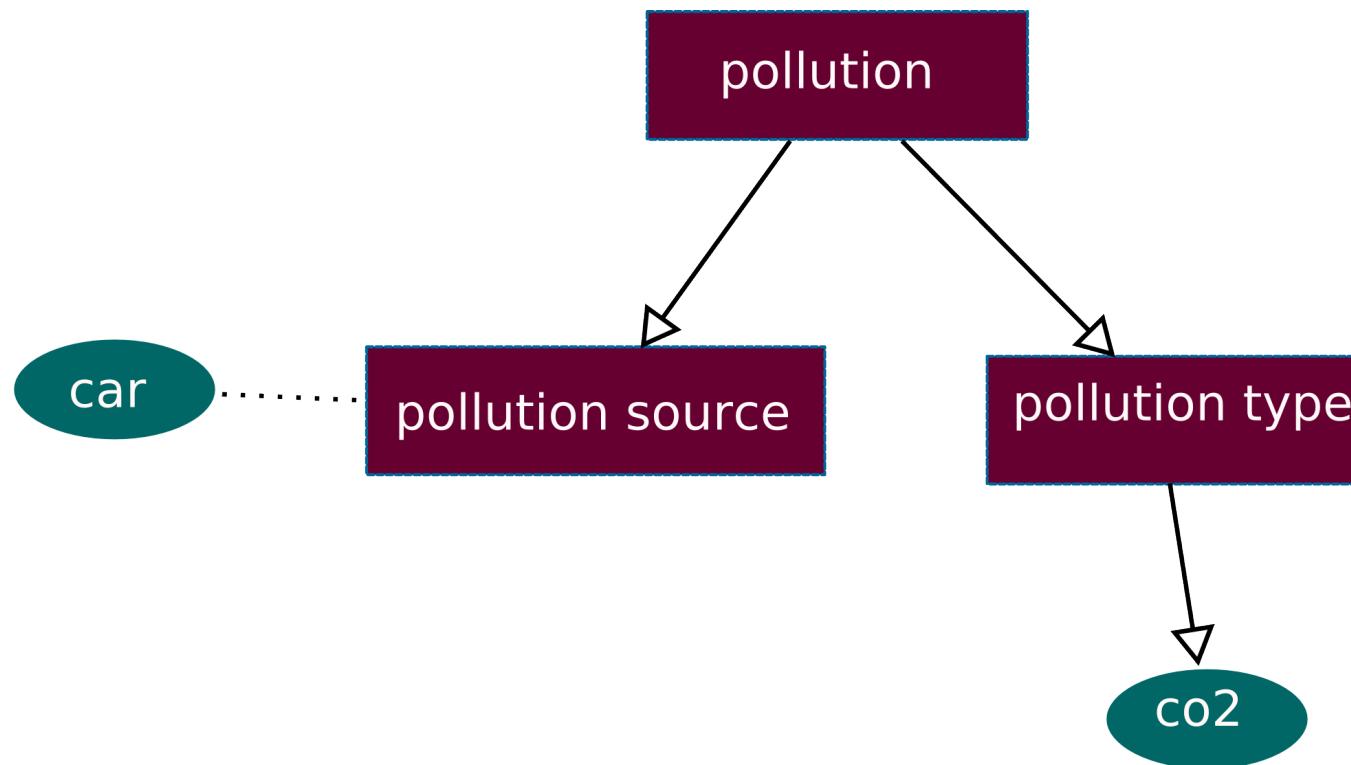
Case 1 : enriching in-house thesaurus



Case 1 : enriching in-house thesaurus

- Most people agreed on a link between "pollution" and "co2"
- Some people disagree on a direct link between "pollution" and "car"

Case 1 : enriching in-house thesaurus



Case 2 : detecting sub-communities of interest

→ For John, "car" and "pollution" have a link,
but not for Paul

>>> 2 sub-communities :
- the "cars' opponents"
- the "cars' defenders"

. Conclusion

Our contributions

- Model to organize collaboratively folksonomies (SRTags)
- Combining automatic handling of
folksonomies and users' expertise
- Micro-edition of folksonomies (SRTag-editor)

