IWAISE 2012

LET TAGGING BE MORE INTERESTING

Samia Beldjoudi[1], Hassina Seridi[2], Catherine Faron-Zucker[3],

[1,2] Laboratory of Electronic Document Management LabGED, Badji Mokhtar University Annaba, Algeria

[3] KEWI Group, I3S Laboratory, University of Nice, France

Plan

- Introduction
- Folksonomies and the social web
- Description of the proposed approach
- Experiment Results and Evaluation
- Conclusion and Future works

INTRODUCTION

- Web 2.0 technologies have created the conditions for new usages on the web which has become a social web.

- Users create, annotate, share and make public what they find interesting on the web.

WEB 2.0 AND THE SOCIAL ASPECT

Web 2.0 is the web of users

Folksonomy is a web 2.0 technology

Folksonomies

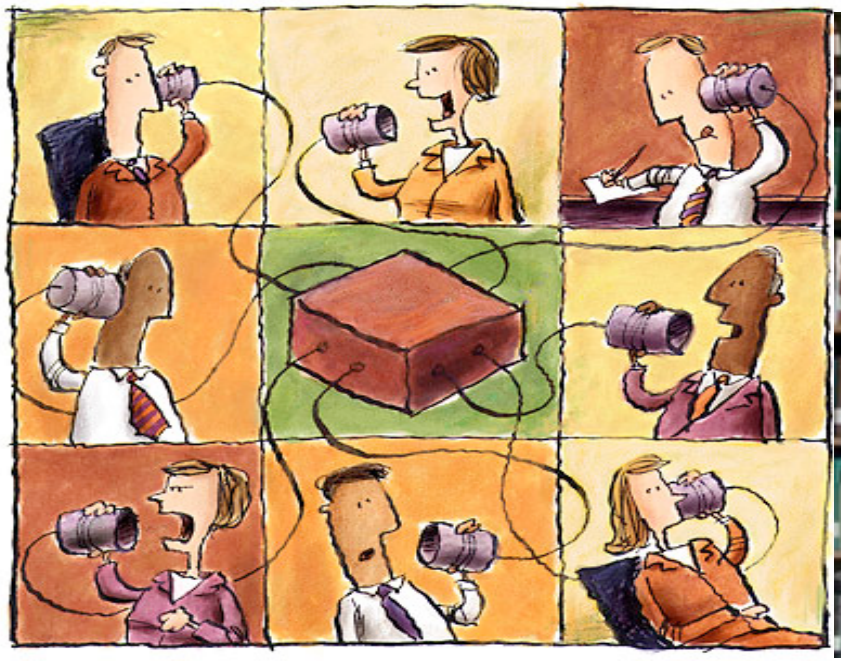
=

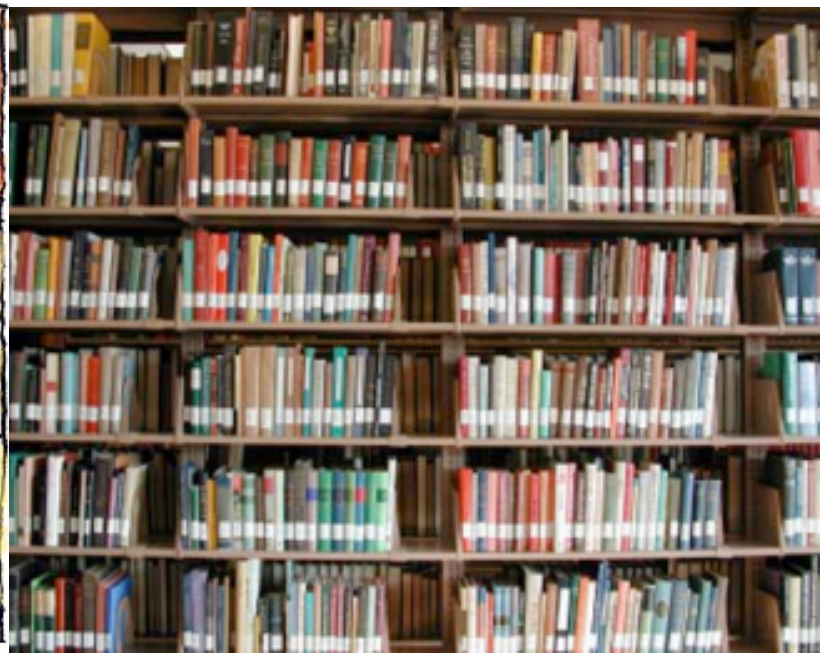
Collaborative Tagging

=

Social Tagging

FOLKSONOMY IS A WEB 2.0 TECHNOLOGY





Folks + **Taxonomy**

Folksonomy

=

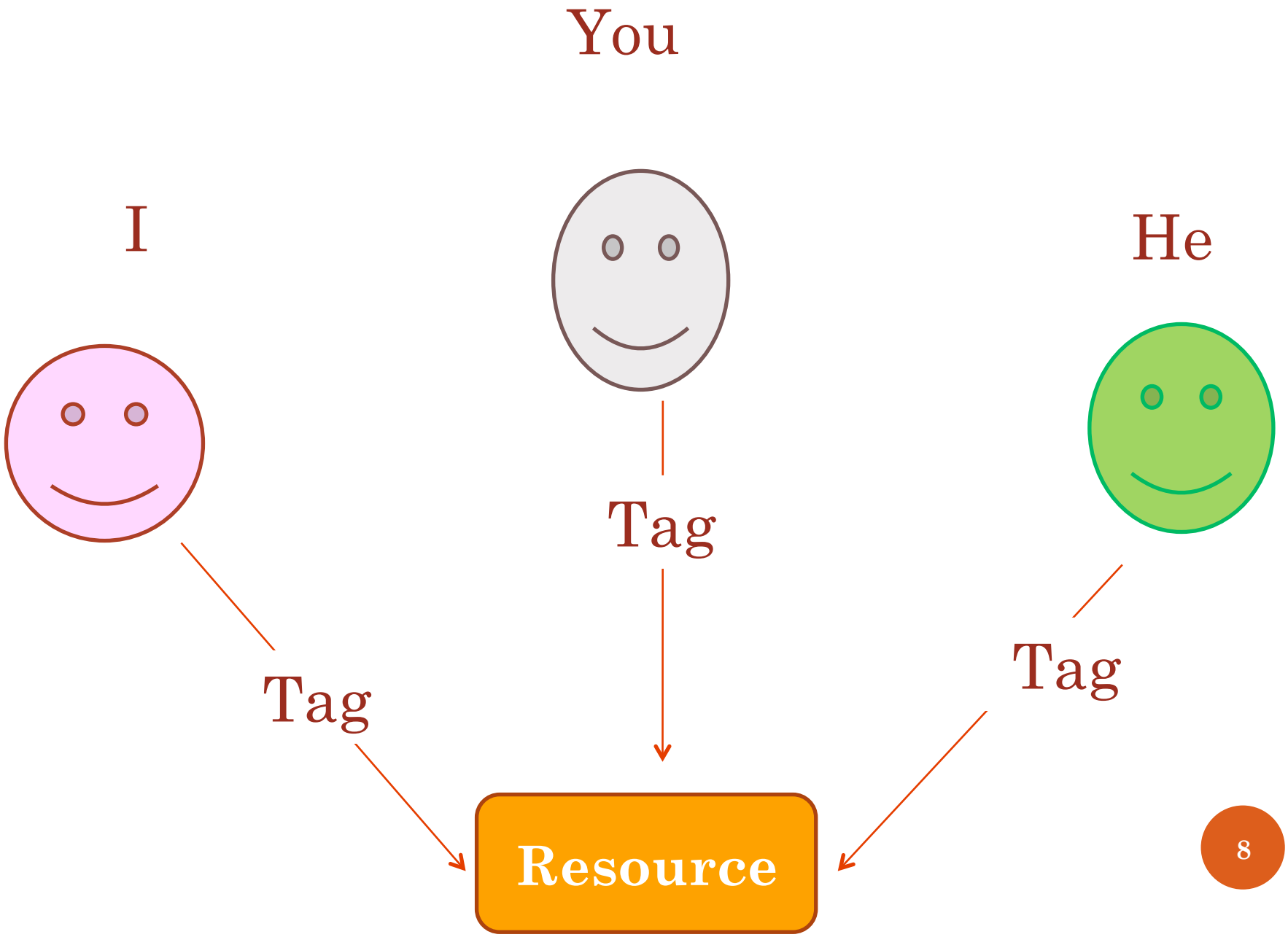
Tag (Keyword)

+

Tagger (User)

+

Resource


You
I
He
Tag
Tag
Tag
Resource

Examples of the best popular folksonomies

- Del.icio.us: www.del.icio.us
- Flickr: www.flickr.com
- YouTube: www.youtube.com
- Dailymotion: www.dailymotion.com
- Myspace: www.myspace.com
- Odeo: www.odeo.com

ADVANTAGES

Simple, easily and instantaneous human indexation

