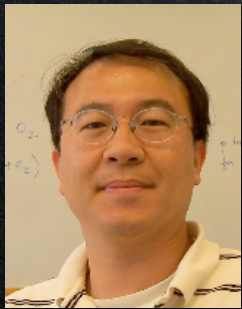


6S: Collaborative Web Search via an Adaptive Peer Network



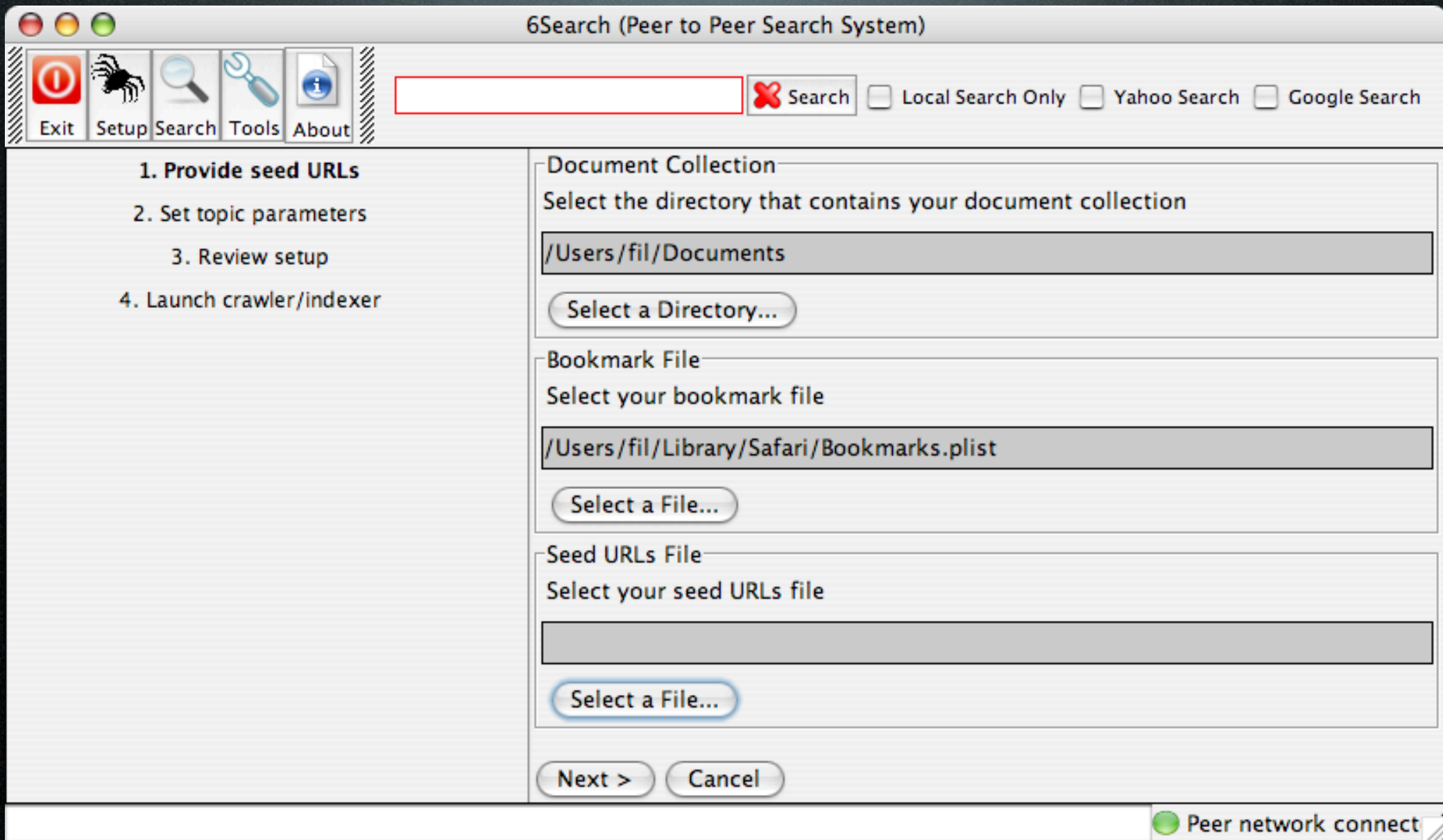
Filippo Menczer,
Le-Shin Wu, Ruj Akavipat



Informatics & Computer Science
Indiana University



<http://homer.informatics.indiana.edu/6S/>



<http://homer.informatics.indiana.edu/6S/>

6Search (Peer to Peer Search System)

Exit Setup Search Tools About

Local Search Only Yahoo Search Google Search

1. Provide seed URLs
2. Set topic parameters
3. Review setup
4. Launch crawler/indexer

Crawling Progress
11/09 14:53:34

Indexing Progress
11/09 14:56:28

Information Summary

Please verify the information you entered.

Document Collection:
/Users/fil/Documents/Homepages/Informatics

Bookmark File:
/Users/fil/Library/Safari/Bookmarks.plist

Crawling Topic:
web mining modeling search peer crawl networks agents

Seed URLs File:

Number of Pages:
100

Your Email:
fil@indiana.edu

Tag for URLs:
web mining modeling search peer crawl networks agents

Peer network connect

<http://homer.informatics.indiana.edu/6S/>

6Search (Peer to Peer Search System)

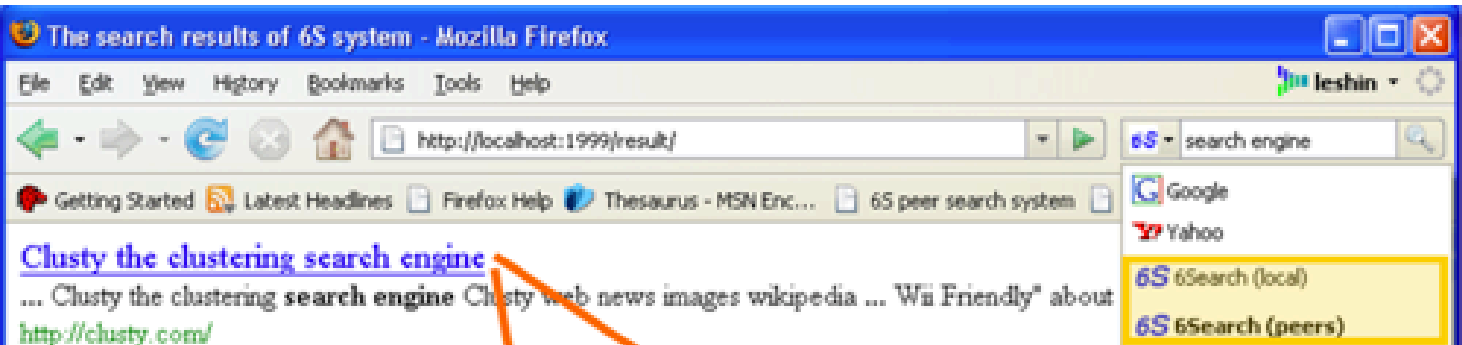
Exit Setup Search Tools About

Search Local Search Only Yahoo Search Google Search

Indexing Date	Document Title	Document Url	Document Tag
2006/11/09 ...	ClickZ Internet Marketing Solutions for Marketers	http://www.clickz.com/	web mining modeling...
2006/11/09 ...		http://www.di.fc.ul.pt/sobre/bibtex.php?entr...	web mining modeling...
2006/11/09 ...		http://wwwconf.ecs.soton.ac.uk/perl/oai2?ve...	web mining modeling...
2006/11/09 ...	Geeking with Greg	http://glinden.blogspot.com/	web mining modeling...
2006/11/09 ...	Crawling the Web Speakers	http://www.round.table.com/speakersbureau...	web mining modeling...
2006/11/09 ...	Things an Observatory should have (institutiona...	http://www.sdsc.edu/NEON/jan2000/broadc...	web mining modeling...
2006/11/09 ...	Luminous Landscape Forum (Powered by Invisio...	http://www.luminous-landscape.com/forum/	web mining modeling...
2006/11/09 ...	VDS Technologies - Map Data	http://www.vdstech.com/map_data.htm	web mining modeling...
2006/11/09 ...	Mark Pincus Blog	http://markpincus.typepad.com/	web mining modeling...
2006/11/09 ...	W3C I18N Tutorial: CSS3 and International Text	http://www.w3.org/International/tutorials/css...	web mining modeling...
2006/11/09 ...	OnoTech	http://onotech.blogspot.com/	web mining modeling...
2006/11/09 ...	Search	http://www.roundtablegroup.com/about/sear...	web mining modeling...
2006/11/09 ...	First Tuesdays at the Urban Institute 03/2002	http://www.urban.org/Presentations/FT/FT3-...	web mining modeling...
2006/11/09 ...	Clusty Cloud Creator	http://cloud.clusty.com/	web mining modeling...
2006/11/09 ...	ピッツバーグ便り	http://mawaha.exblog.jp/	web mining modeling...
2006/11/09 ...	Powerset: See what people are saying right now...	http://www.technorati.com/tag/Powerset/	web mining modeling...
2006/11/09 ...	AI Resources and Links	http://labs.rightnow.com/resources.php	web mining modeling...
2006/11/09 ...	The Next Net: The Search for the Next Big Leap...	http://business2.blogs.com/business2blog/2...	web mining modeling...
2006/11/09 ...	slides.dvi	http://www.comp.hkbu.edu.hk/WI05/speake...	web mining modeling...
2006/11/09 ...	Amazon.com: The Search: How Google and Its ...	http://www.amazon.com/Search-Rewrote-Bu...	web mining modeling...
2006/11/09 ...	Industry Solutions from RightNow Technologies ...	http://www.rightnow.com/products/industry-...	web mining modeling...
2006/11/09 ...	ODP - Open Directory Project	http://www.dmoz.org/	web directory

Advanced Tool

Peer network connect



[Clusty the clustering search engine](#)

... Clusty the clustering search engine Clusty web news images wikipedia ... 'Wii Friendly' about

<http://clusty.com/>

Contributors: bebe#4,beast#7,El@homer#1

[Similar pages \(Powered by GiveALink\)](#)

[Ask.com Search Engine - Better Web Search](#)