Futur works

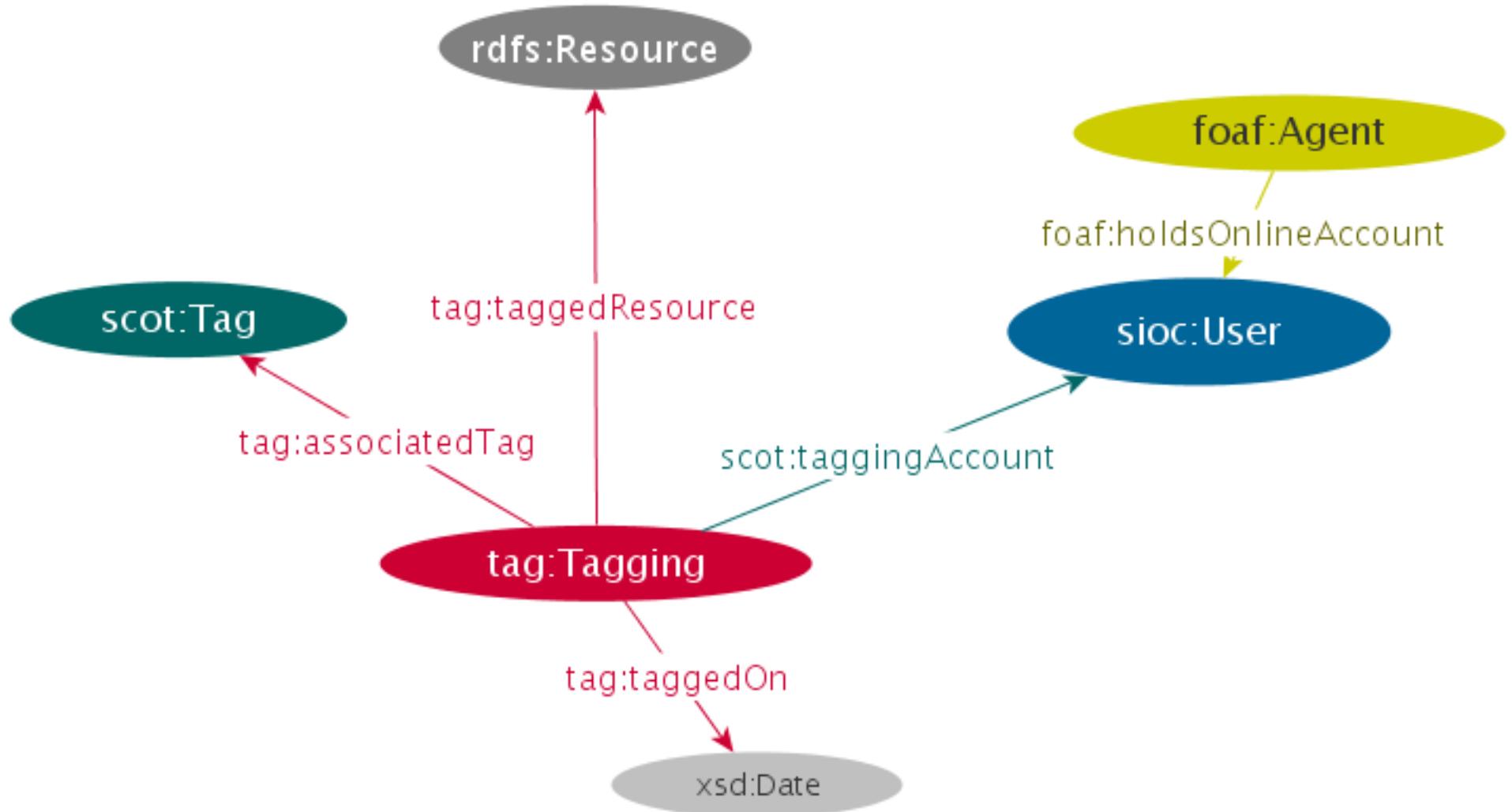
- integrating documents and ontological resources (bootstrapping)
- Exploiting individuals scale results and providing services and global visualizations
- Experiments with documentation services and thesaurus administrators

Thank you for your attention.

