Beyond Google: A Quest to Understand Information Seeking Across Time and People

Chirag Shah
School of Information & Library Science (SILS)
University of North Carolina (UNC) at Chapel Hill
chirag@unc.edu

The “Google effect”

Google michelle obama Search Advanced Search

SafeSearch: Moderate

Web › **Images** Show options... Results **1** - **20** of about **14,000,000** (**0.08** seconds)

Offensive Search Results
www.google.com/resultsinfo
Sometimes our search results can be offensive. We agree. Read more.

President Obama Posters
www.AllPosters.com Vast Selection of President **Obama** Posters. Shop Today and Save!

Barack Obama Fathead®
www.Fathead.com Life-Size Wall Graphics For The Ultimate **Obama** Supporter.

Sponsored Links

Related searches: **michelle obama ape**

AOL says Michelle
300 x 387 - 25k
polichicksonline.com
Find similar images

michelle obama | Leave
380 x 500 - 84k - jpg
politicalkudzu.com
Find similar images

Michelle Obama stokes
413 x 310 - 29k - jpg
weblogs.sun-sentinel.com
Find similar images

Michelle Obama Begins
400 x 300 - 19k - jpg
obamafamilywhitehouse.com
Find similar images

Michelle Obama
330 x 400 - 26k - jpg
freerepublic.com
Find similar images

Michelle Obama's
500 x 500 - 23k - jpg
topnews.in
Find similar images

michelleobama.jpg
298 x 431 - 14k - jpg
powerlineblog.com
Find similar images

Michelle Obama Photo
392 x 629 - 50k - jpg
dancewithshadows.com
Find similar images

A quest for beyond “Google”

- Single-shot to multi-session information seeking
- Solo to social and collaborative activities
- Searching to seeking and discovering
- Retrieving to resolving and synthesizing

Takeaways

- Many situations call for multi-session and/or multi-person information seeking.
- Often we need more information (contextual, related) about an object to make sense of it.
- Social aspect could play a key role in information seeking.
- Need to consider integration, cognitive load, and adoption factors.

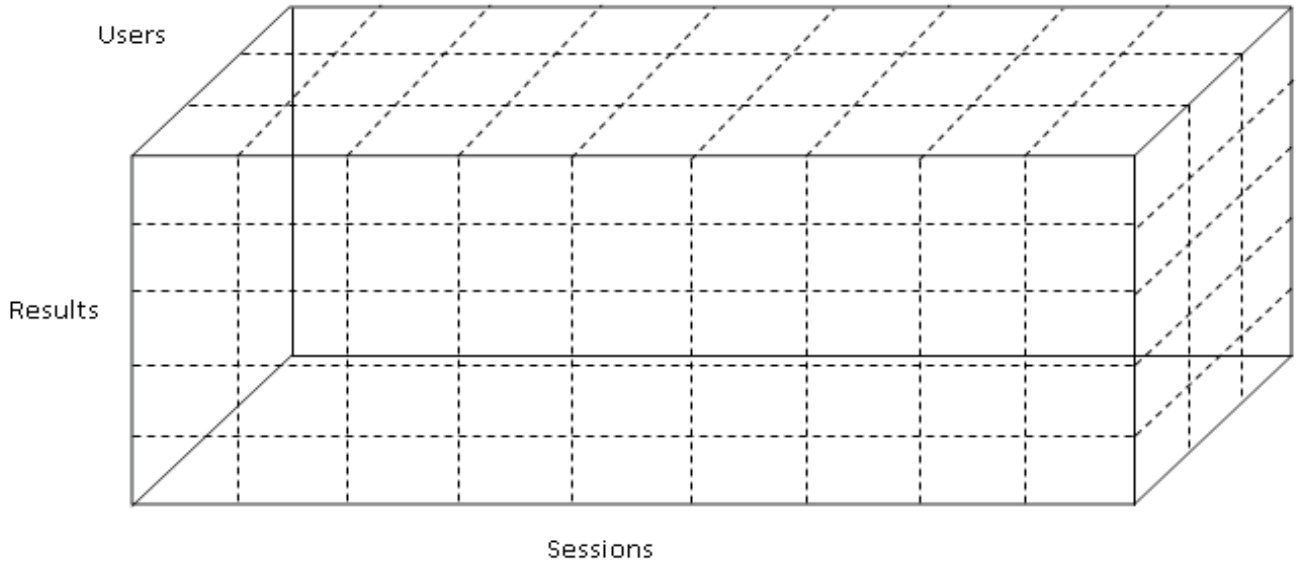
Information Seeking and Usage Over Time, Across People