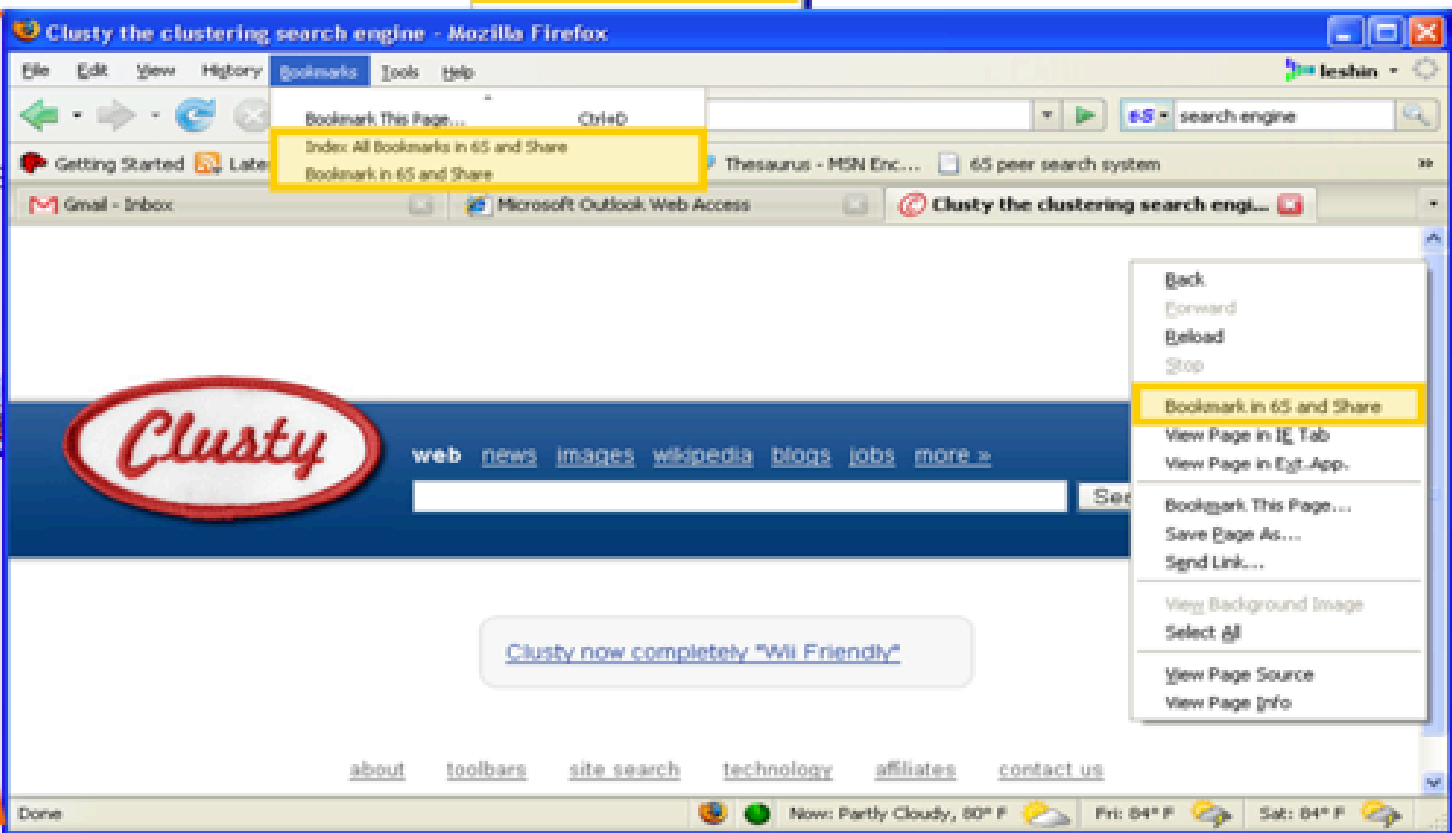
... Ask.com Search Engine - Better Web Search Mys

& Media Search ...

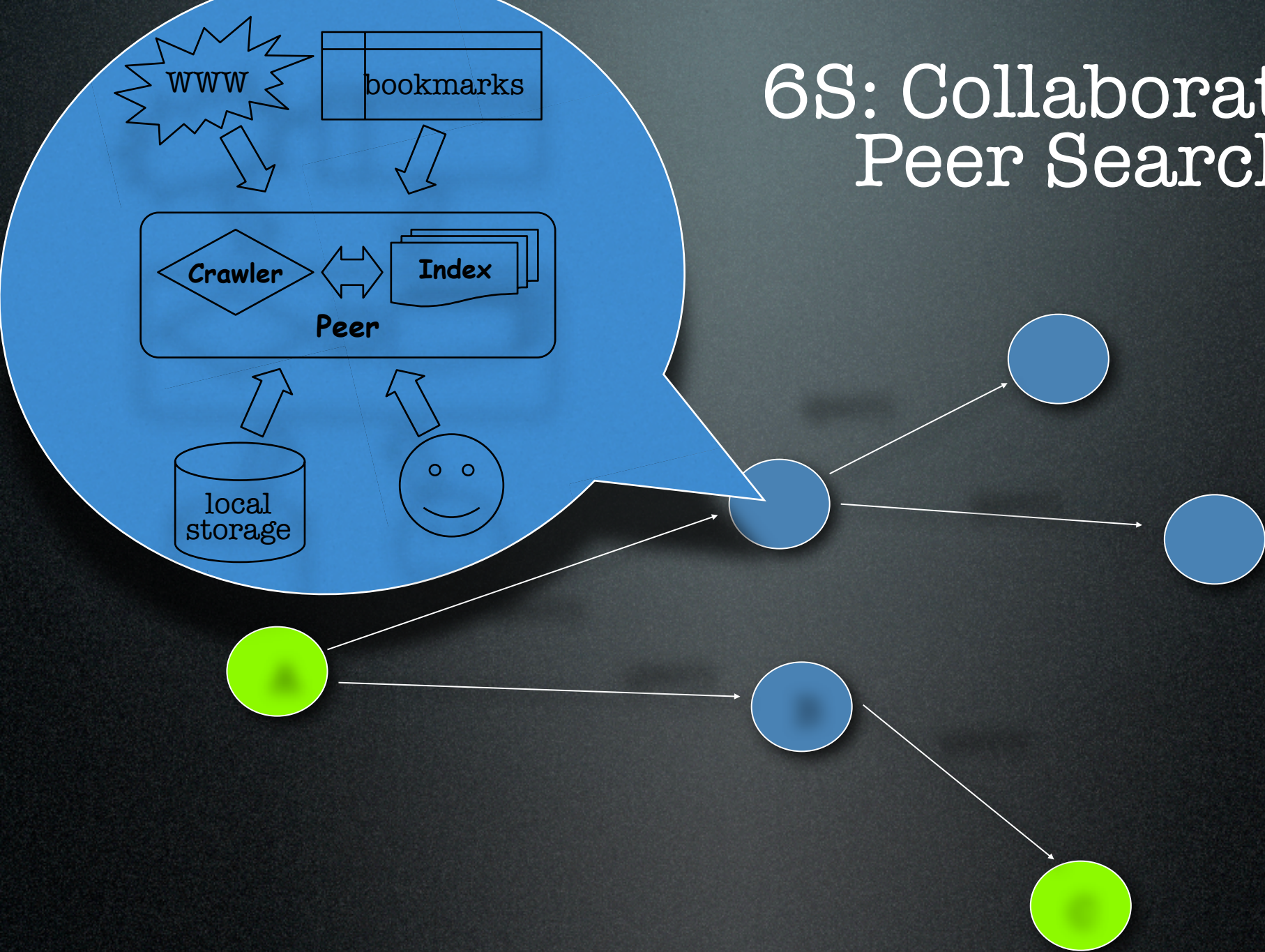
<http://www.ask.com/>

Contributors: My,bebe#4,beast#7

[Similar pages \(Powered by GiveALink\)](#)

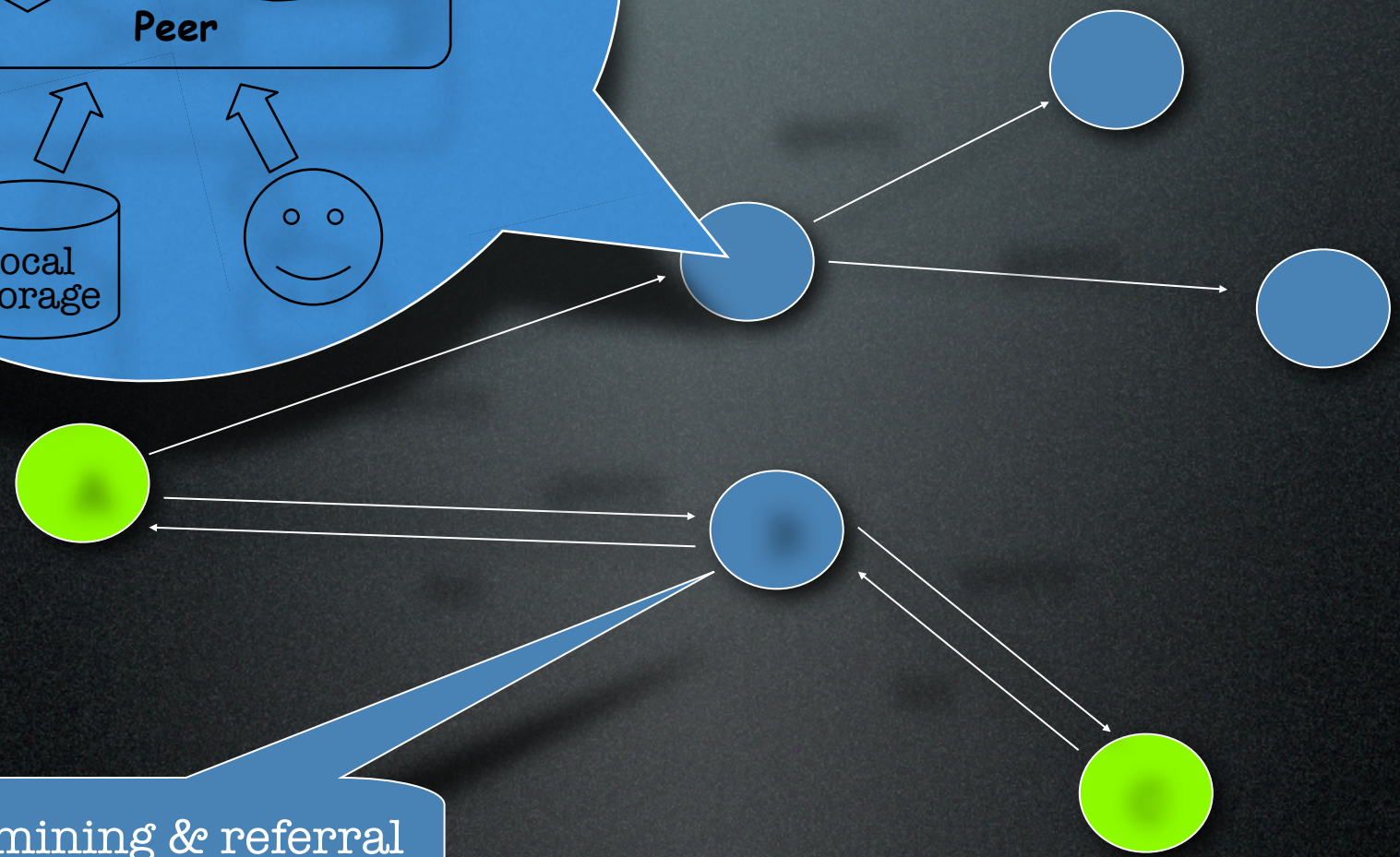
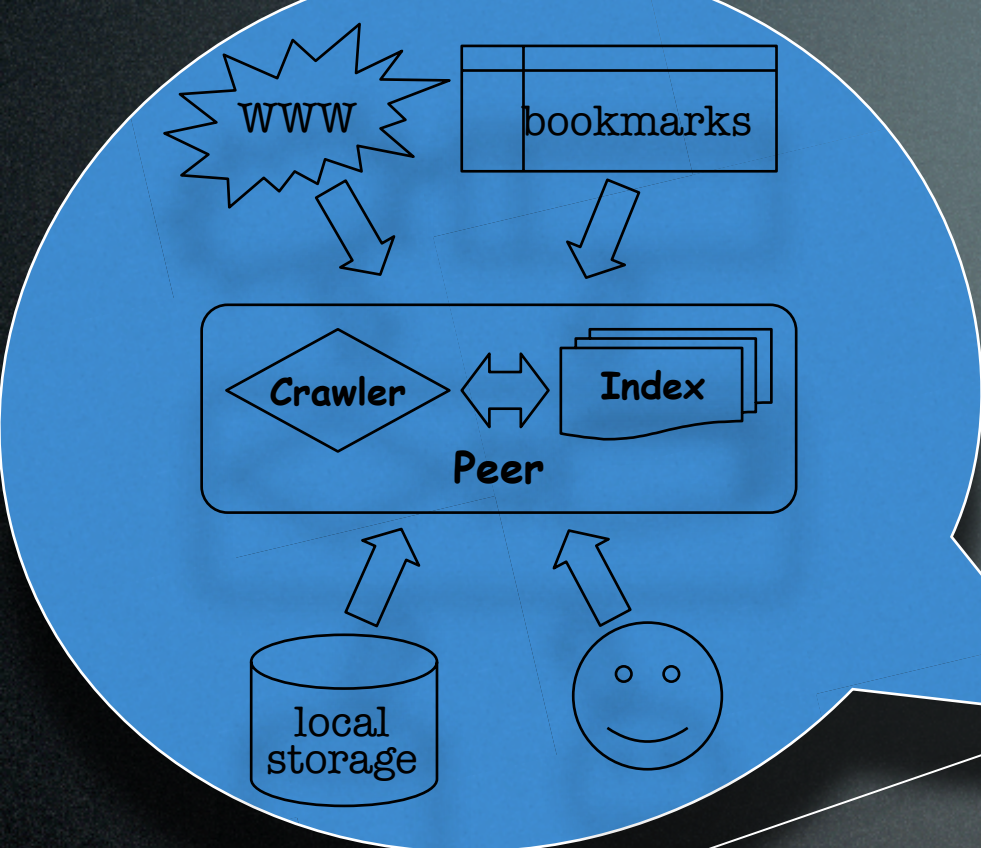


