Web browsing support for cross-community activities

Tomohiro Oda

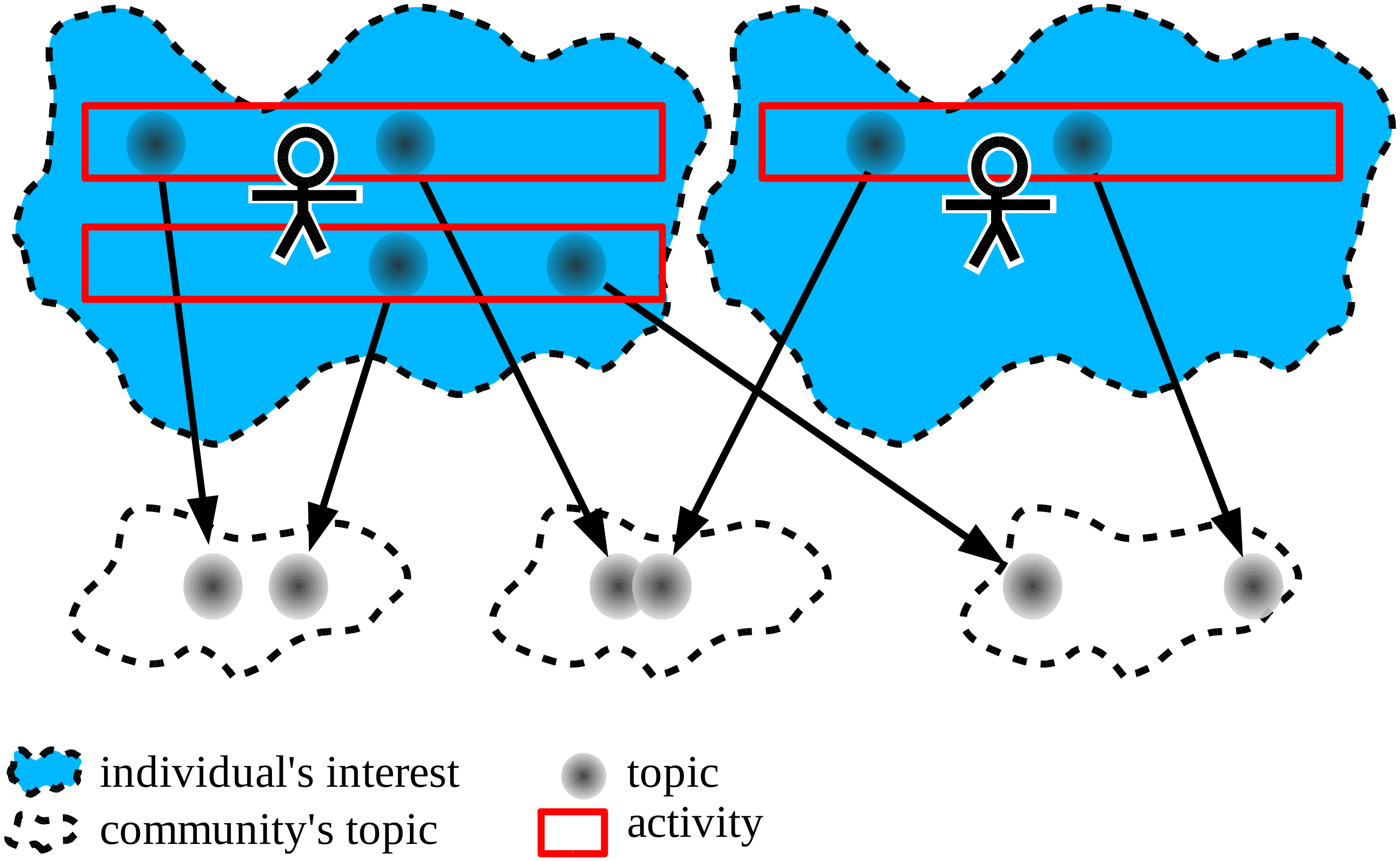
Agenda

- cross-community activity
- cross-community activity and DynC
- difficulties in supporting cross-community activities
- cSuite: web browsing support tool for cross-community activities
- cSuite for DynC

Cross-community activity

- Definition:
  - Activity either
    - needs support of multiple communities, or
    - contributes to multiple communities.
- Examples:
  - standard graph format
  - developing OpenGL interface in Smalltalk for CAD system of a ship constructor.