Social aspect

Gain time and economize the expertise of users

INCONVENIENT

Tags' Ambiguity (Polysemy)

Apple ➔ Fruit

Apple ➔ Society

INCONVENIENT

Spelling Variations (Synonymy)

Cat =? Chat

Related work

- Resource recommandation: [De Meo et al. 2010]
- Tag recommandation: [Schmitz et al. 2006]
- Resolving tags' ambiguity: [Mika 2005] [Gruber 2005] [Buffa et al. 2008] [Pan et al. 2009] [Limpens 2011] …

➔**These approaches are promising but expensive and not trivial; also they didn't personalize the tag-based recommendation in folksonomies according to each user profile.**

OBJECTIVES (1)

- Personalization & Socialization:

➔ Join together the semantic and the social aspect in folksonomies.

➔The proposed idea aimed to allow each community's member to benefit from resources judged similar to his preferences.

OBJECTIVES (2)

- Personalization & Recommendation:

➔Resources' Recommendation in folksonomies

➔Resolving the Tags' ambiguity problem

CONTRIBUTION

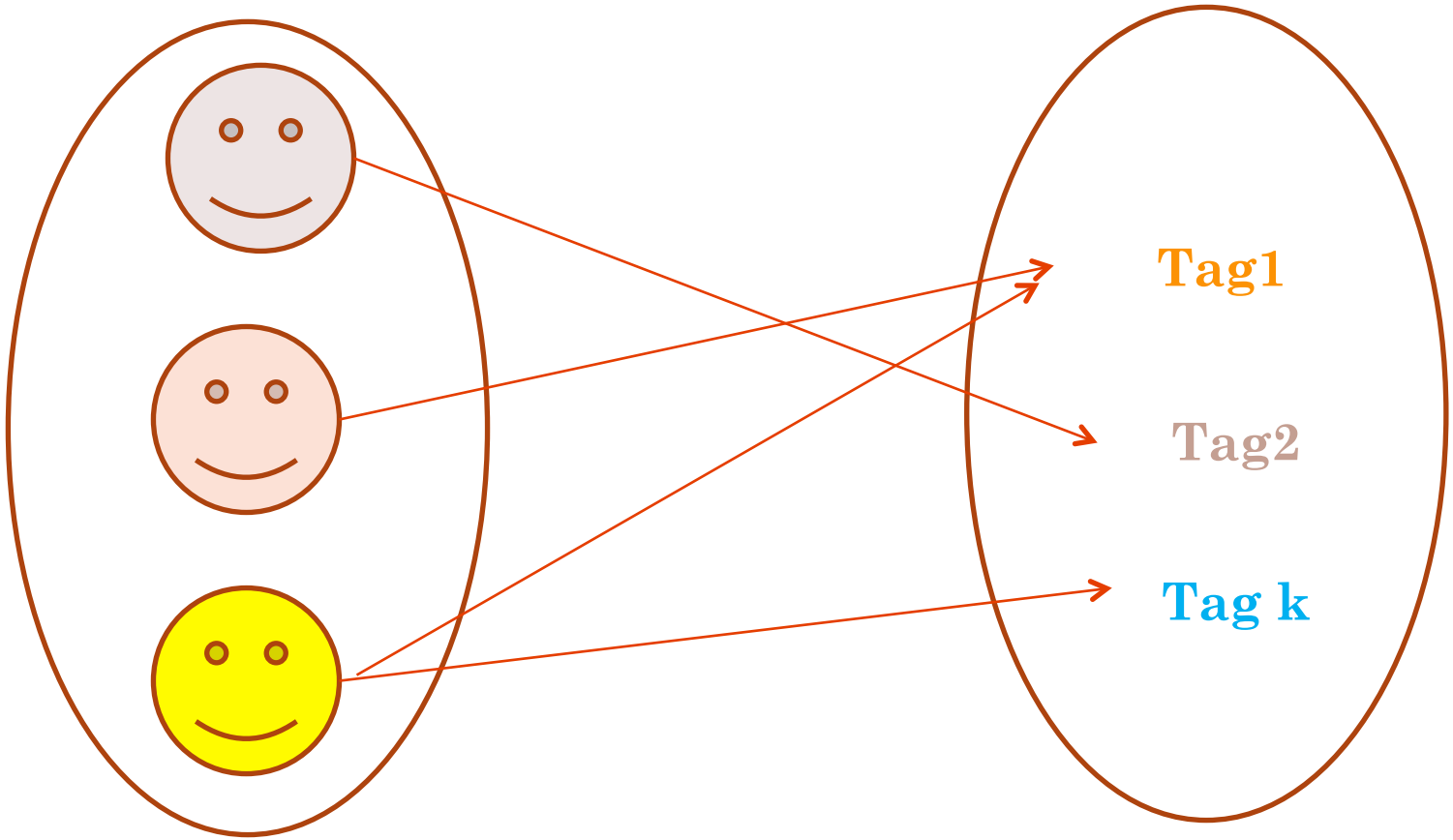
- Resource recommendation based on association rules, without soliciting the user's expertise
- Personalizing and Improving Tag-Based Recommendation in Folksonomies
- Resolution of tag ambiguity based on social similarities without explicitly using ontologies
- Tags classification based on association rules