6S: Collaborative Peer Search



WWW2004
WWW2005
WTAS2005
P2PIR2006

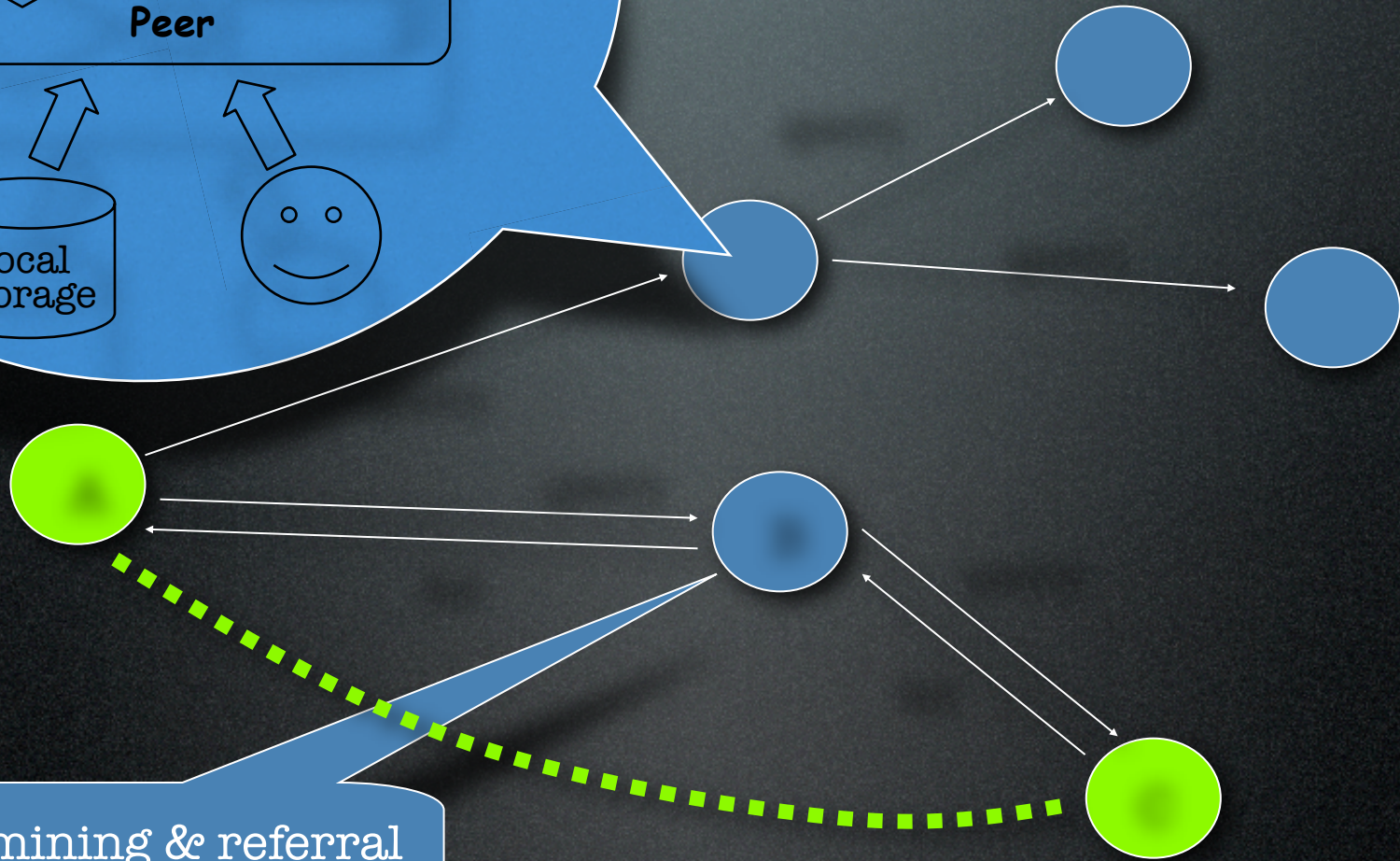
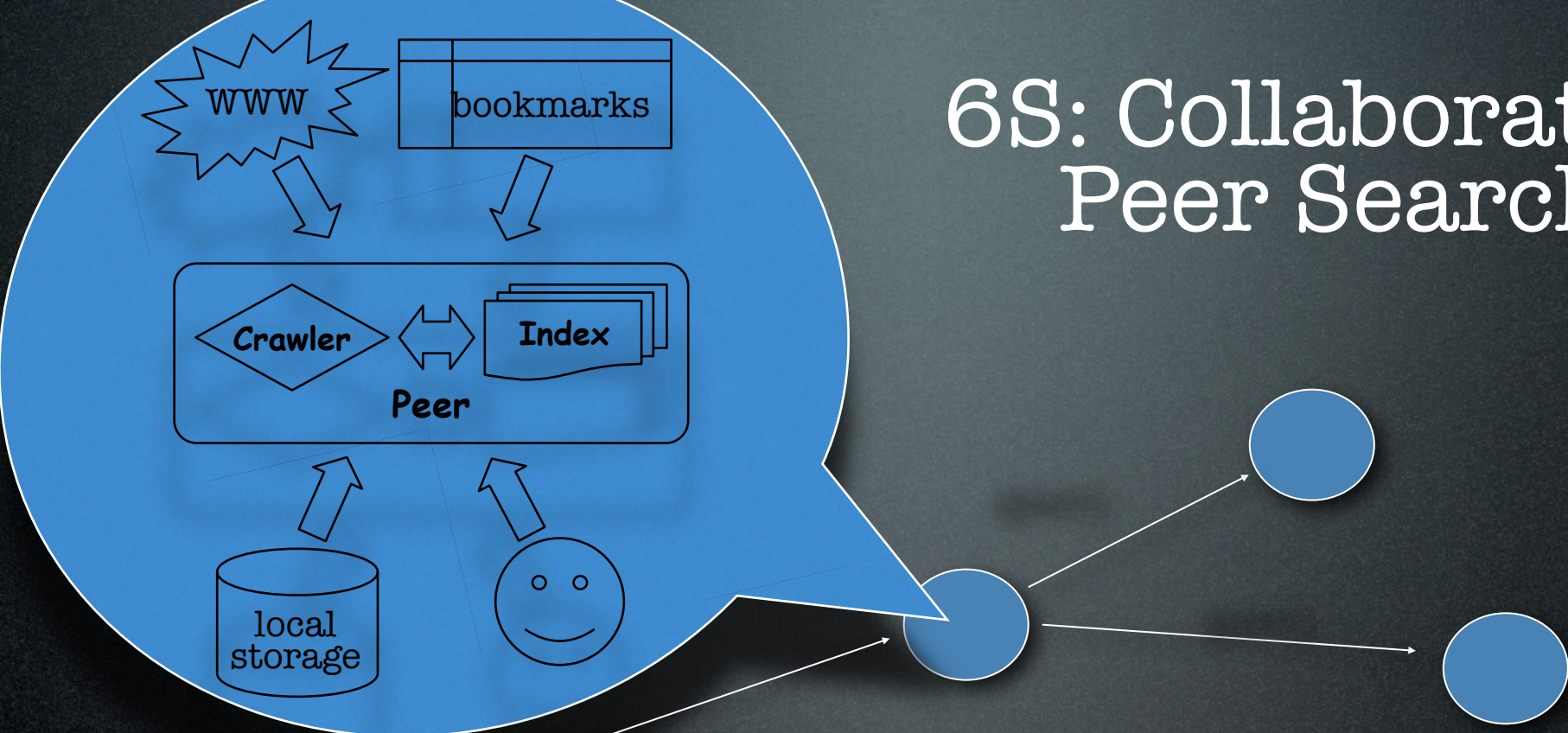
6S: Collaborative Peer Search