Communities, individuals, activities, and interests



individual's interest

community's topic

topic

activity

Comparison with DynC

- Similarities
  - Focused on each individual's tasks or activities
  - not for community, but for individuals
- Differences
  - supportive community v.s. supportive person
  - assuming pre-existing communities
    v.s. forming a new short-term community

Difficulties in supporting cross-community activities

- A task needs knowledge of multiple communities.
  - None of each community covers the whole task.
  - It is hard to identify/describe the task from each community's viewpoint.
- It is difficult to recommend collaborators/related artifacts
  - different motivations, interests, and goals on a same topic

Example difficulties: web browsing

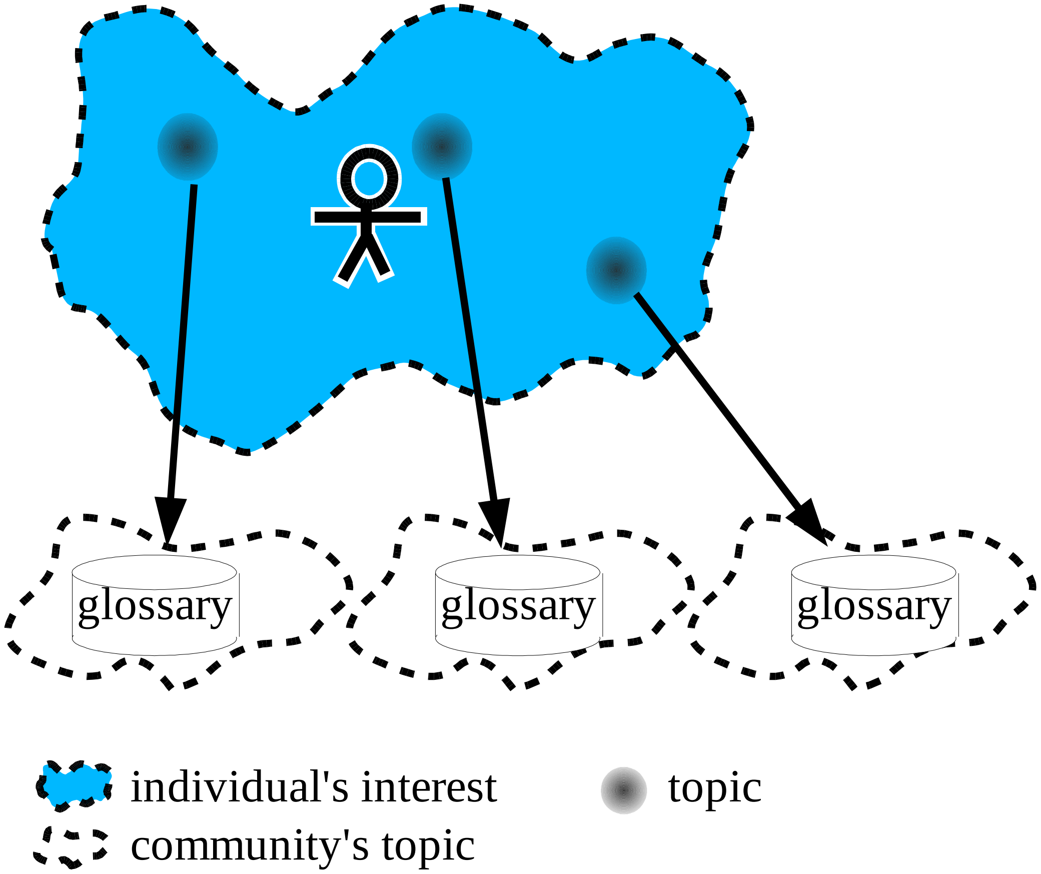
- difficult to identify tasks
  - Browsing a community's website does not mean the user is working on a task covered by the community.
    e.g. A CAD programmer is reading the C-99 specification. Does the C language community cover CAD programming?
- difficult to recommend collaborators/related artifacts
  - Browsing the same document does not mean sharing the same interests and goals.
    e.g. Two programmers are reading HOW-TO of Linux-2.6 device driver. One is a FreeBSD kernel hacker, and another is an ethernet board manufacturer.

cSuite: cross-community support using HTTP proxy

- Each community provides "glossary" as community's knowledge.