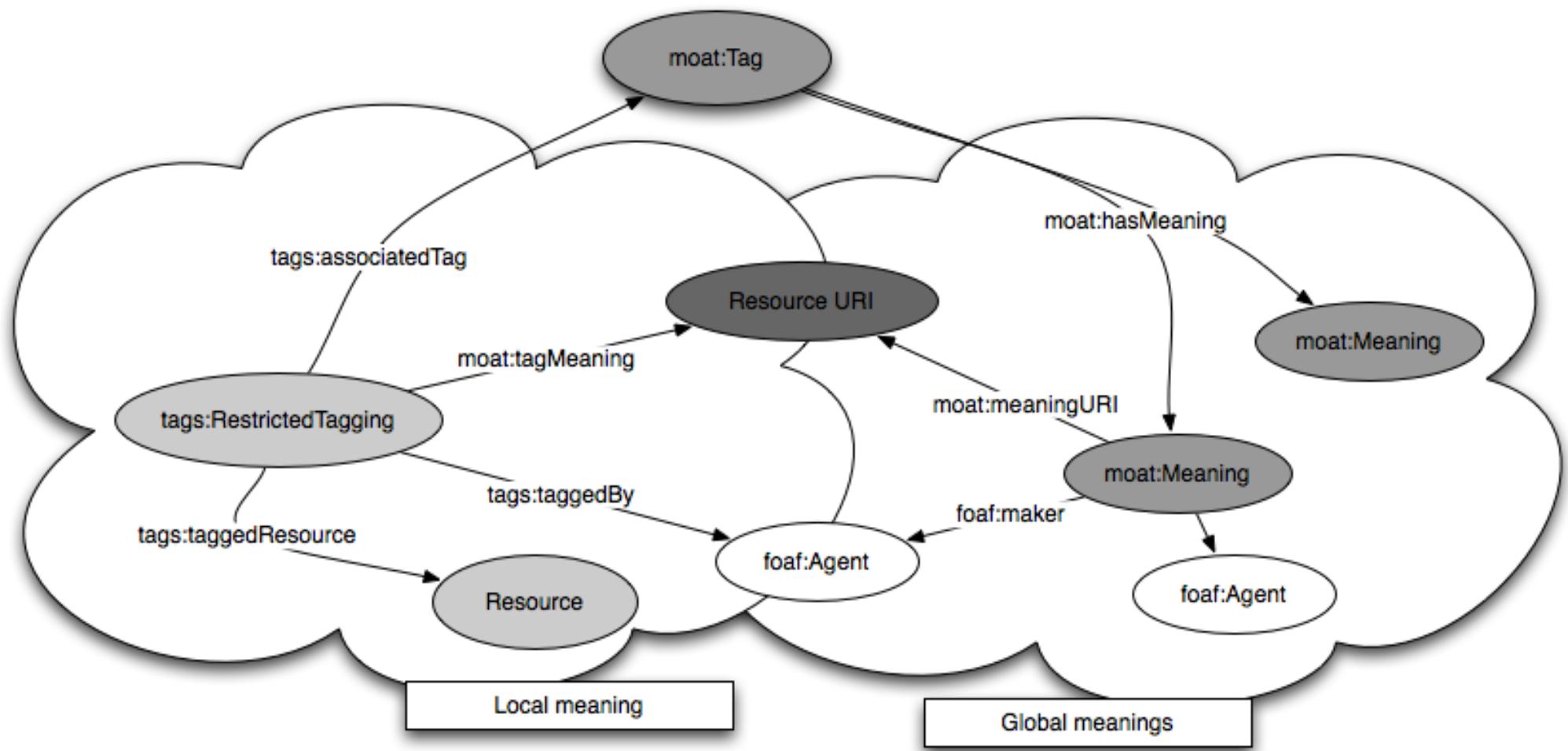
State of the art

- ☒ "web 2.0-like" ontologie construction [Braun *et al.*, 2007]
- ☒ Tagging tags [*Extreme tagging*, Tanasescu & Streibel, 2007]
- ☒ Collaborative and multi-points of view approaches
[Cahier *et al.*, 2007]
- ☒ Formalizing tags meanings [MOAT, Passant & Laublet, 2008]