Data mining & referral opportunities

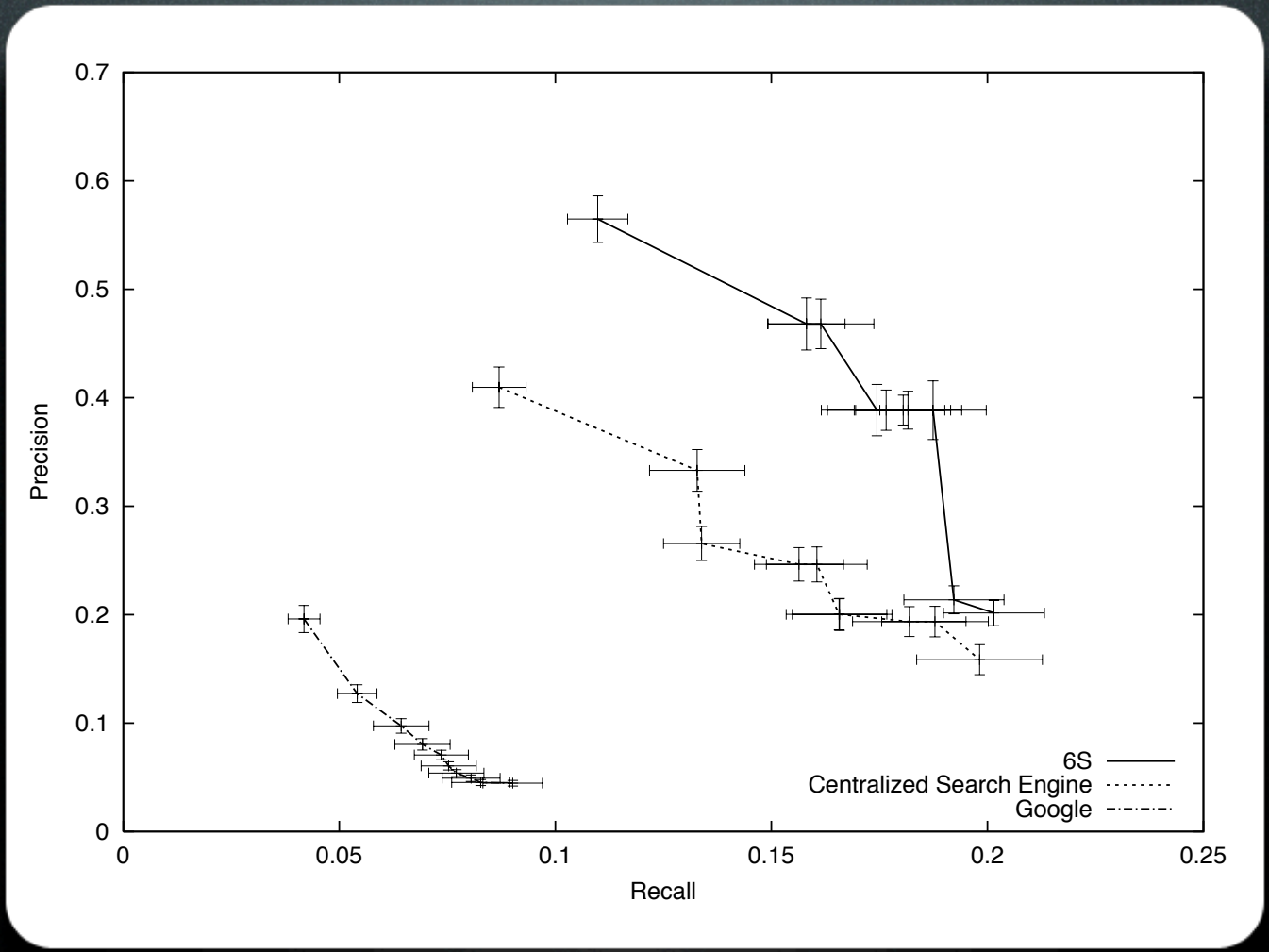
WWW2004
WWW2005
WTAS2005
P2PIR2006

6S: Collaborative Peer Search



Data mining & referral opportunities

WWW2004
WWW2005
WTAS2005
P2PIR2006

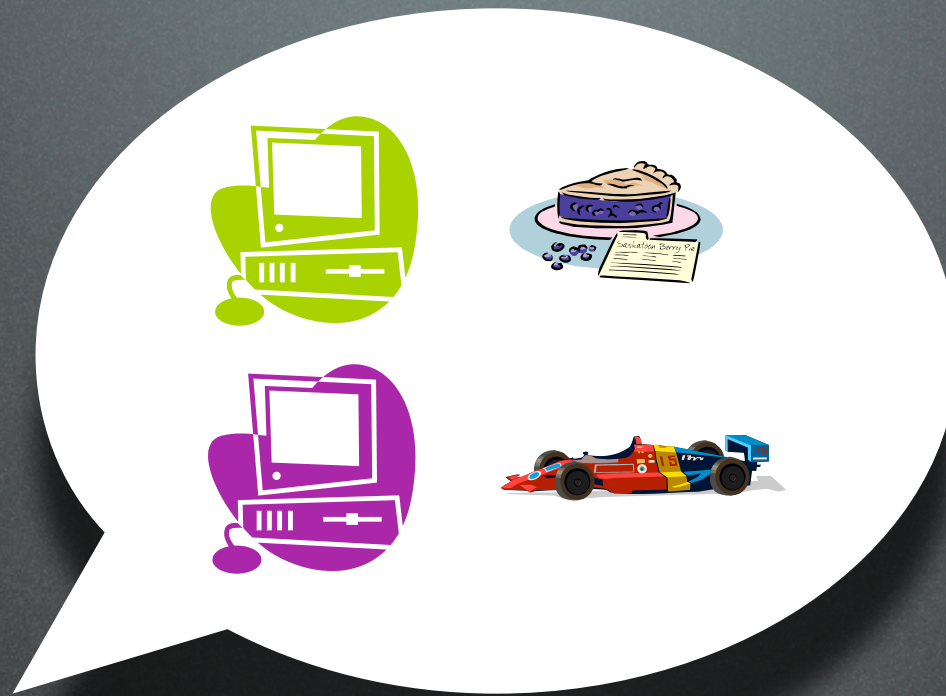


Distributed vs Centralized

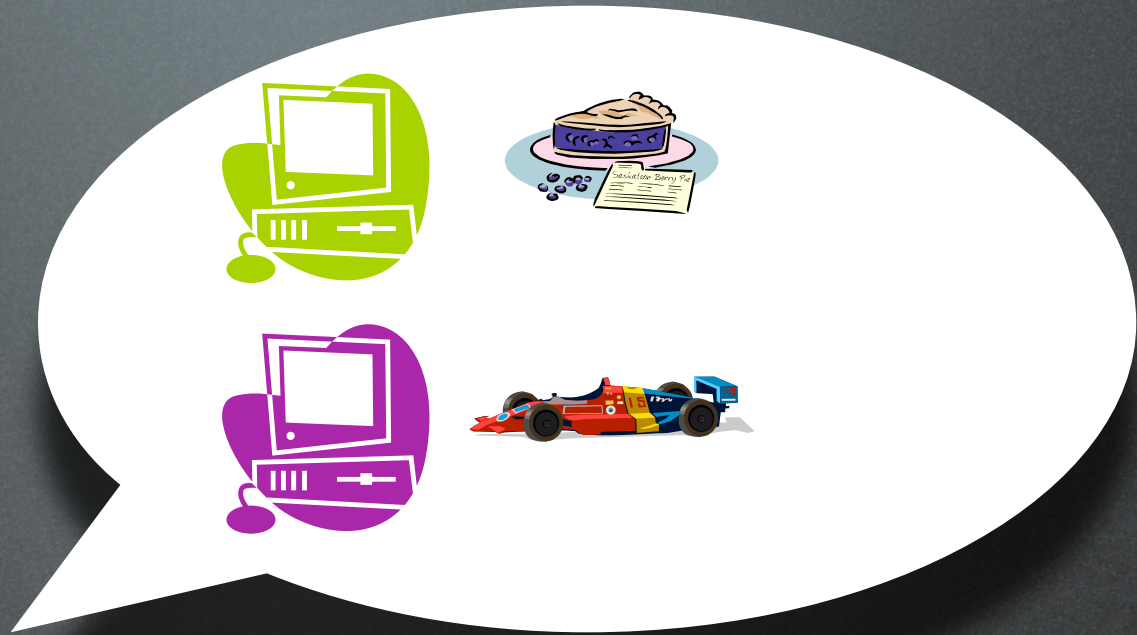
Algorithm 1: Random Known



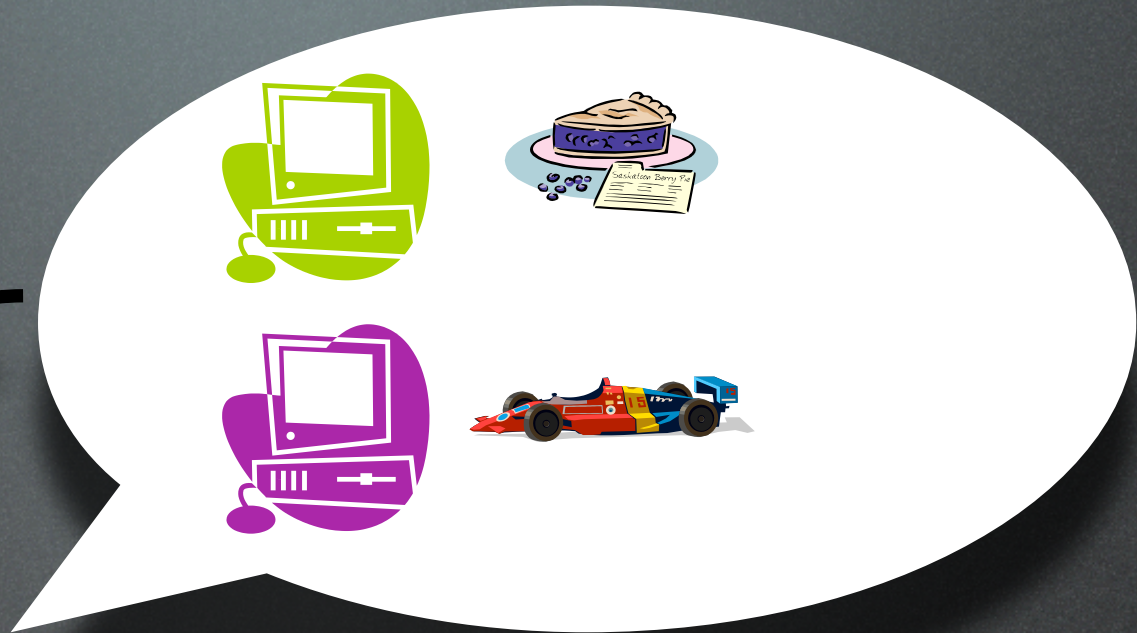
Algorithm 2: Greedy



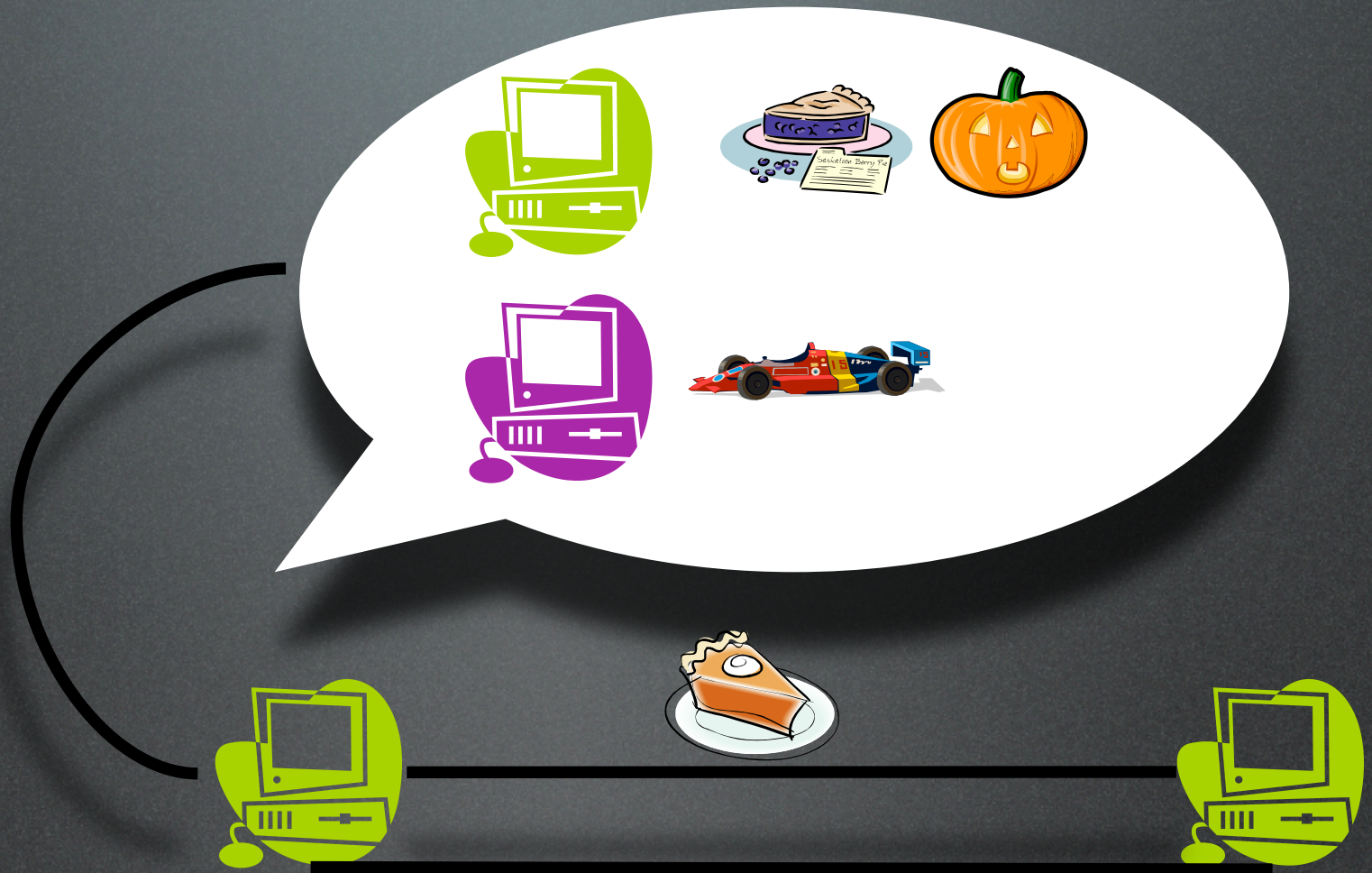
Algorithm 3: Reinforcement



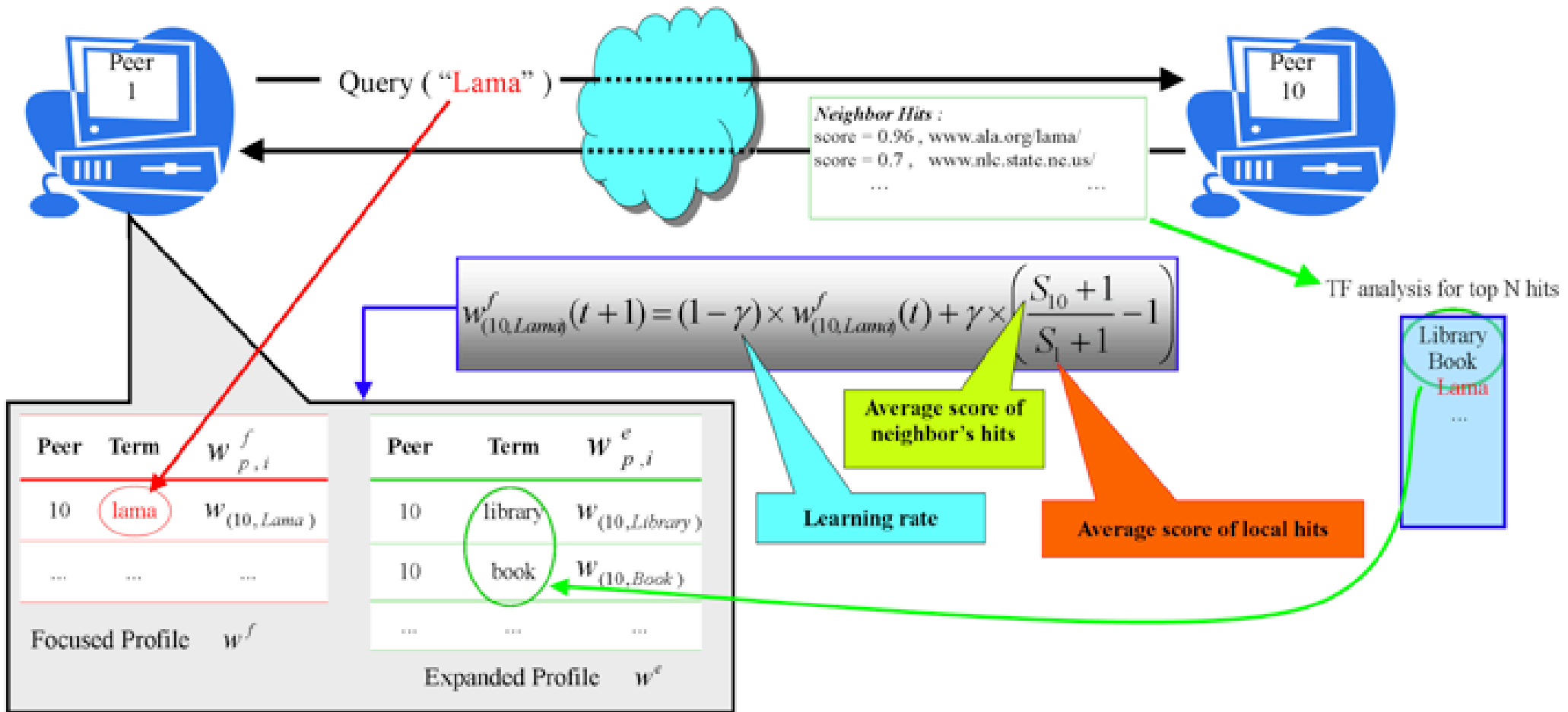
Algorithm 3: Reinforcement