- A user specifies a list of glossary servers that the user is interested in.

- cSuite provides additional information to HTML documents.

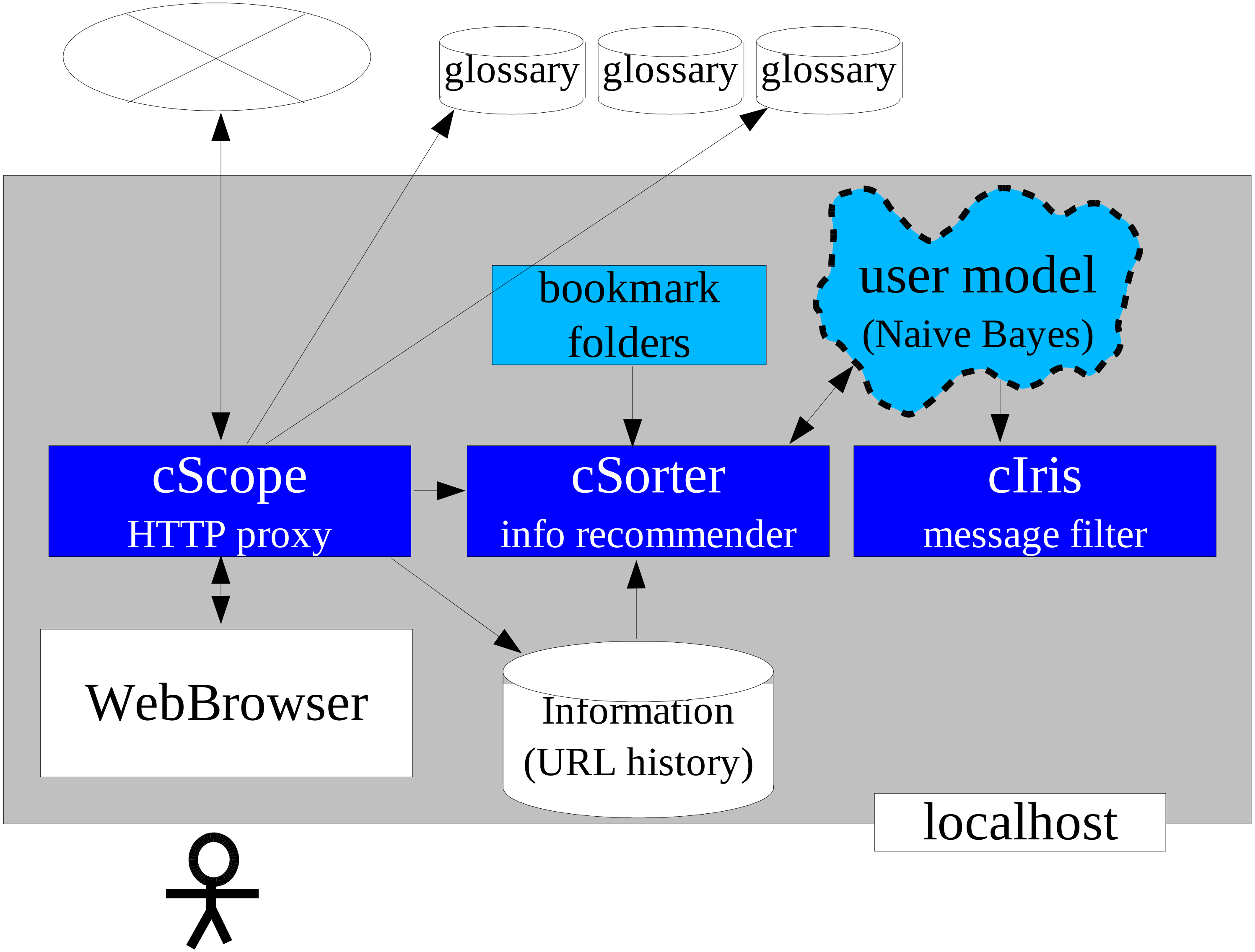


The development of cSuite is sponsored by IPA, Japan.