Formal description of folksonomies

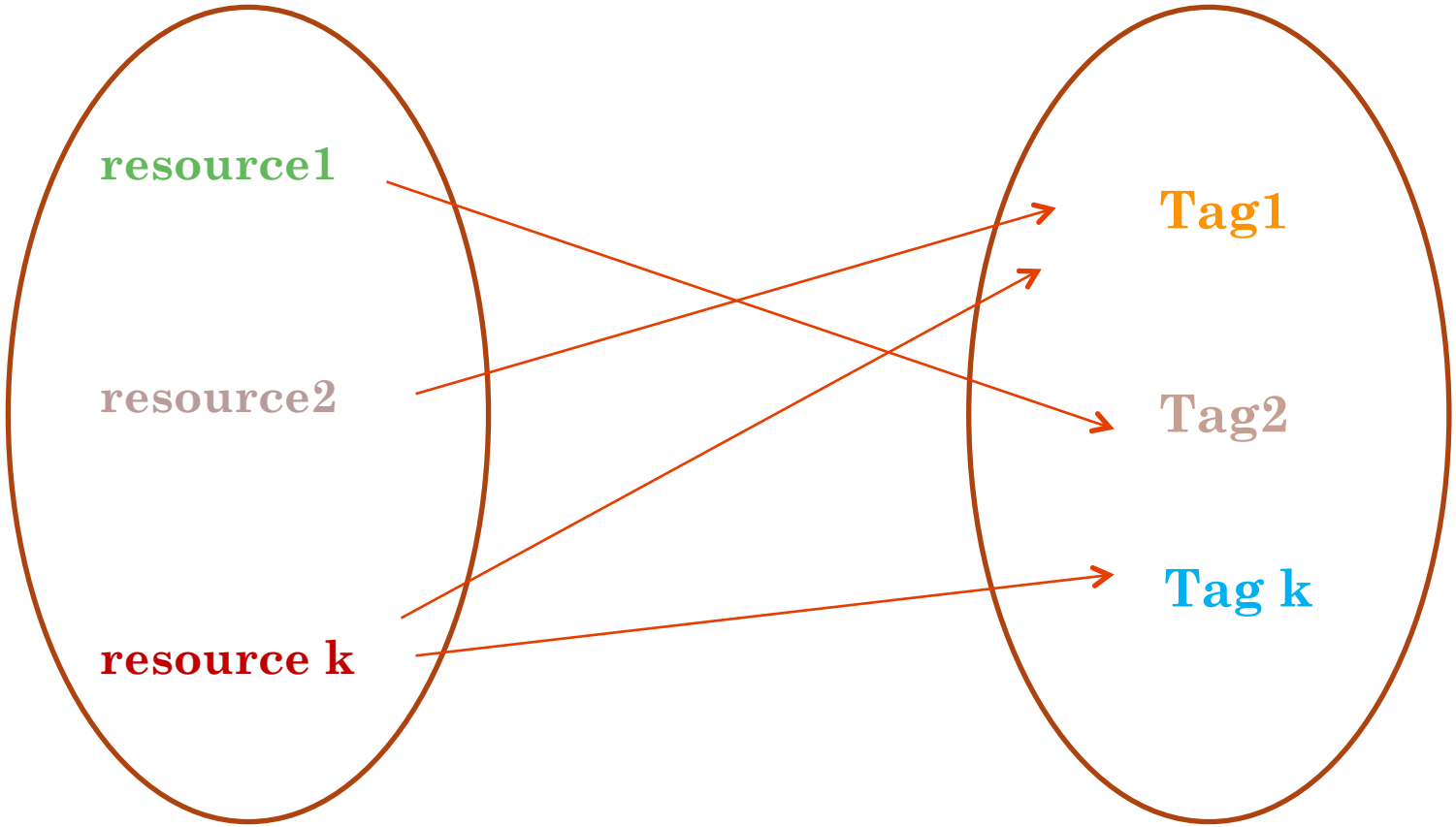
- A folksonomy: <U, T, R, A> with:

  - U: user
  - T: tag
  - R: resource
  - $A \subseteq U \times T \times R$

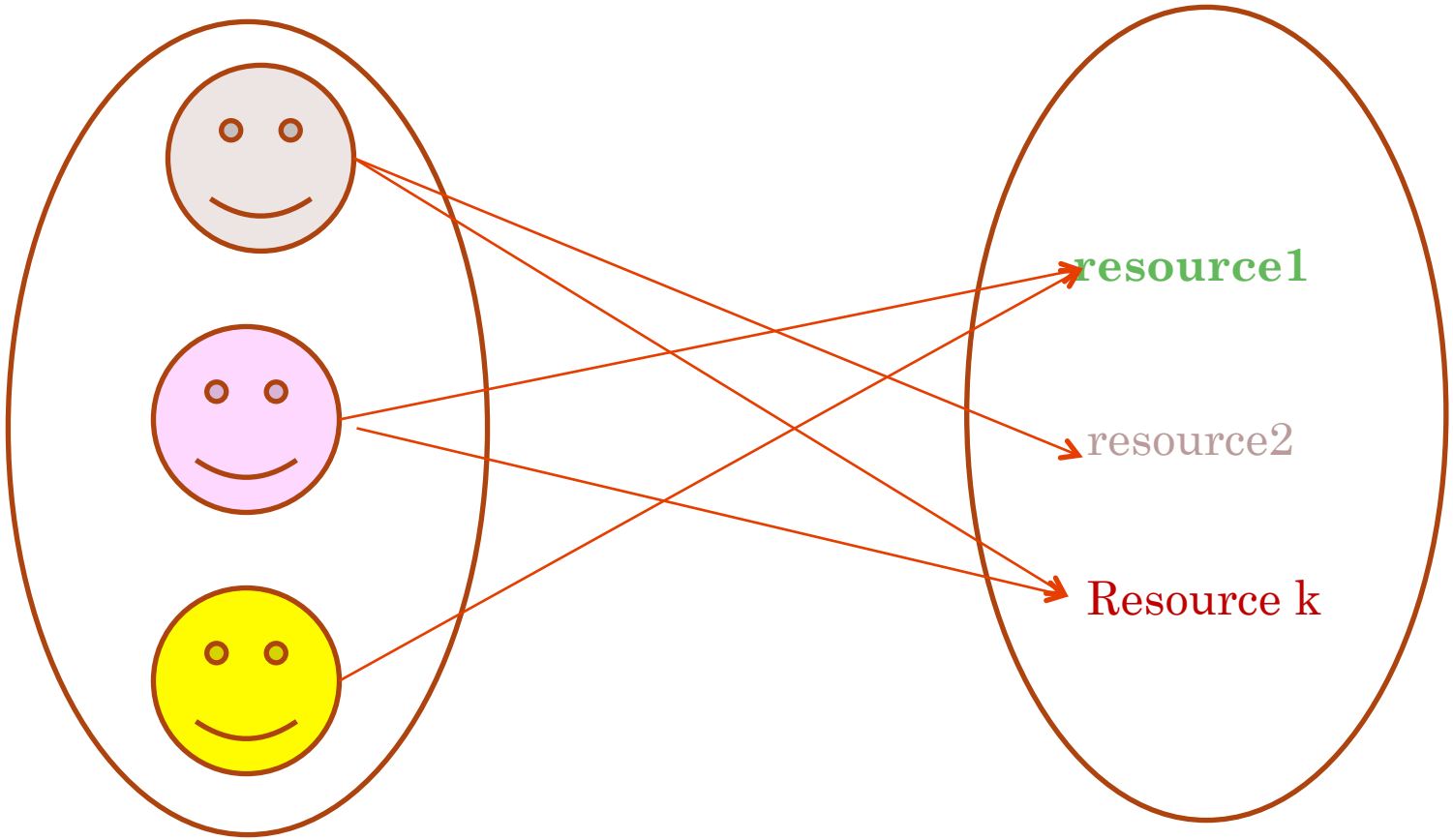
THE FIRST NETWORK



THE SECOND NETWORK



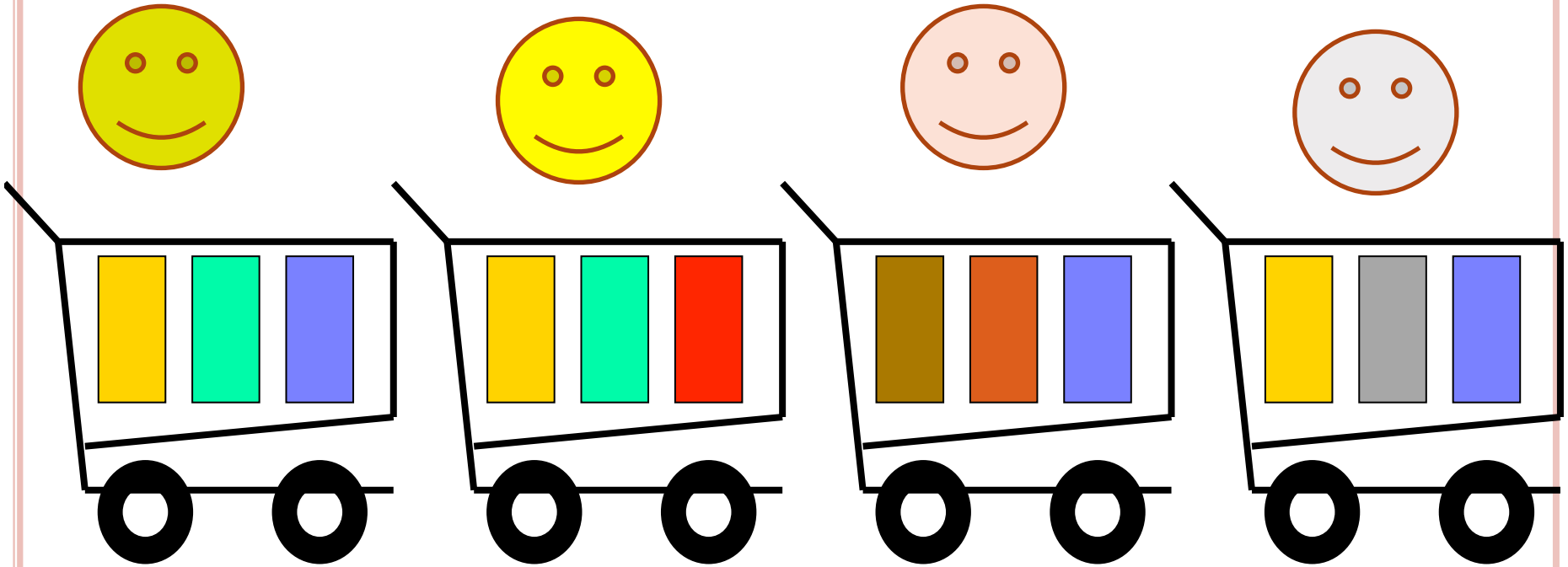
THE THIRD NETWORK



FORMAL DESCRIPTION OF FOLKSONOMIES

- TU = $[X_{ij}]$ where $X_{ij} = 1$ if $\exists\ r \in R, \langle u_i, t_j, r \rangle \in A$
  0 otherwise
- TR = $[Y_{ij}]$ where $Y_{ij} = 1$ if $\exists\ u \in R, \langle u, t_i, r_j \rangle \in A$
  0 otherwise
- UR = $[Z_{ij}]$ where $Z_{ij} = 1$ if $\exists\ t \in R, \langle u_i, t, r_j \rangle \in A$
  0 otherwise