Algorithm 3: Reinforcement



Algorithm 3: Reinforcement



Query Routing



Query ("Lama library book")

$$\sigma(p, Query) = \sum_{i \in Query} [\alpha \cdot w_{p,i}^f + (1 - \alpha) \cdot w_{p,i}^e]$$

N top ranked among known peers are selected as neighbors and sent Query.

Forwarding the query



Peer	Term	$w_{p,i}^f$
10	lama	.98
...

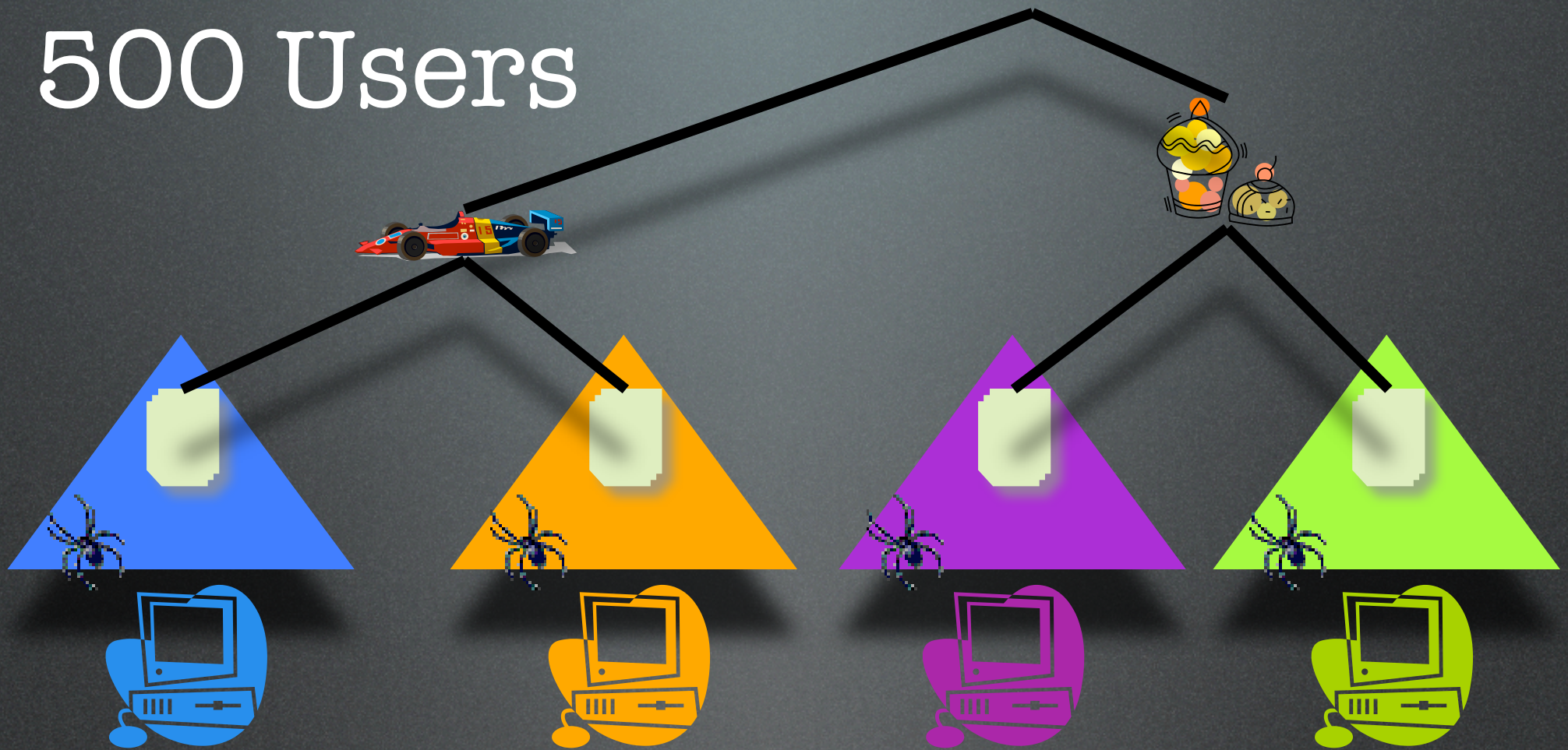
Focused Profile w^f

Peer	Term	$w_{p,i}^e$
10	library	.87
10	book	.67
...

Expanded Profile w^e

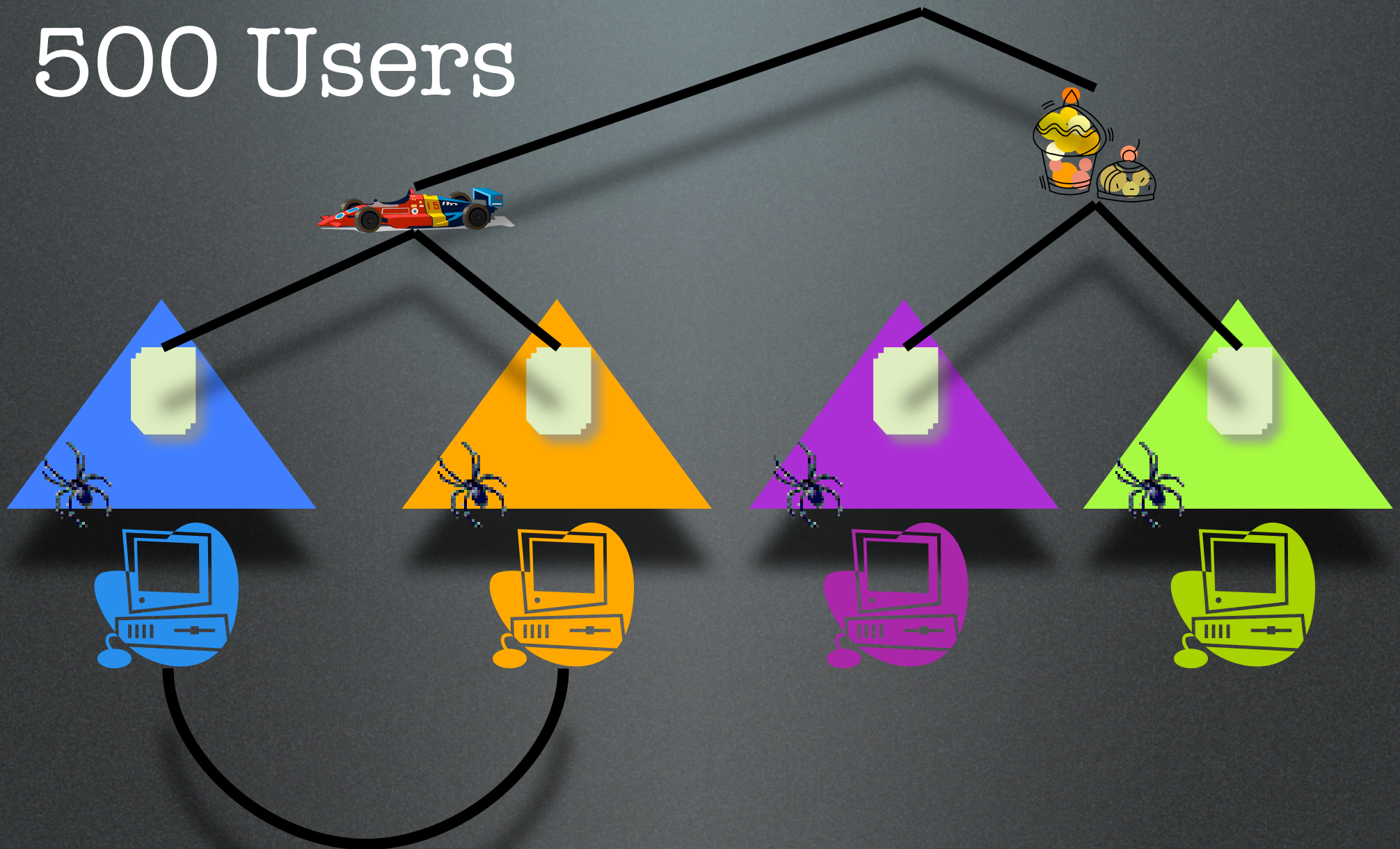
Simulating 500 Users

ODP (dmoz.org)