Basic ideas of cSuite

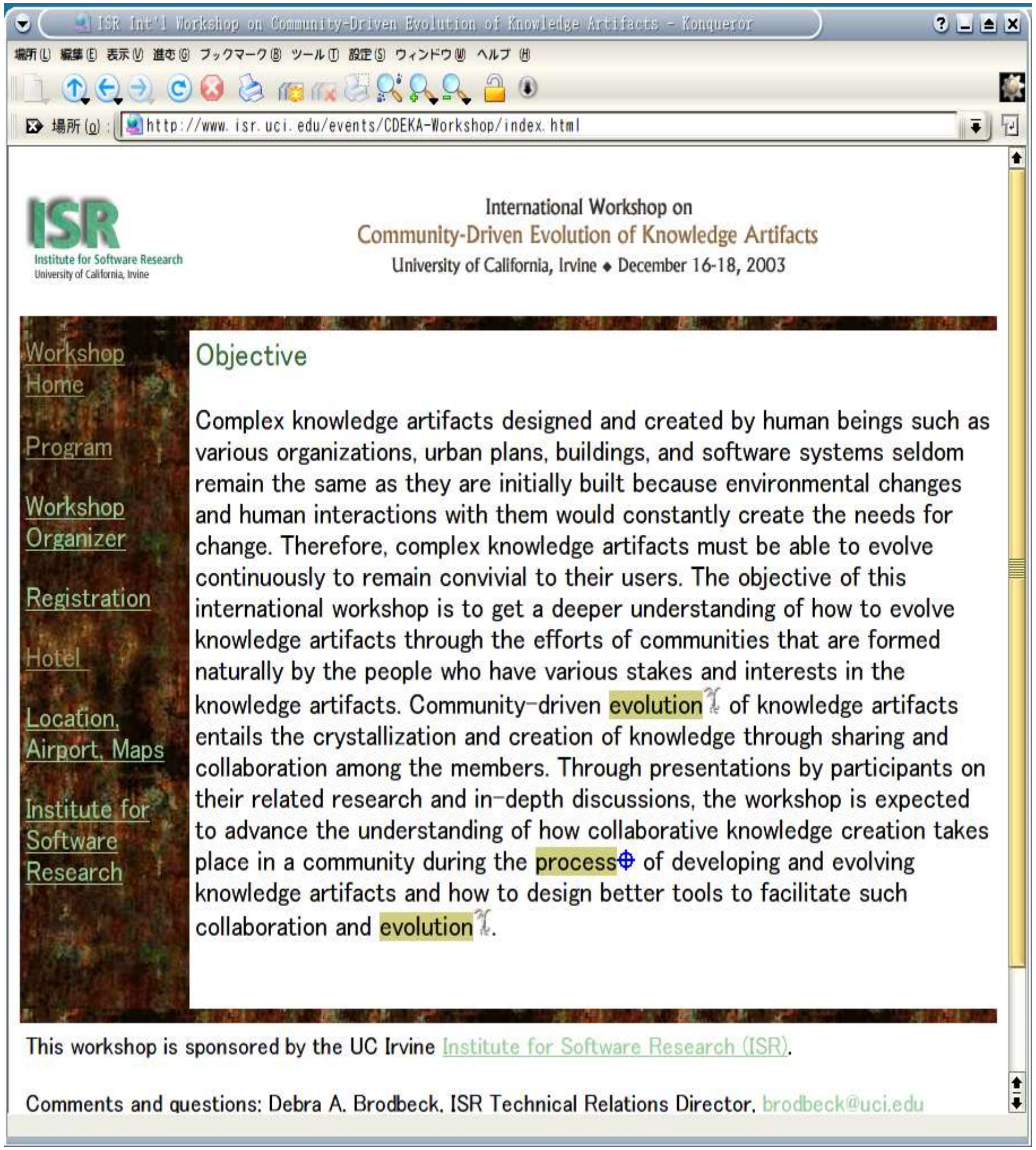
- One possible way to identify user's task and to find supportive persons/related documents:
  - Words are very important clues of user's tasks.
  - Many communities provide their glossaries as
    - FAQs
    - Tutorials
  - Natural Language Processing techniques like
    - Text classification
    - Word disambiguation

Architecture of cSuite



cScope: HTTP proxy

- cScope is a private HTTP proxy server which works on localhost.
- cScope wiretaps all "GET" requests and returned HTML documents.
- cScope inserts icons to each occurrence of keywords.
- Each icon represents a community.