ASSOCIATION RULES



Association rules in Data mining

Transaction ID	ItemSet
1	Bread, Cream, Water
2	Cream
3	Bread, Cream, Milk
4	Water
5	Cream, Water

Association rule: Bread ⇒ Cream:

indicates that if a customer buys Bread, he is likely to also buy Cream

ASSOCIATION RULES IN FOLKSONOMIES

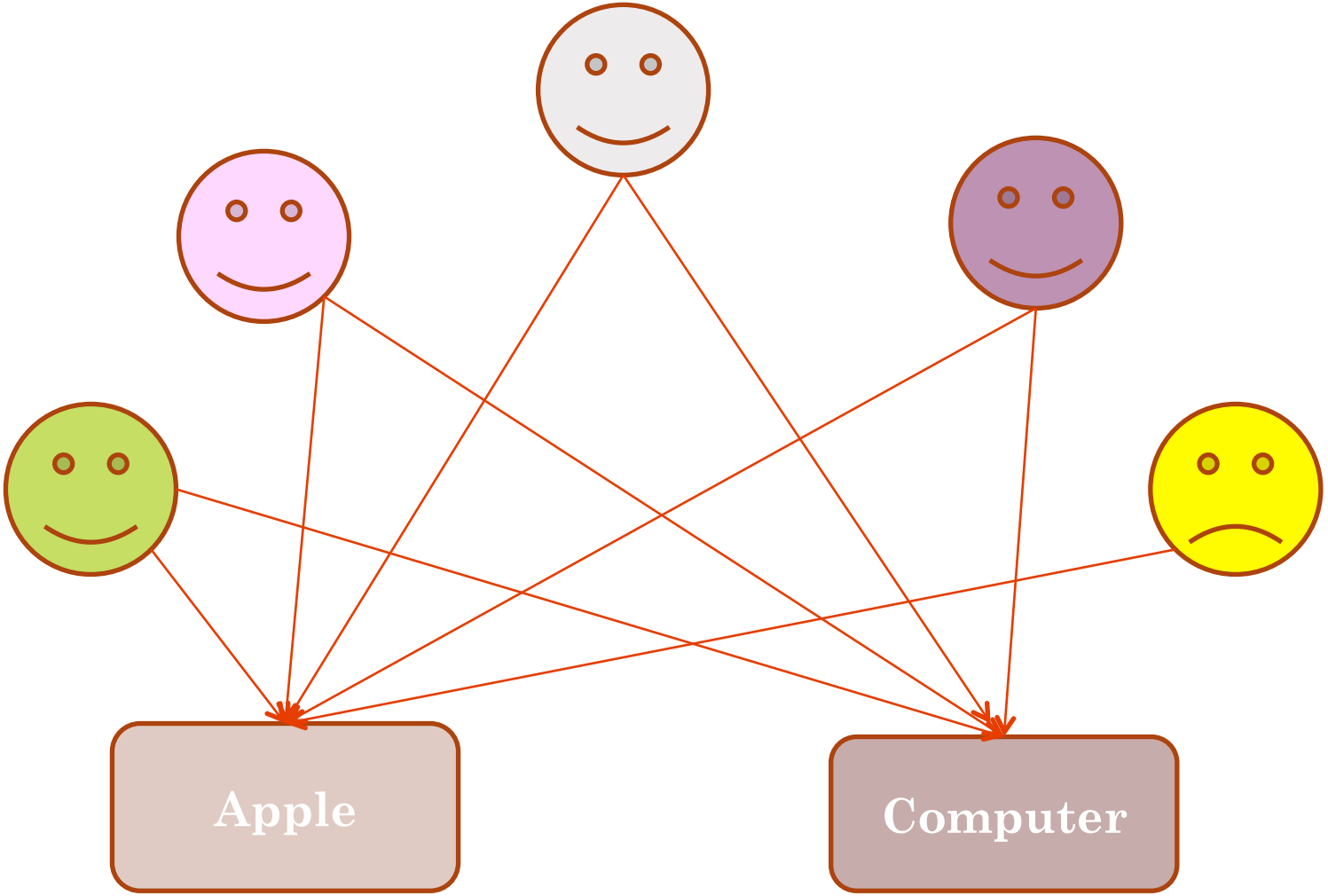
Transaction ID	ItemSet
User1	Computer, Programming, Apple
User2	Computer, Apple
User3	Kitchen, Apple
User4	Programming
User5	Kitchen