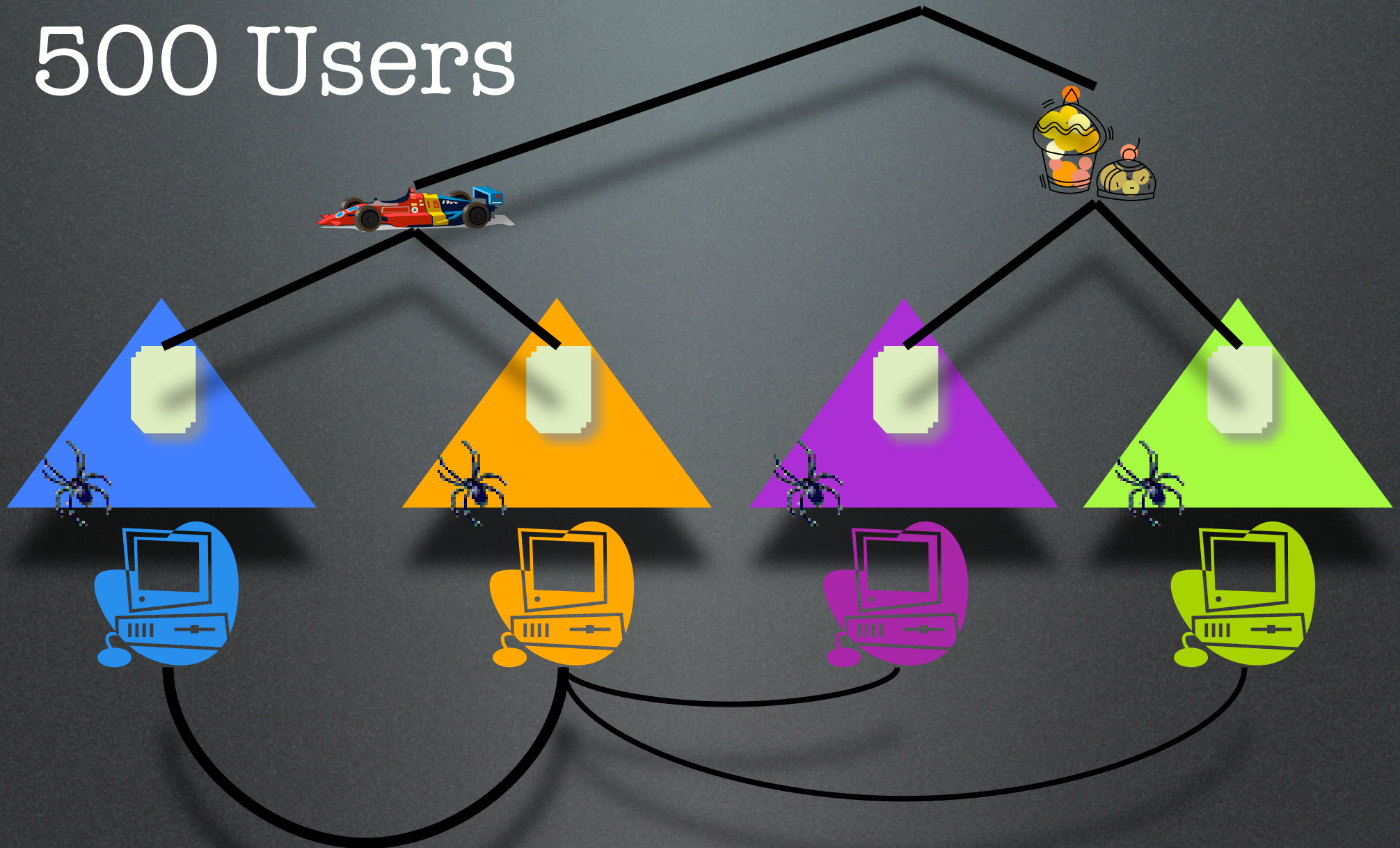
Simulating 500 Users

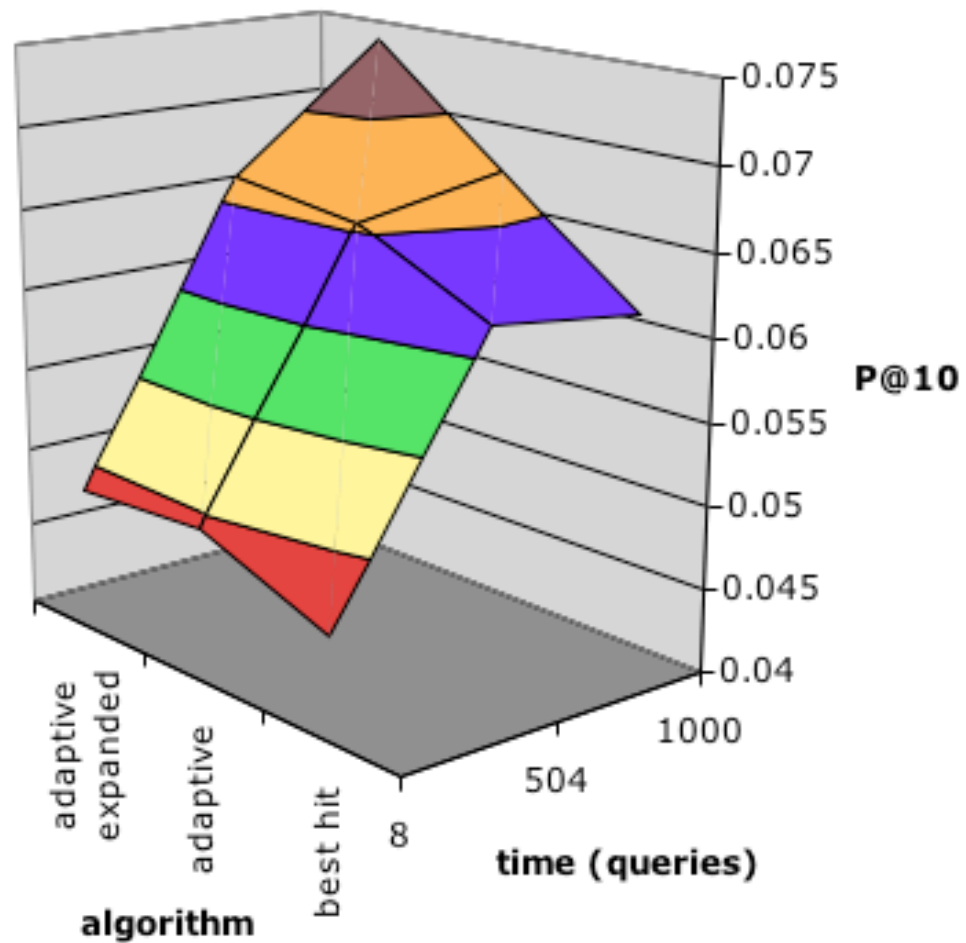
ODP (dmoz.org)



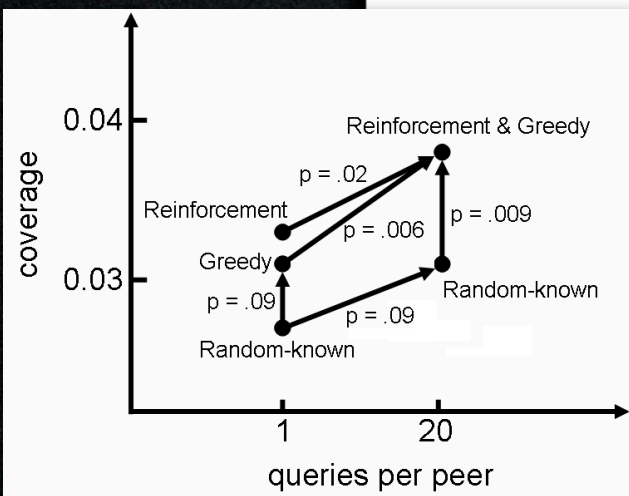
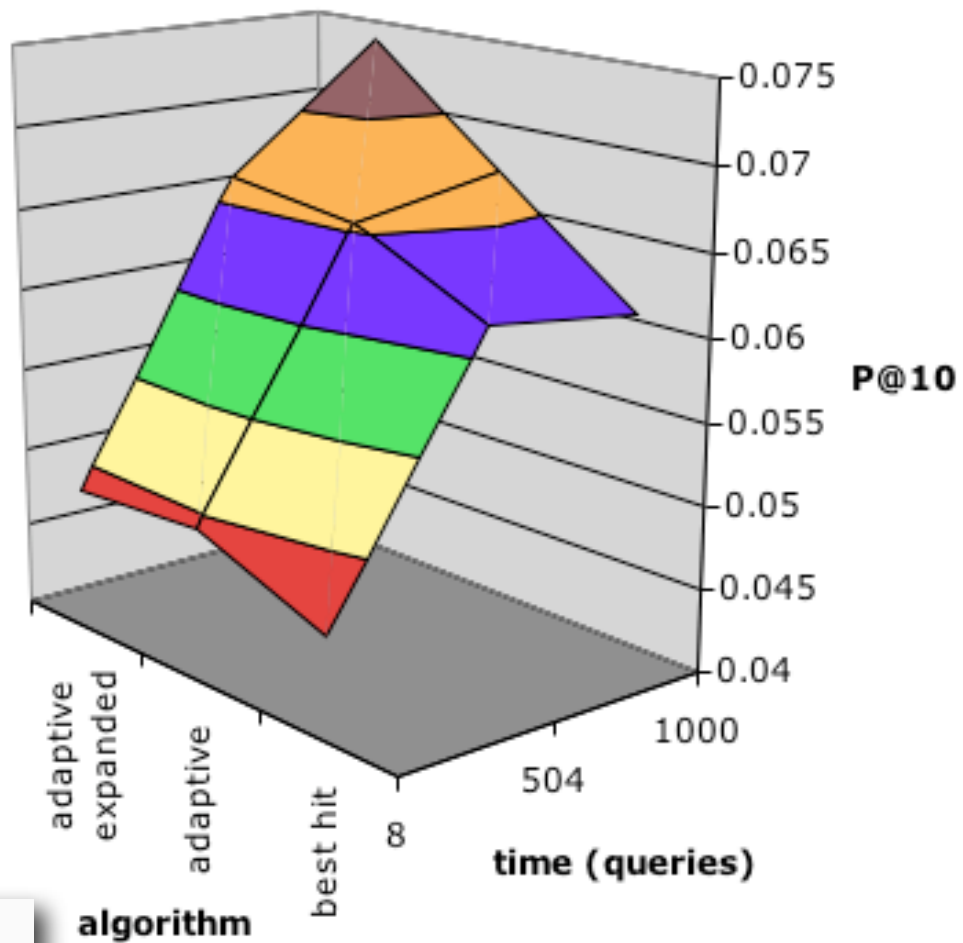
Simulating 500 Users

ODP (dmoz.org)

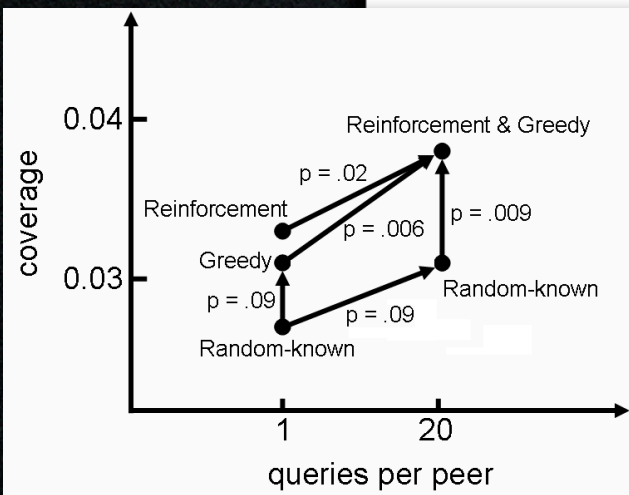
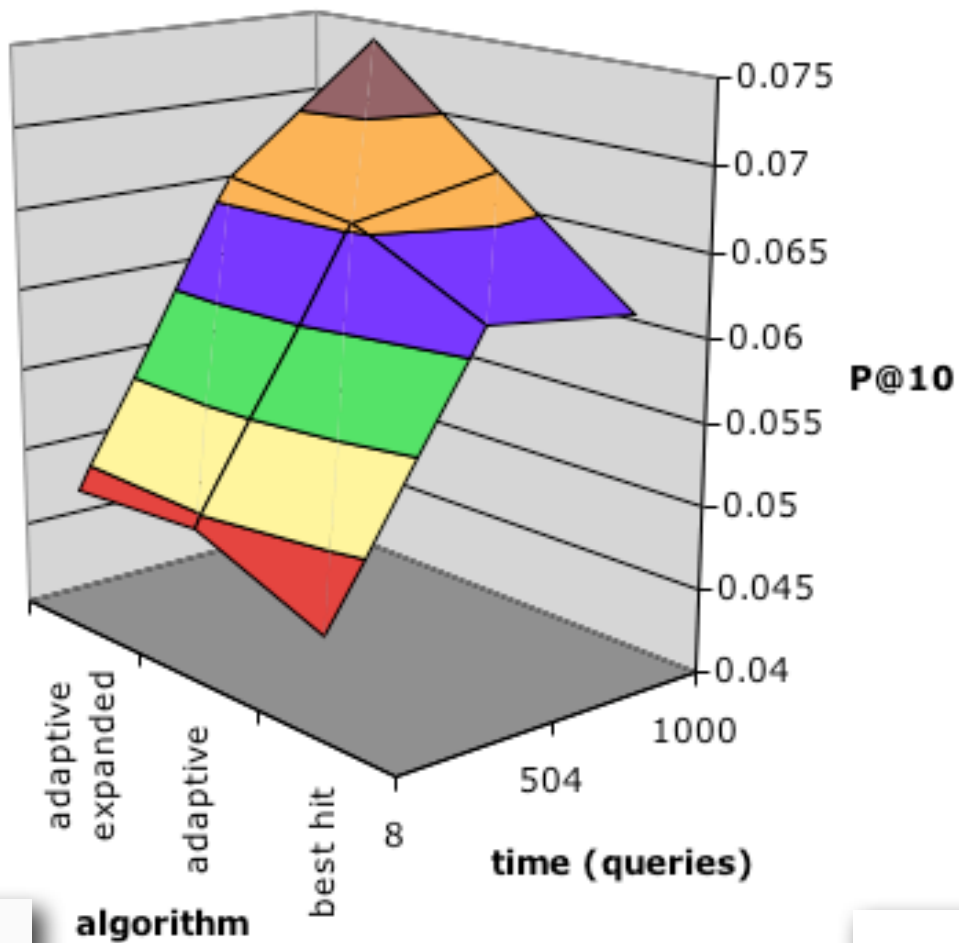