Tagging resources



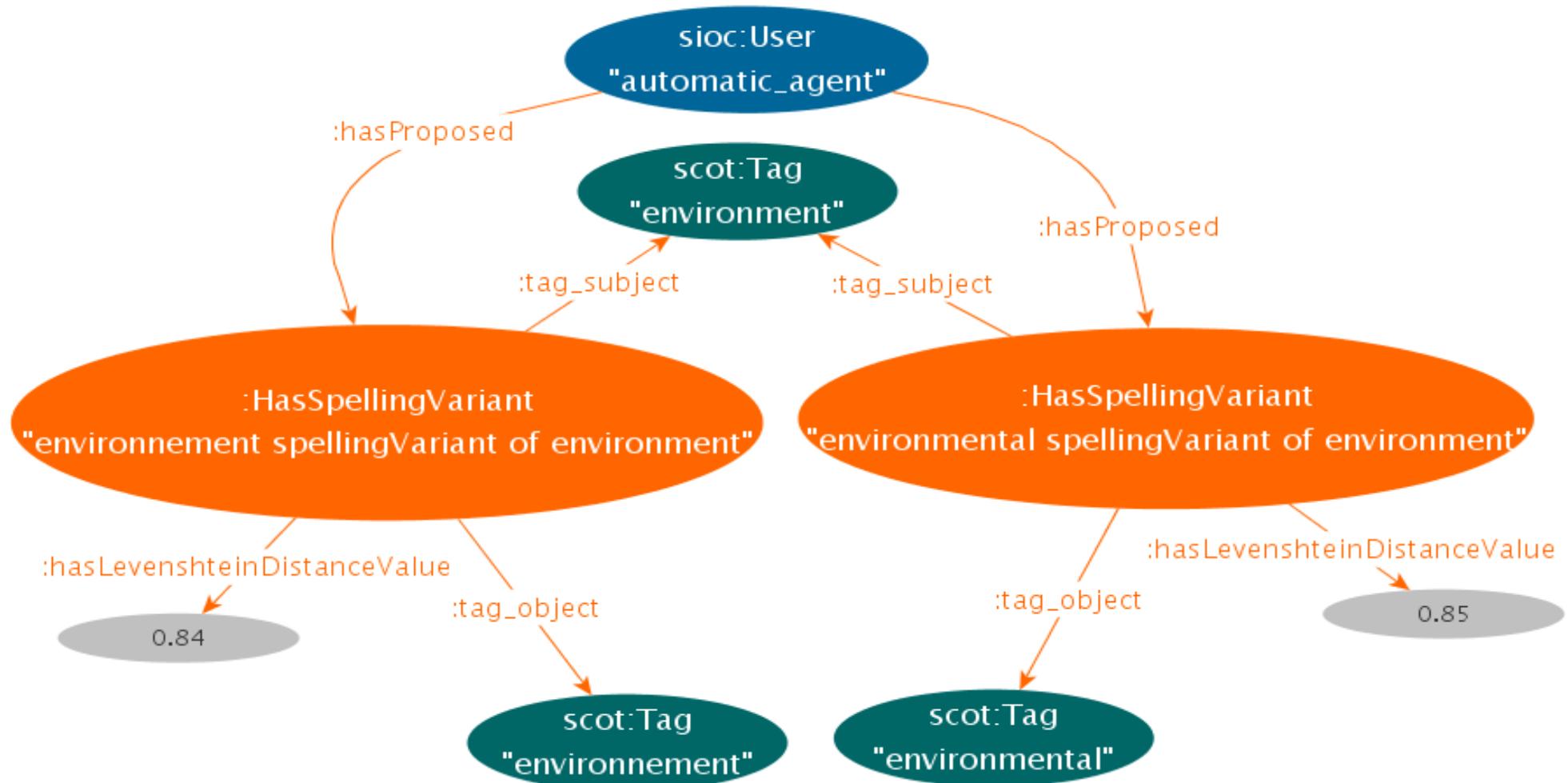
Linking tags to « meanings » : MOAT



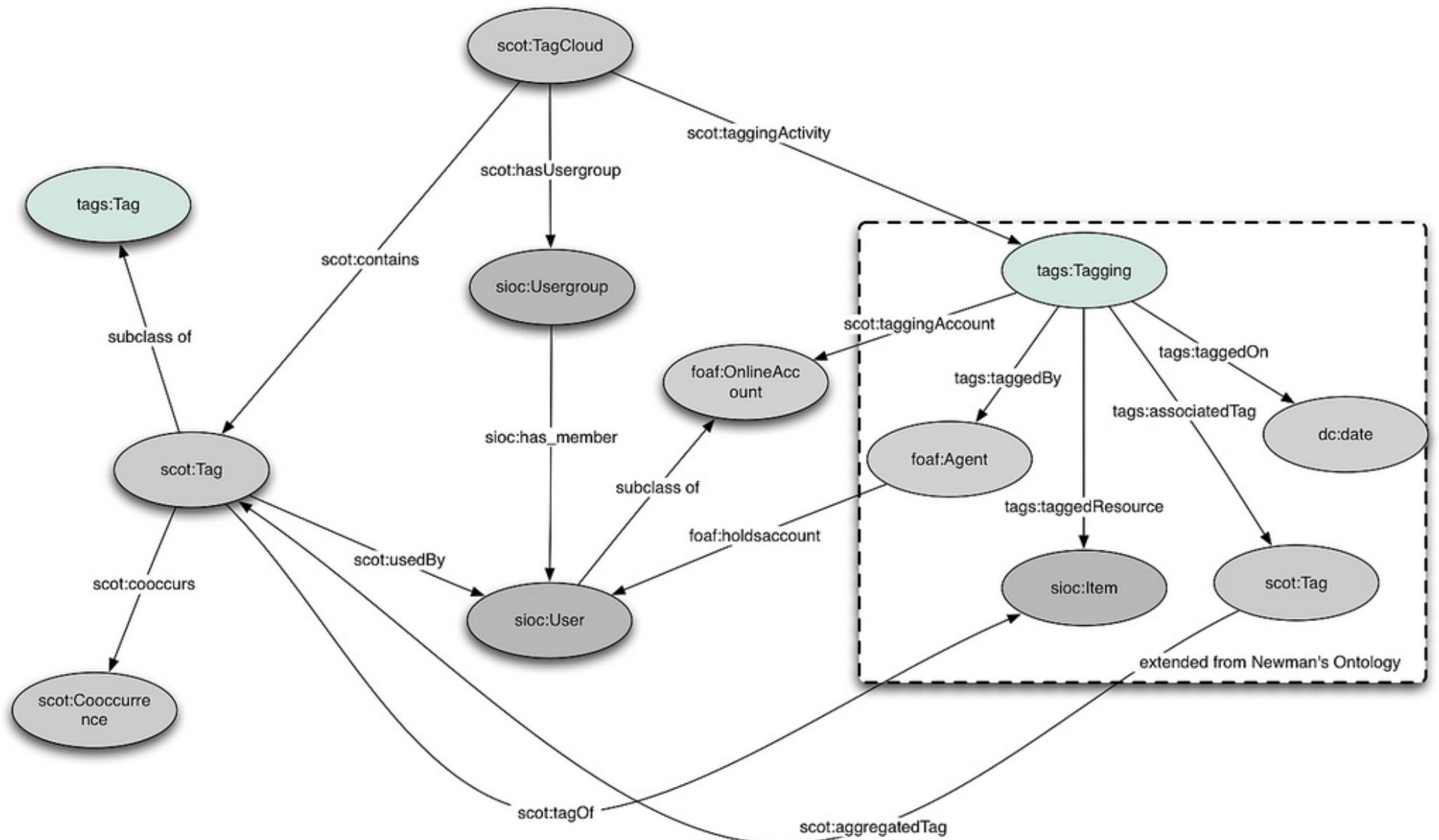
Detecting "spelling variants"

Tag1	Tag2	Levenshtein distance
geographie	geographique	0.83
environment	environnement	0.84
déchets	déchet	0.85
industrie	industriel	0.9
développementdurable	développement-durable	0.95

Detecting "spelling variants"



Système de gestion de bookmarks



Model de tagging > SCOT