- Extraction of association rules on tags from TU
  - e.g. Apple ⇒ Computer

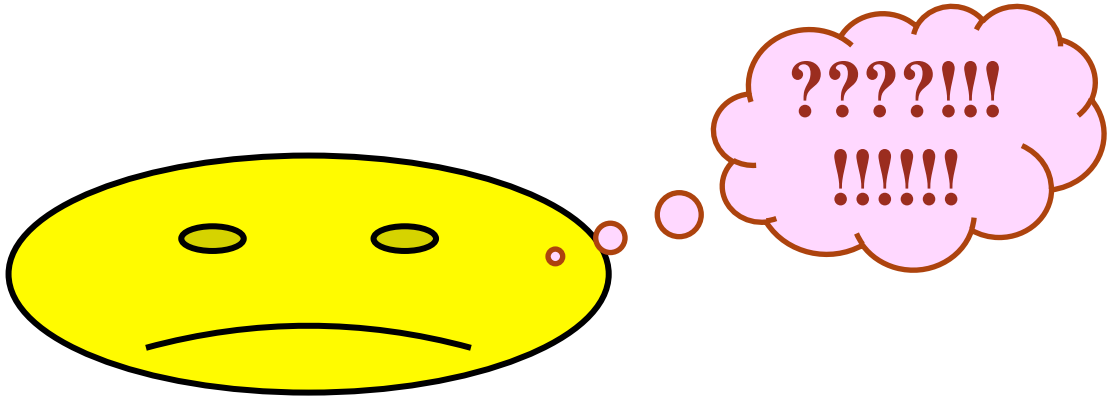
Resource Recommandation

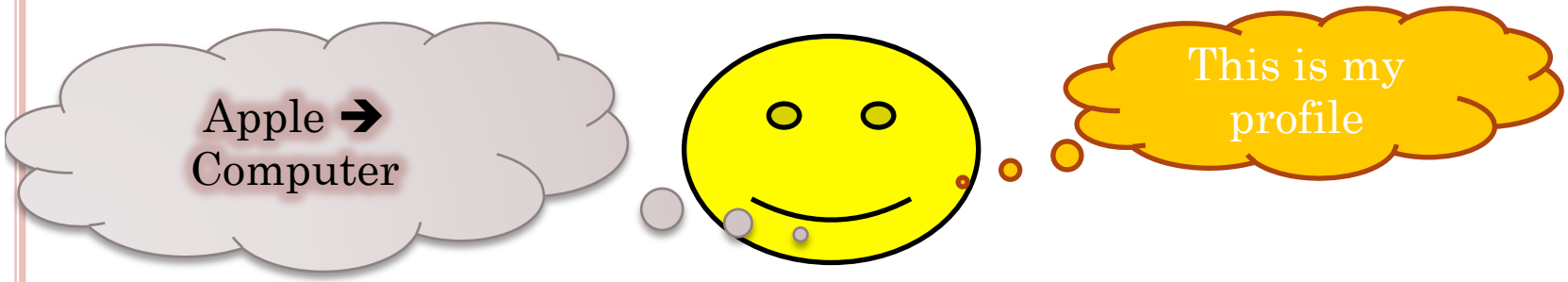
- Based on association rules on tags :

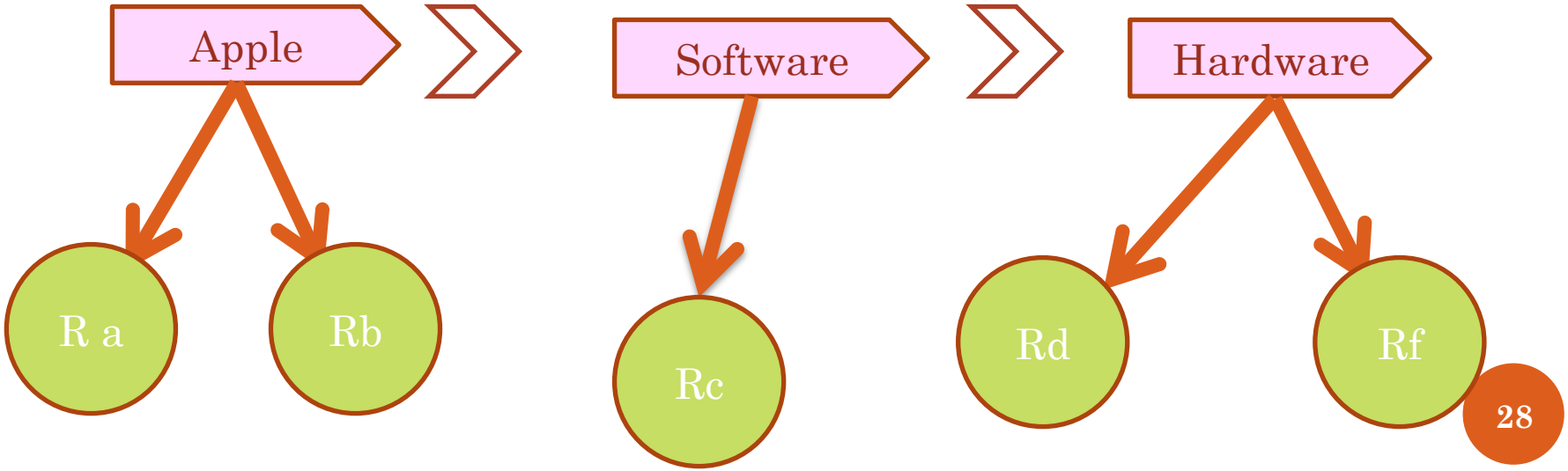
A user may be recommended the resources associated to the tags occurring in the consequent of the association rules which antecedent contain his tags