ResultsSpace framework

Gary Marchionini
Information & Library Science
UNC Chapel Hill

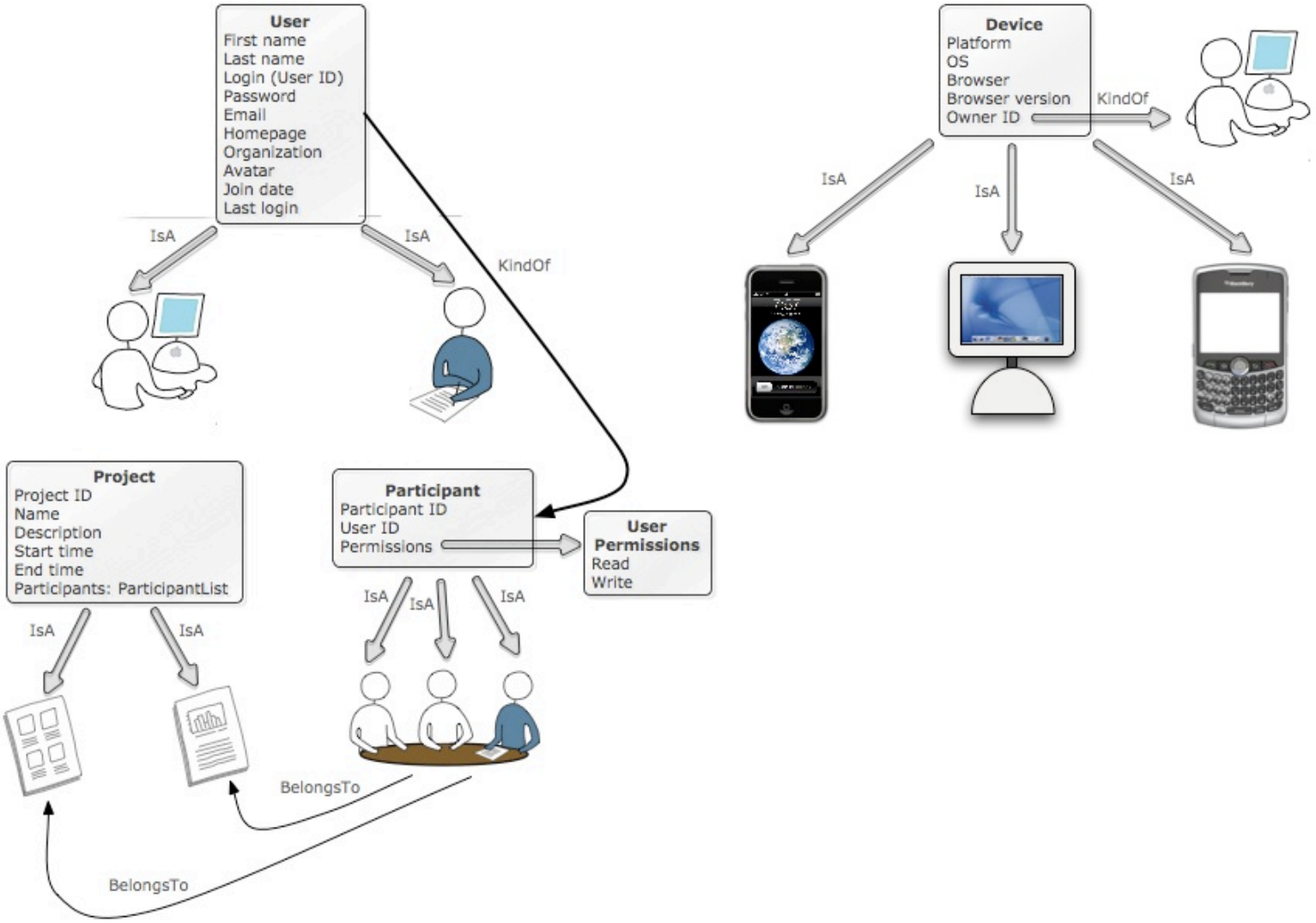
Rob Capra
Information & Library Science
UNC Chapel Hill

NSF supported project starting Fall 2008.