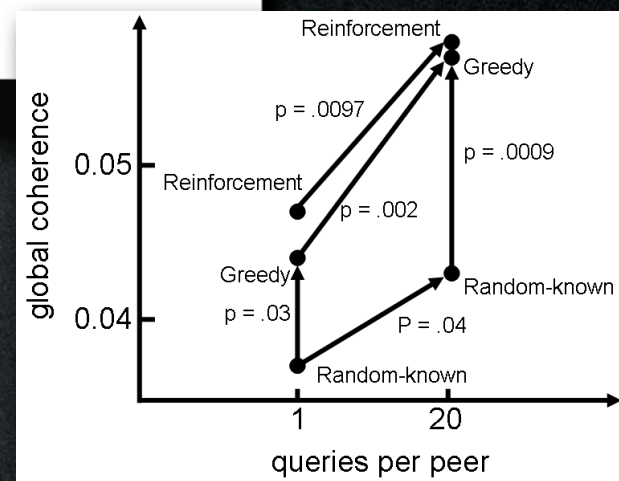
P@10

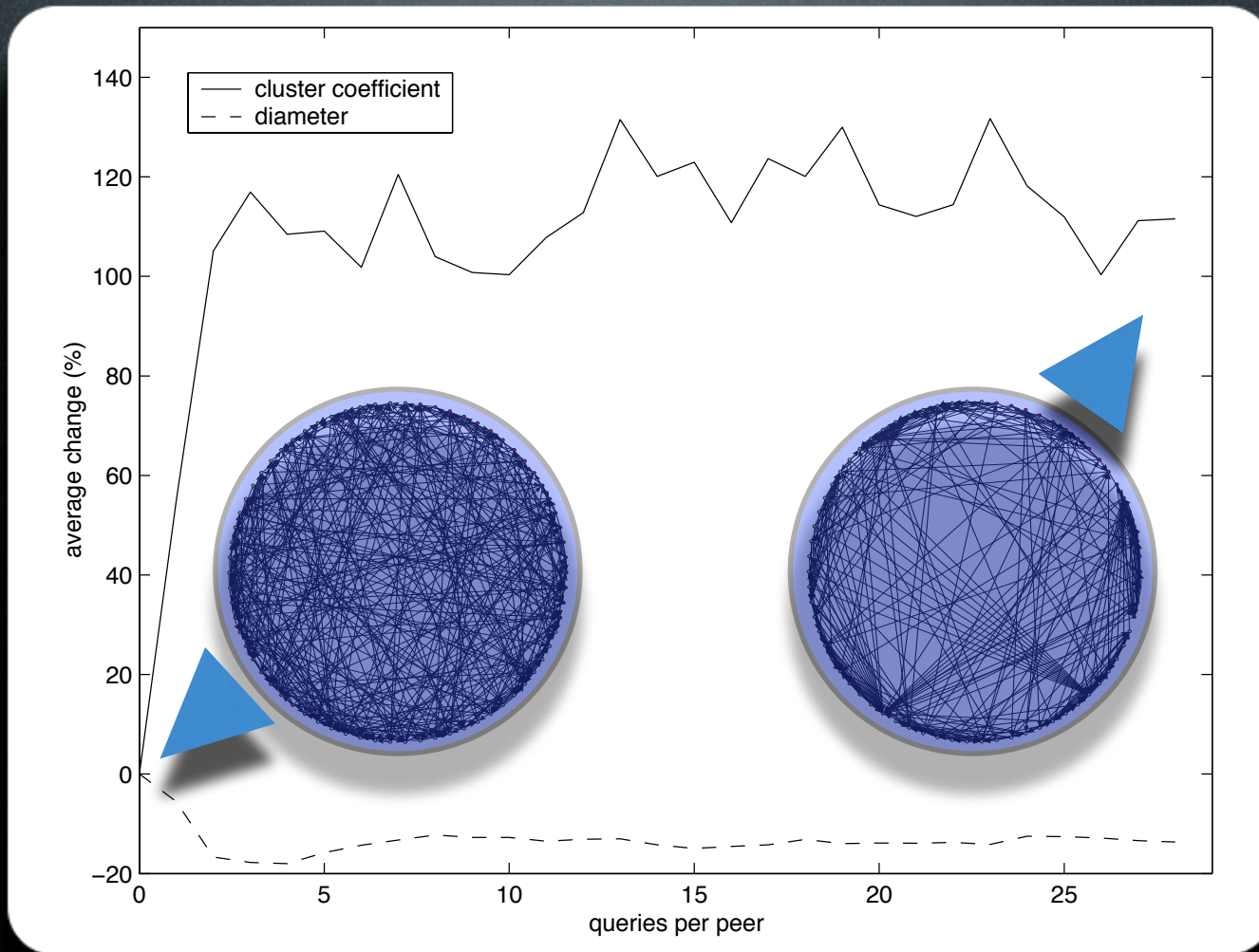


P@10



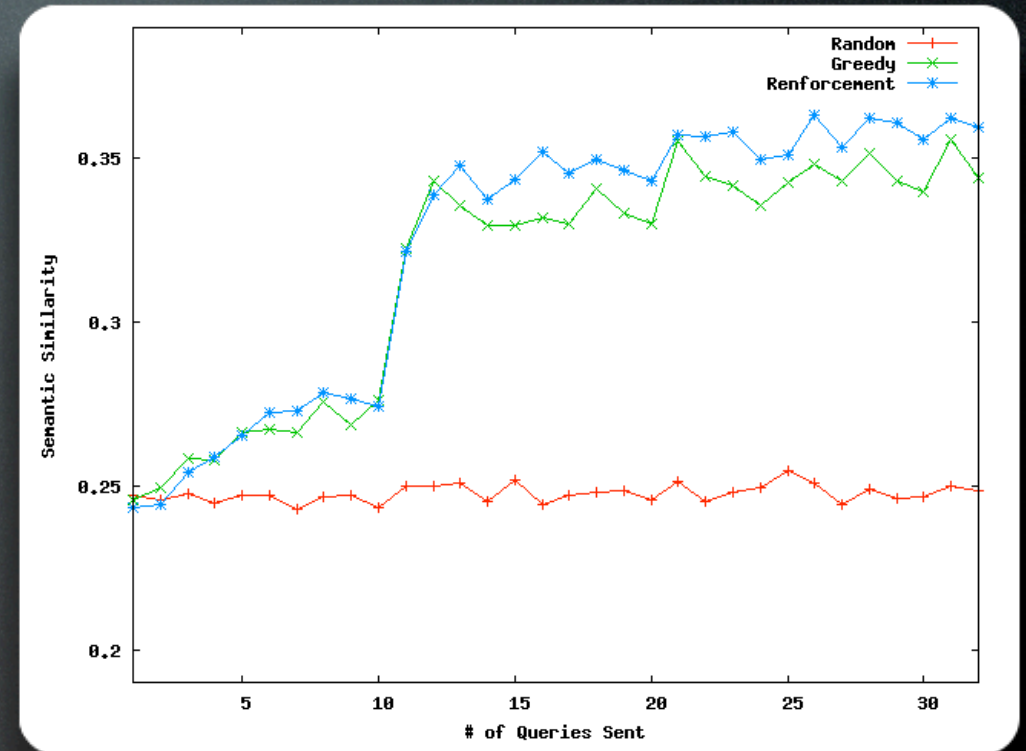
P@10



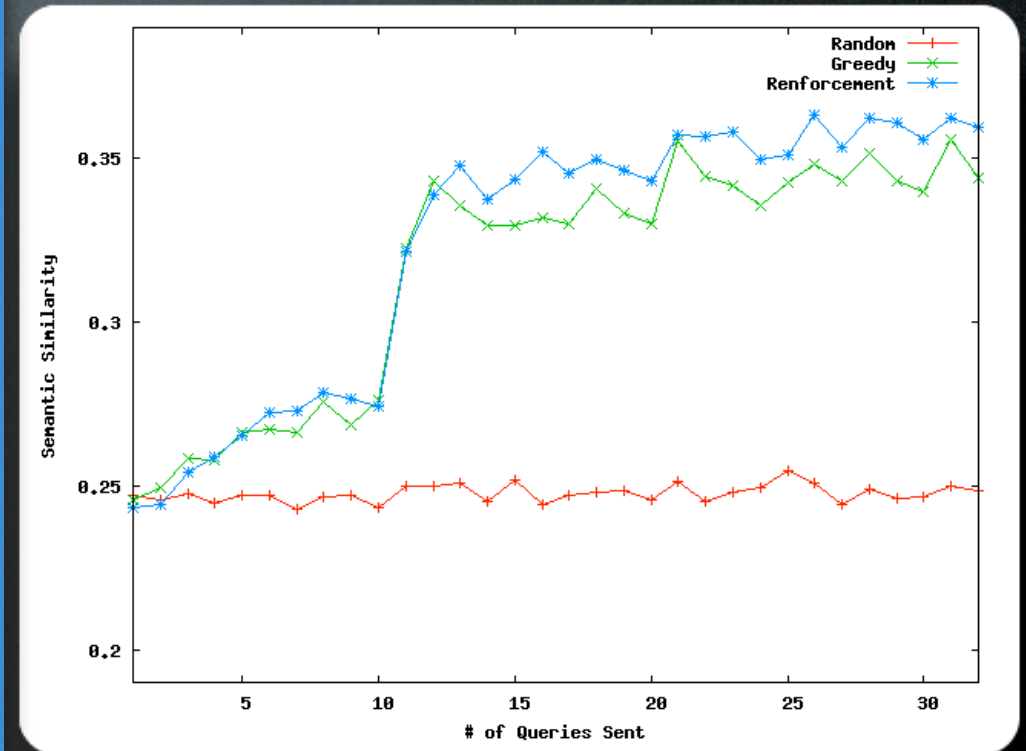


Small-world

Semantic Similarity

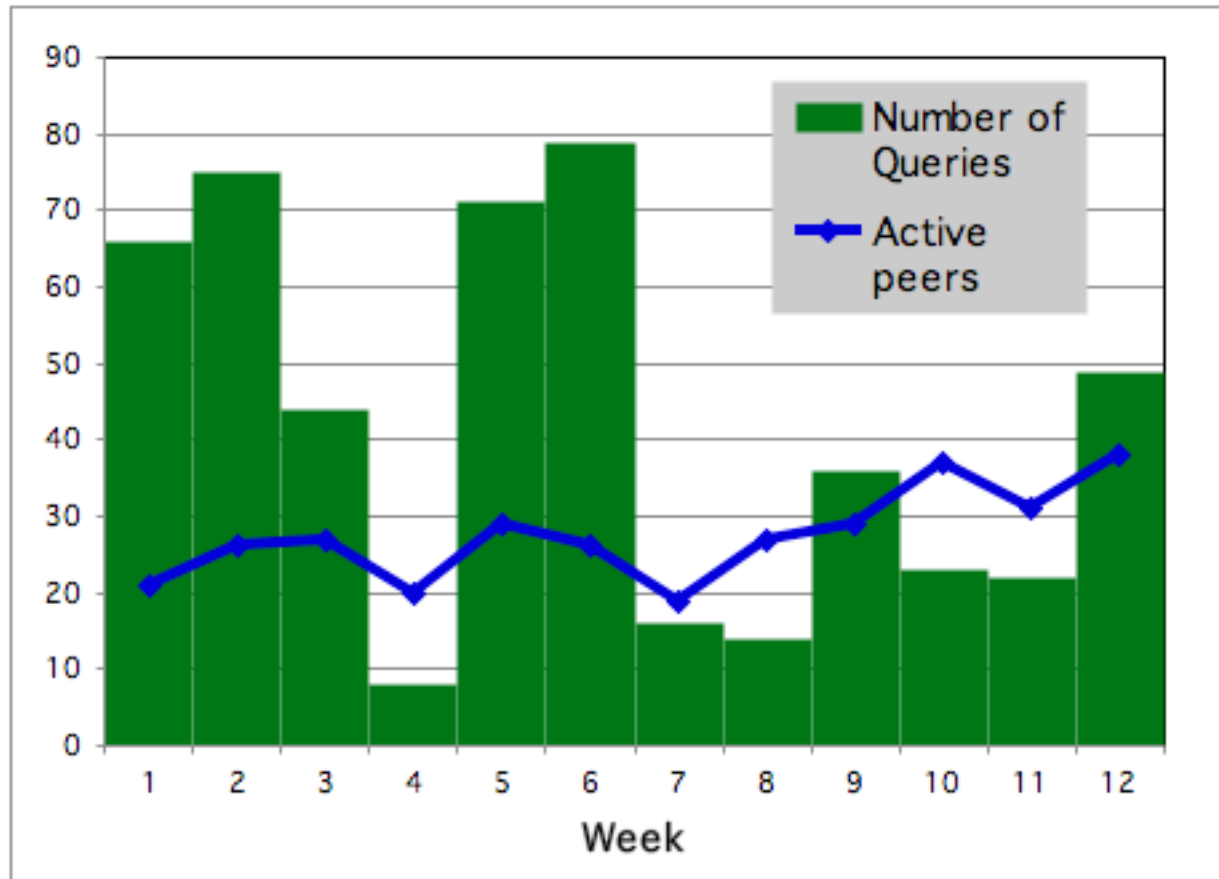


Semantic Similarity

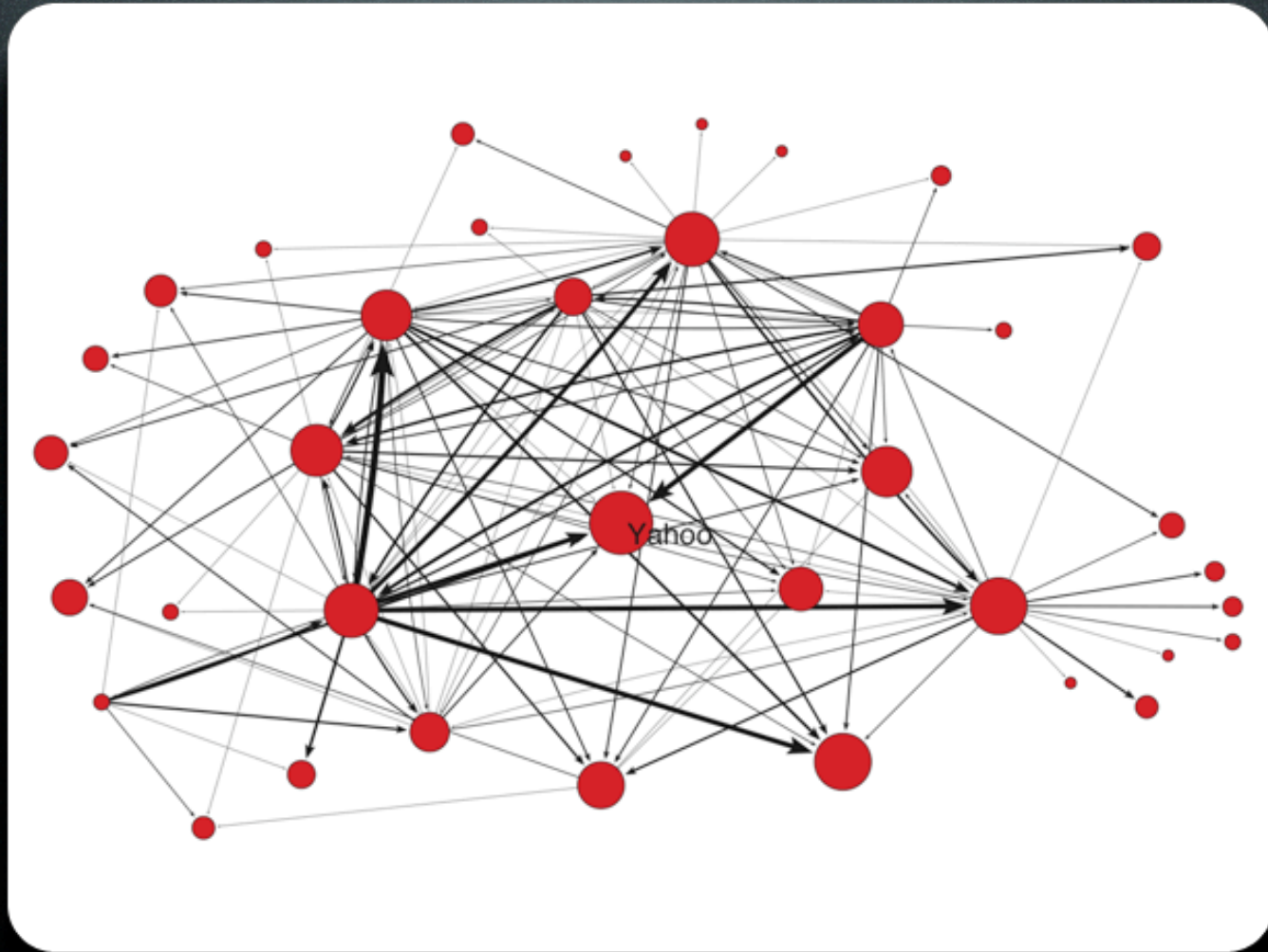


Ongoing Work

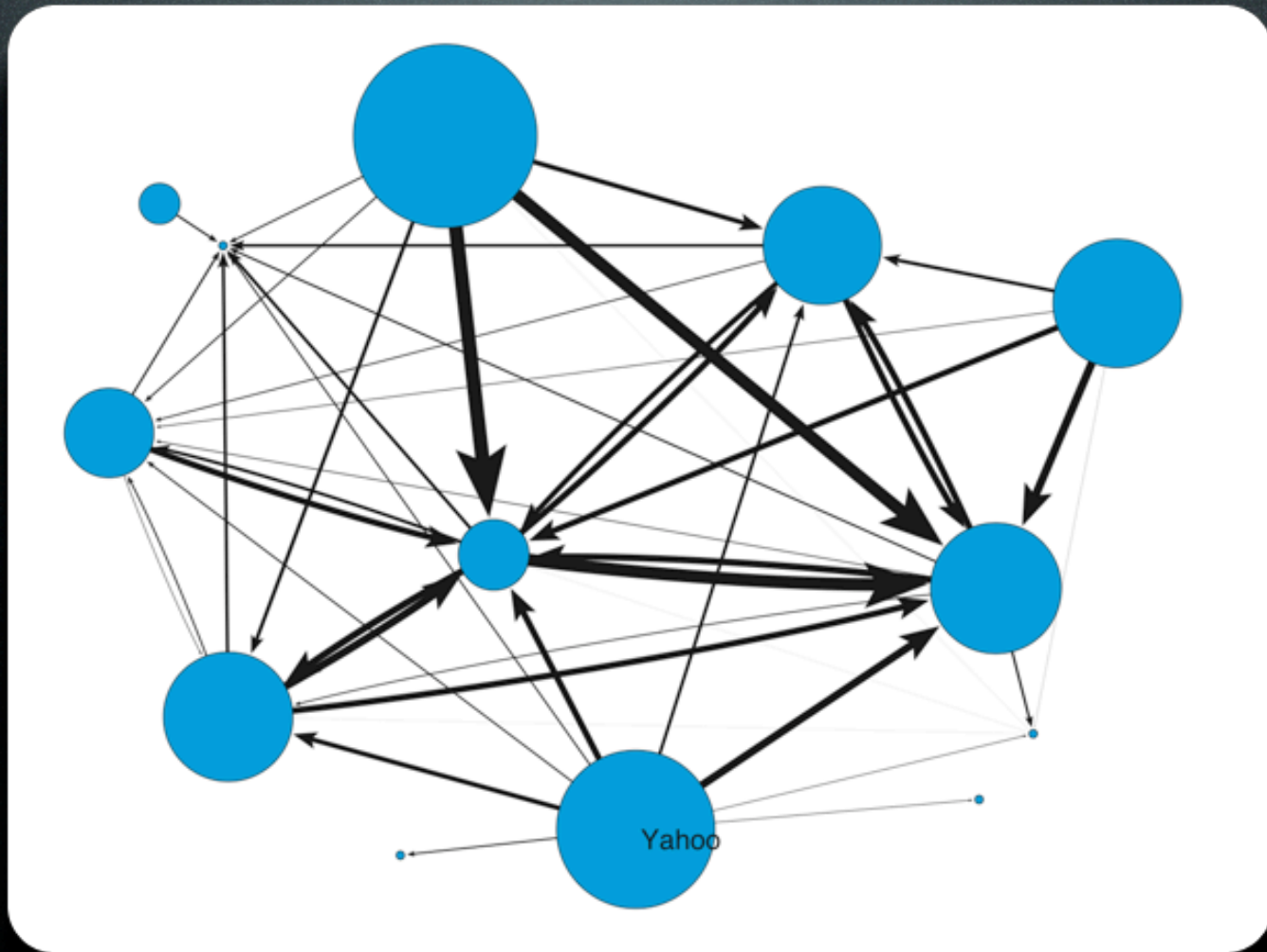
- TREC based evaluation framework (Klampanos et al. 2005)
- User study with 6S application
- Improve coverage/diversity in query routing algorithm
- Spam protection: trust/reputation subsystem



User study



Query network



Result network

Questions?

<http://homer.informatics.indiana.edu/6S/>

The screenshot displays a web browser window with the title "6S peer search system". The address bar shows the URL <http://sulu.informatics.indiana.edu/6S/network/index.html>. The main content area features a red banner with the text "6S-Six Degree of Separation Search System" in white. Below the banner is a navigation menu with buttons for "Home", "Members", "Download", "Publications", "Getting Started", "Network Viewer", and "Forum". A section titled "6S Network" is visible, followed by a network diagram. The diagram consists of green nodes connected by blue arrows, representing a network of connections. Labeled nodes include "qqq", "xt", "RA", "WCCW", "LS", "homea...", "eigmac", "lir", and several "Unkno..." nodes. The browser's status bar at the bottom indicates "Loading 'http://sulu.informatics.indiana.edu/6S/network/index.html', completed 25 of 28 items".