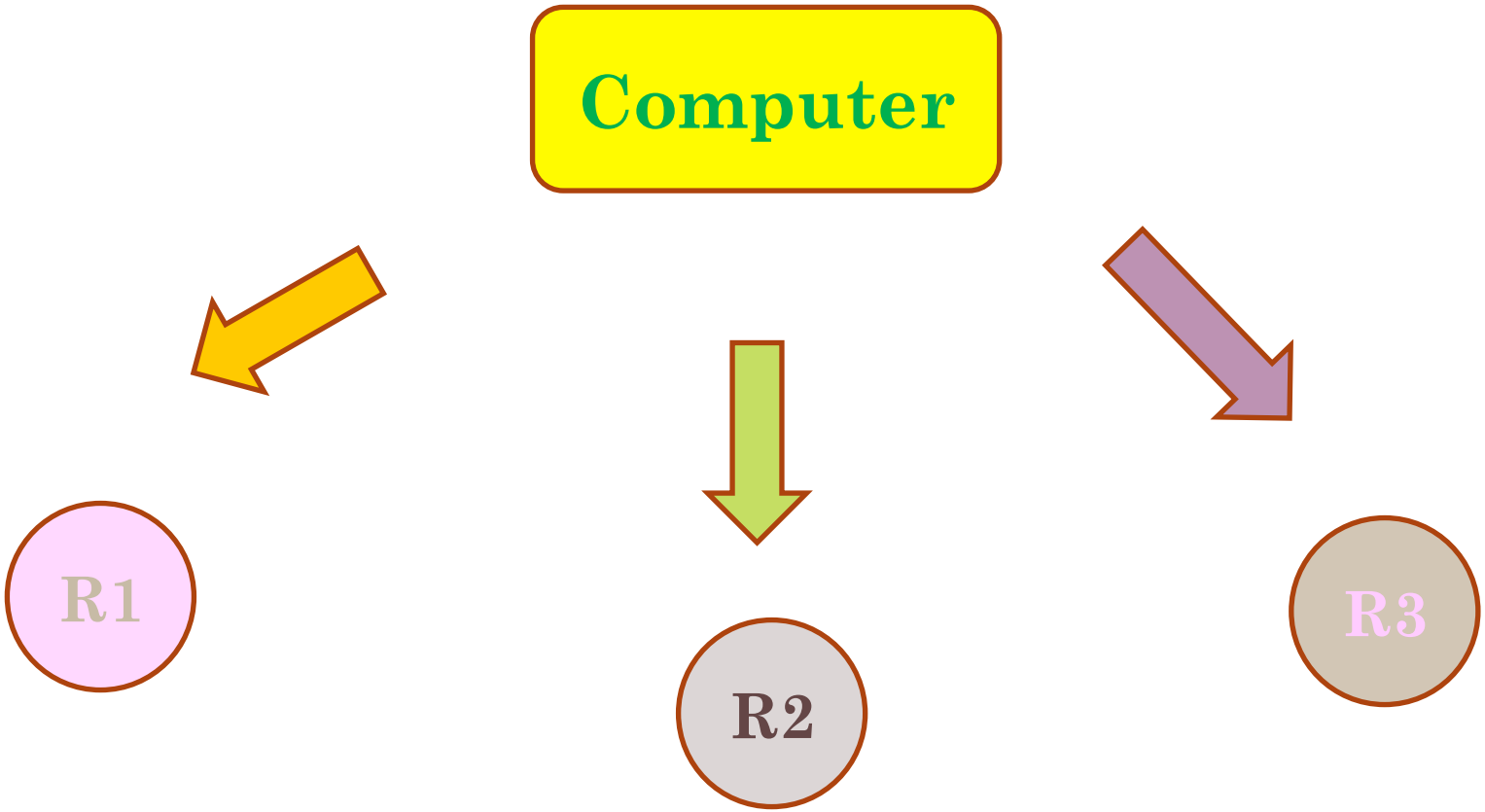

Apple
Computer

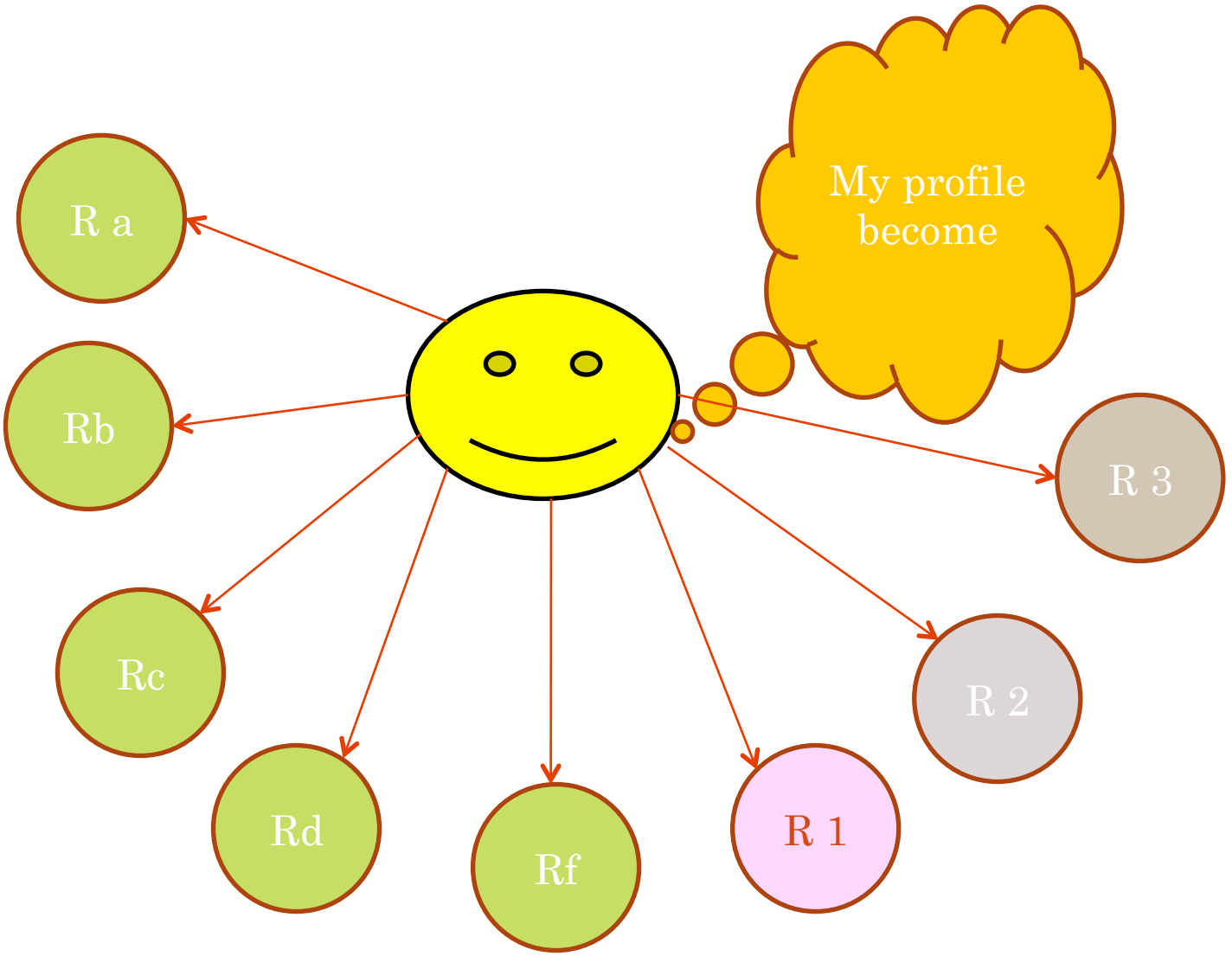
Apple → Computer



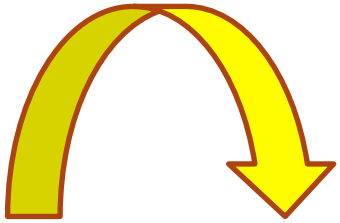

Apple ➔
Computer
This is my
profile


Apple
Software
Hardware
R a
Rb
Rc
Rd
Rf


Computer
R1
R2
R3


My profile become
R a
Rb
Rc
Rd
Rf
R 1
R 2
R 3

Resource Recommandation



The effectiveness of the recommendation depends on the resolution of tag ambiguity

RESOLVING TAG AMBIGUITY

- Measuring the similarity between users, to specify those who have similar preferences.

- Similarity between users:

  $sim(u_1, u_2) = cos(v_1, v_2)$ with v_1, v_2 extracted from UR

- Levels of recommendation depending on the similarity between users are associated to any proposed resource. Each either resource is highly recommended; simply recommended or weakly recommended according to profile of each user.

TO AVOID THE COLD START PROBLEM

- Similarity between resources:

  $sim(r_1, r_2) = cos(v_1, v_2)$ with v_1, v_2 extracted from TR

- A user u_1 is recommended the resources associated to the tags occurring in his query if these resources are close to those already recommended to him.

TAGS' CLASSIFICATION

- **Equivalent tags:**

The two tags T1 and T2 are equivalents (i.e. T1 ↔ T2) iff (T1 →T2) and (T2 → T1).