ISR Int'l Workshop on Community-Driven Evolution of Knowledge Artifacts - Konqueror
場所(L) 編集(E) 表示(V) 進む(G) ブックマーク(B) ツール(T) 設定(S) ウィンドウ(W) ヘルプ(H)
場所(O):
http://www.isr.uci.edu/events/CDEKA-Workshop/index.html
ISR
Institute for Software Research
University of California, Irvine

International Workshop on
Community-Driven Evolution of Knowledge Artifacts
University of California, Irvine ◆ December 16-18, 2003

Workshop Home
Program
Workshop Organizer
Registration
Hotel
Location, Airport, Maps
Institute for Software Research

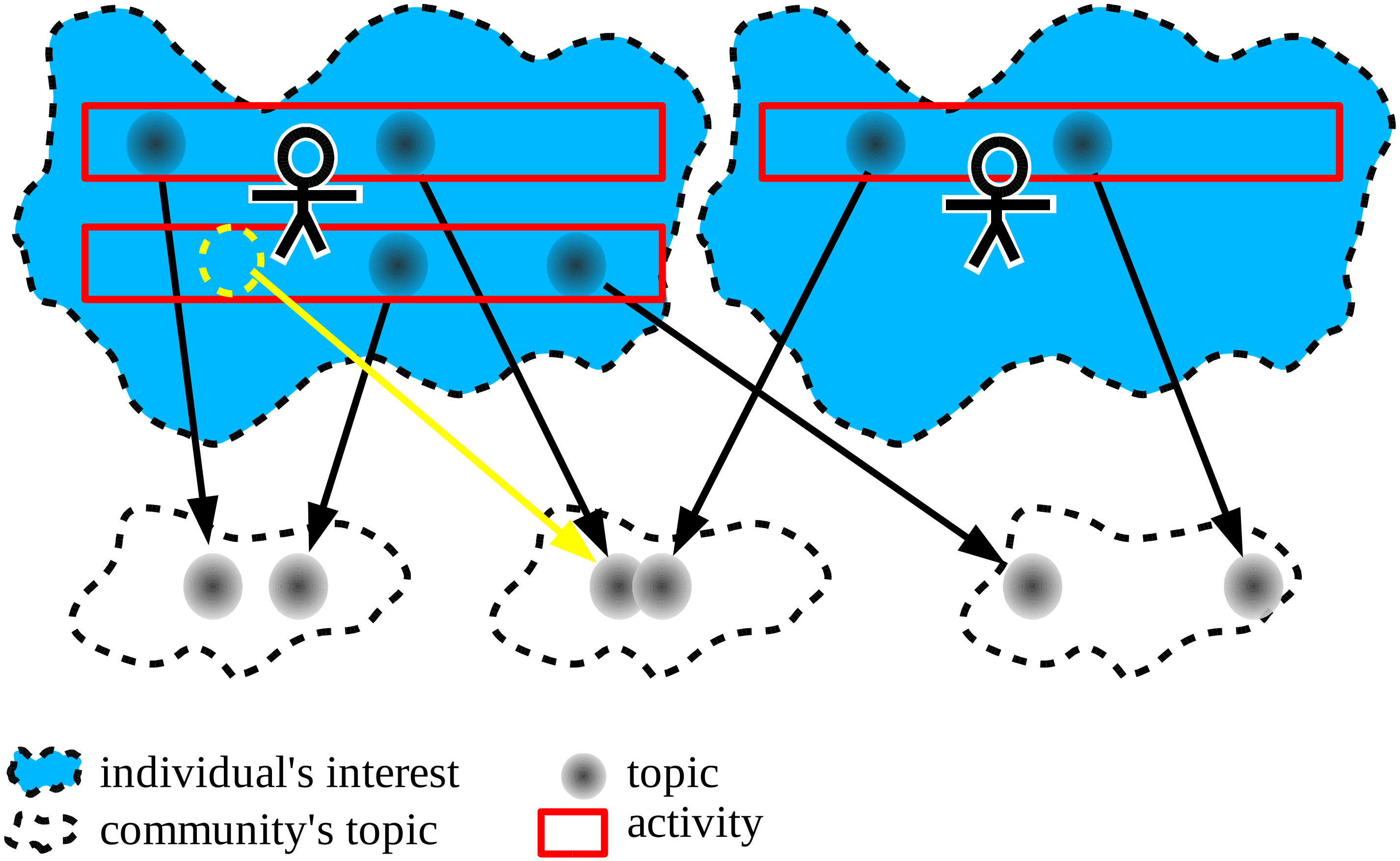
Objective

Complex knowledge artifacts designed and created by human beings such as various organizations, urban plans, buildings, and software systems seldom remain the same as they are initially built because environmental changes and human interactions with them would constantly create the needs for change. Therefore, complex knowledge artifacts must be able to evolve continuously to remain convivial to their users. The objective of this international workshop is to get a deeper understanding of how to evolve knowledge artifacts through the efforts of communities that are formed naturally by the people who have various stakes and interests in the knowledge artifacts. Community-driven evolution of knowledge artifacts entails the crystallization and creation of knowledge through sharing and collaboration among the members. Through presentations by participants on their related research and in-depth discussions, the workshop is expected to advance the understanding of how collaborative knowledge creation takes place in a community during the process of developing and evolving knowledge artifacts and how to design better tools to facilitate such collaboration and evolution.

This workshop is sponsored by the UC Irvine Institute for Software Research (ISR).

Comments and questions: Debra A. Brodbeck, ISR Technical Relations Director, brodbeck@uci.edu

Context delivery



individual's interest

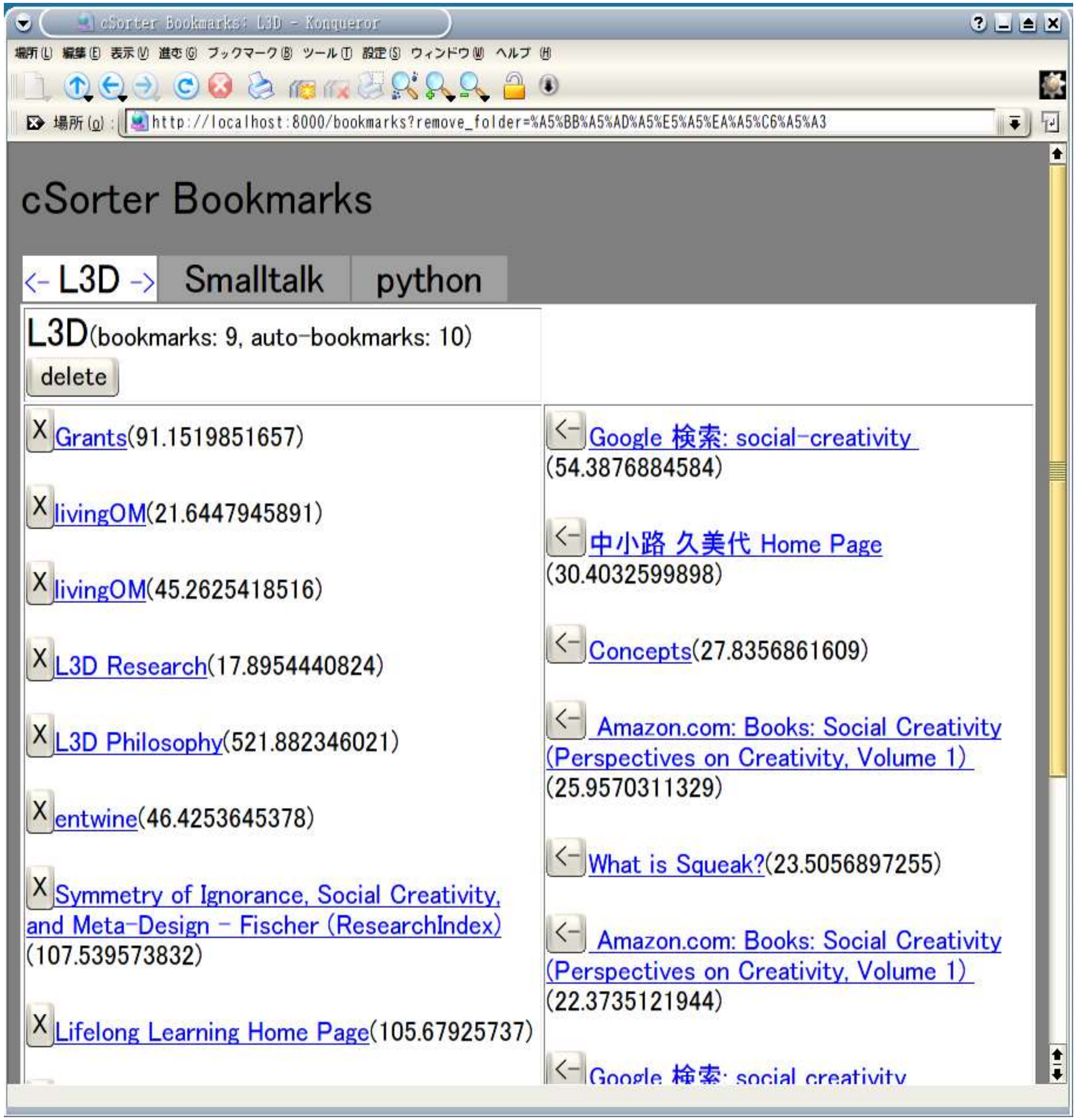
community's topic

● topic

□ activity

cSorter: datamining user's interests

- A user provides "categories", which represents user's interests.
- The user also gives bookmarks in each category, which are sample documents of each interest.
- cSorter recommends documents for each category using Naive Bayes (from URL history).



cSorter Bookmarks: L3D - Konqueror
場所(L) 編集(E) 表示(V) 進む(G) ブックマーク(B) ツール(T) 設定(S) ウィンドウ(W) ヘルプ(H)
場所(o):

http://localhost:8000/bookmarks?remove_folder=%A5%BB%A5%AD%A5%E5%A5%EA%A5%C6%A5%A3

cSorter Bookmarks

<- L3D -> | Smalltalk | python

L3D(bookmarks: 9, auto-bookmarks: 10)

delete

X Grants(91.1519851657)

X livingOM(21.6447945891)

X livingOM(45.2625418516)

X L3D Research(17.8954440824)

X L3D Philosophy(521.882346021)

X entwine(46.4253645378)

X Symmetry of Ignorance, Social Creativity, and Meta-Design - Fischer (ResearchIndex)(107.539573832)

X Lifelong Learning Home Page(105.67925737)

<- Google 検索: social-creativity (54.3876884584)

<- 中小路 久美代 Home Page (30.4032599898)

<- Concepts(27.8356861609)

<- Amazon.com: Books: Social Creativity (Perspectives on Creativity, Volume 1) (25.9570311329)

<- What is Squeak?(23.5056897255)

<- Amazon.com: Books: Social Creativity (Perspectives on Creativity, Volume 1) (22.3735121944)