- **Directe related tags:**

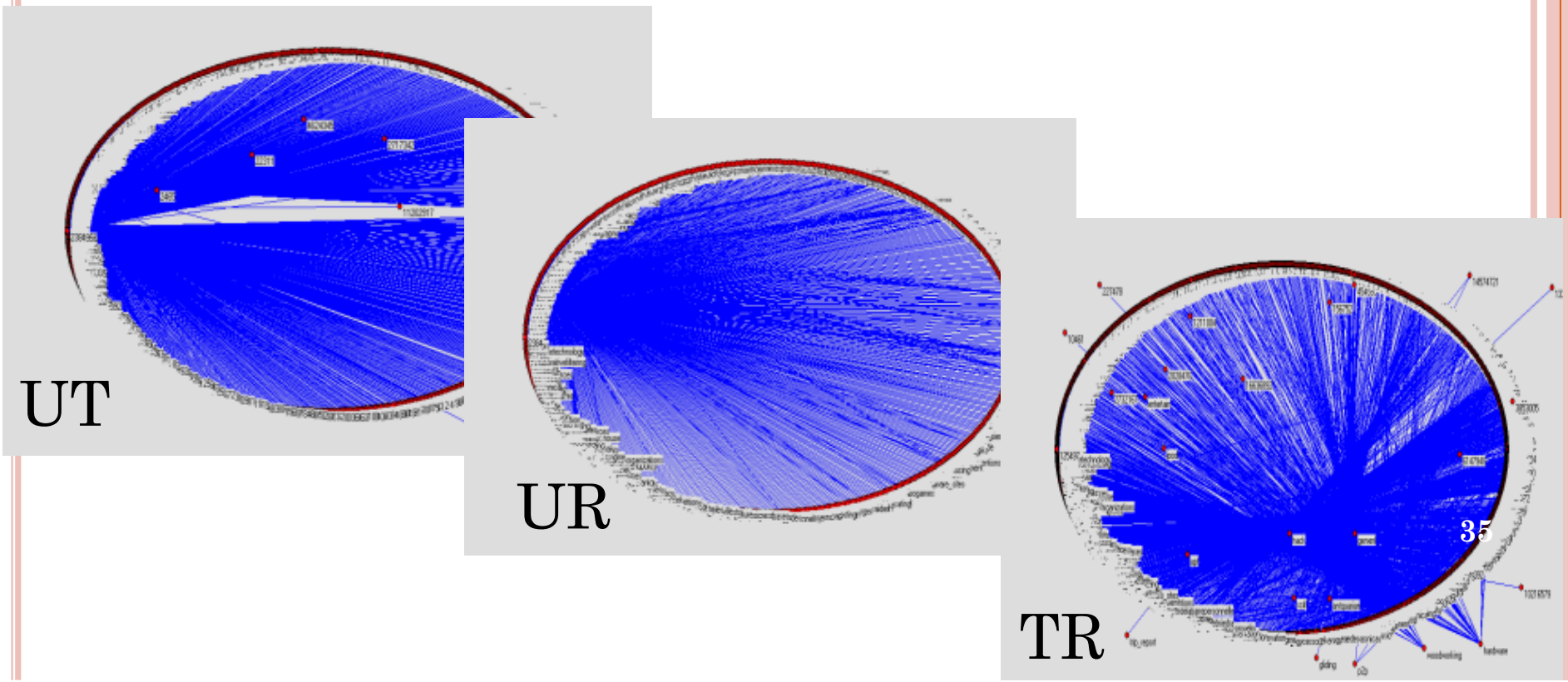
The two tags T1 and T2 are directly related (T1 is directly related to T2) iff (T1 → T2) or (T2 → T1).

- **Indirecte related tags:**

The two tags T1 and T3 are indirectly related (T1 is indirectly related to T3 i.e.: T1 → T3) iff (T1 → T2) and (T2 → T3).

EXPERIMENTS WITH DEL.ICIO.US

- 507 tag assignments involving 51users, 239 tags ,112 resources
- Pajek



EXPERIMENTS WITH DEL.ICIO.US

- Extraction of 65 association rules with $supp_{min} = 0.5$ and $conf_{min} = 0.6$

  Apple ⇒ Computer

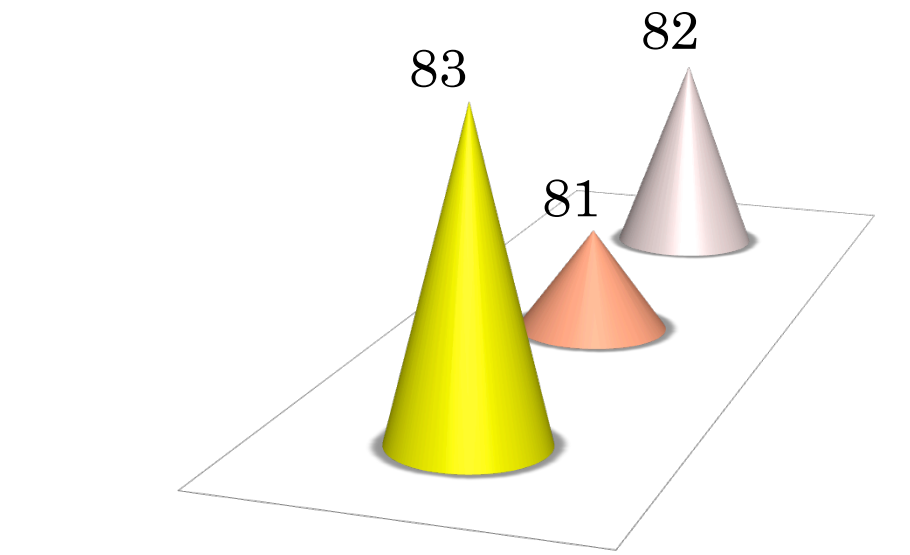
  60% of the users using the tag *apple* also use the tag *computer*

Evaluation

- We distinguish between ambiguous tags and non ambiguous tags
- In rules involving non ambiguous tags, resources associated to these tags are highly recommended
- In rules involving ambiguous tags, resources close to the user interest are highly recommended and those far from his interests have a low level of recommendation

RESULTS

The average of the three metrics



Conclusion

- Personalizing and Improving Tag-Based Recommendation in Folksonomies
- Resolving tag ambiguity without explicitly using ontologies
- Associating levels of recommendation based on similarities between users

FUTURES WORKS

- Enrich the generated association rules by other measures which will help to specify the relevance degree of each rule to each user .

- Validate our approach on larger databases.

Thank
you!