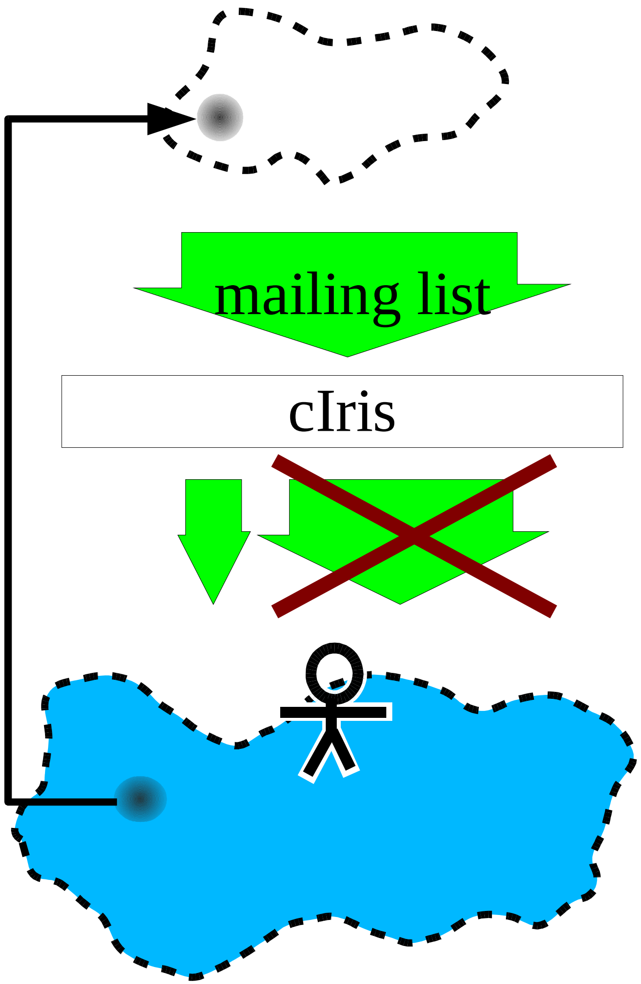
<- Google 検索: social creativity

Interests are dynamic

- The system should catch up updates of user's interests.
  - A user may get interested in a new topic.
  - A user may expand the range of a topic.
  - A user may retract a topic of interest.
  - A user may have different interests on a same document.
  - and so on...

cIris: Information filter at end points

- Many communities provide tons of information via mailing lists.
- Many participates have only partial intests in the community's topics.
- cIris filters documents using the stochastic model developed by cSorter.
- cIris uses distribution of keywords as a user model.
  (similar to distribution of functionality)



Sender's benefit on receiver's filter

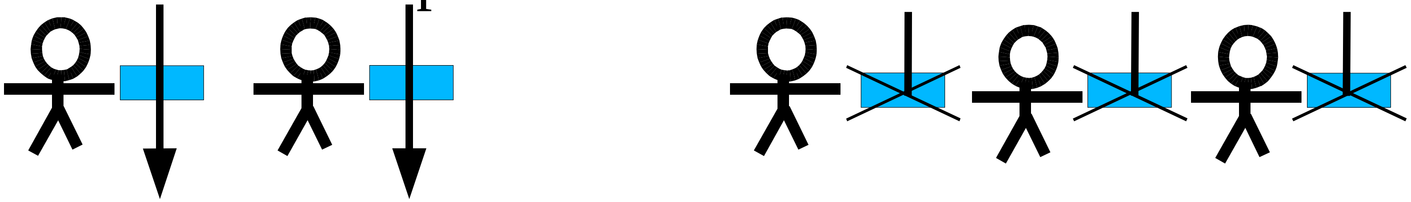
- Suppose that you are sending a message to a mailing list...
  - A sender don't know receivers' interests.
    - You may hesitate to broadcast the message which many receiver can respond to.
    - Or, you may bother people by broadcasting the message which no reciever really care.
  - Using cIris, senders don't have to worry about receivers' interests.

Difficulties revisited

- Identifying task
  - cSorter can classify recent N documents to identify the topic of the current task.
  - cScope can help users to identify potential topic of the current task.
- Recommending artifacts/people
  - cSorter
  - cIris: see the next slide.

cSuite for DynC

- Possible ways to extend cSuite for “dynamic community"
  - Use cIris to screen persons



    - public cIris: Send remote query to cIris of your friends.
      - privacy issue ... cIris has a lot of private information!
    - P2P cIris: Flood the message into a P2P-like network and filter at each node using cIris.

Conclusions

- Cross-community activities need support over multiple communities.
- cSuite is a support tool for cross-community activities focused on individuals:
  - Context delivery suggests potential support of / potential contribution to a community.
  - Document categorization catches up changes of interests.
  - Information filtering at receiver's end.
- Possible extention for DynC