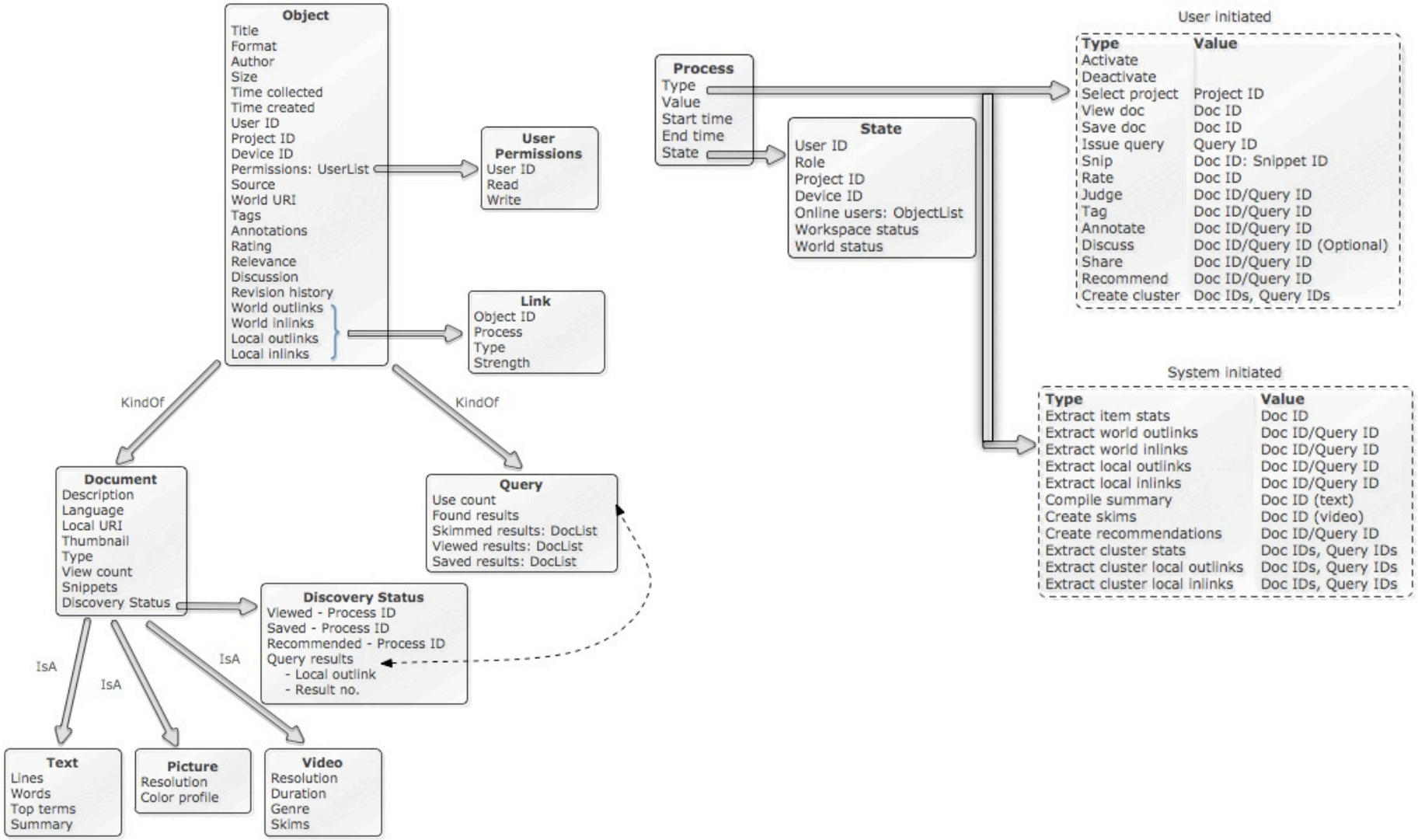


ResultFrame

ResultsSpace entities models



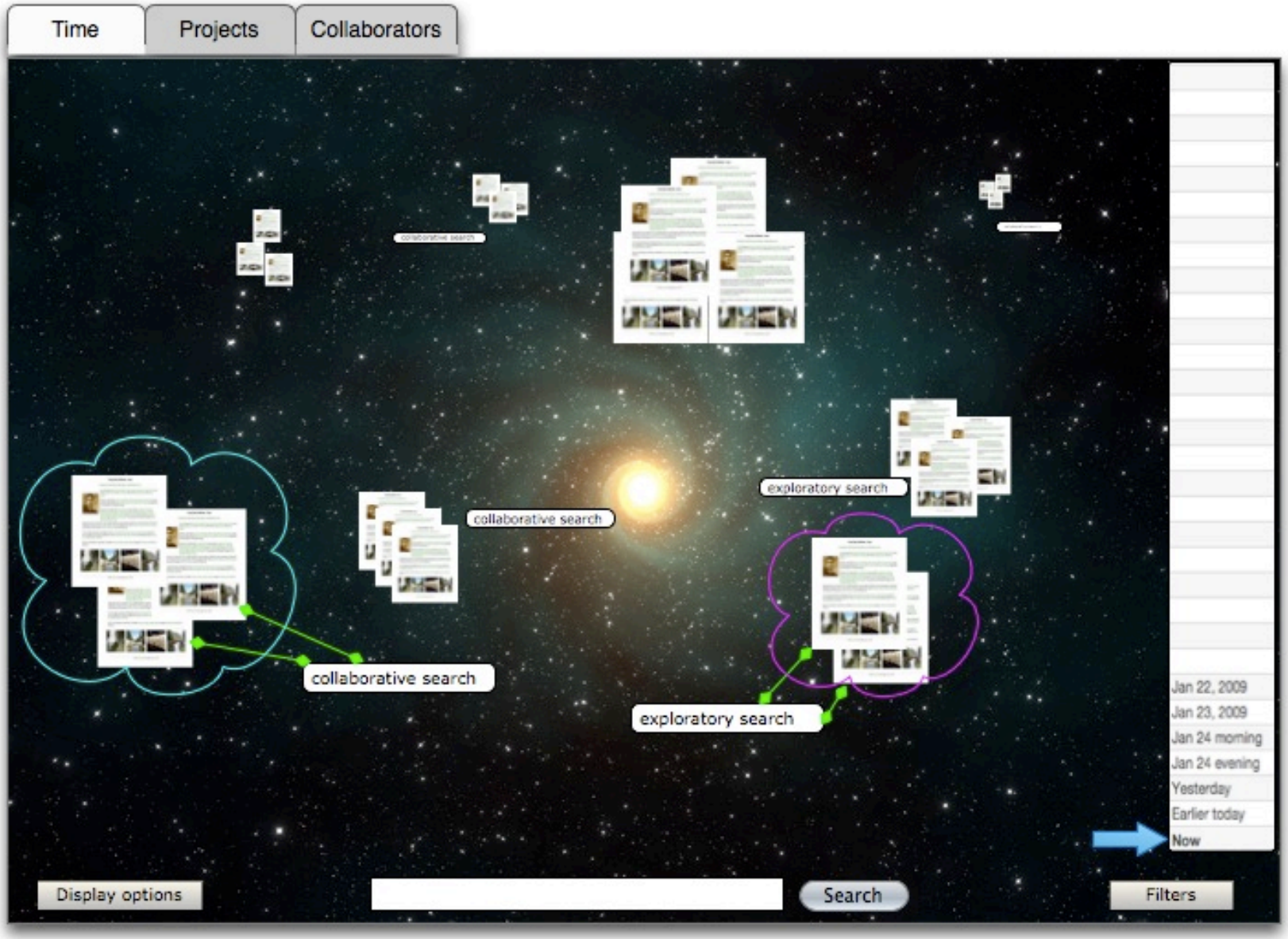
ResultsSpace entities models



Simulation for ResultsSpace

```
=> Searching Live for "type II diabetes"

=> Skimming the results

=> Viewing result # 1
     PageURI: http://search.live.com/health/results.aspx?br=lv&amp;qu=Type II Diabetes

=> Snipping: http://search.live.com/health/results.aspx?br=lv&amp;qu=Type II Diabetes

=> Searching Google for "resultsspace framework"

=> Skimming the results

=> Viewing result # 4
     PageURI: http://www.intute.ac.uk/socialsciences/cgi-bin/search.pl?
     term1=space%20science

=> Tagging document: http://www.intute.ac.uk/socialsciences/cgi-bin/search.pl?
term1=space%20science

=> Annotating document: http://www.intute.ac.uk/socialsciences/cgi-bin/search.pl?
term1=space%20science

=> Judging document http://www.intute.ac.uk/socialsciences/cgi-bin/search.pl?
term1=space%20science
     Thumbs up
```

ResultsSpace mockup



Issues and challenges

- Developing robust data models
- Designing intuitive interfaces
- Addressing cognitive load
- Considering adoption
- Facilitating integration into existing practices

Collaborative Information Seeking

Collaboration in Information Seeking

Requirement or setup

e.g., merger

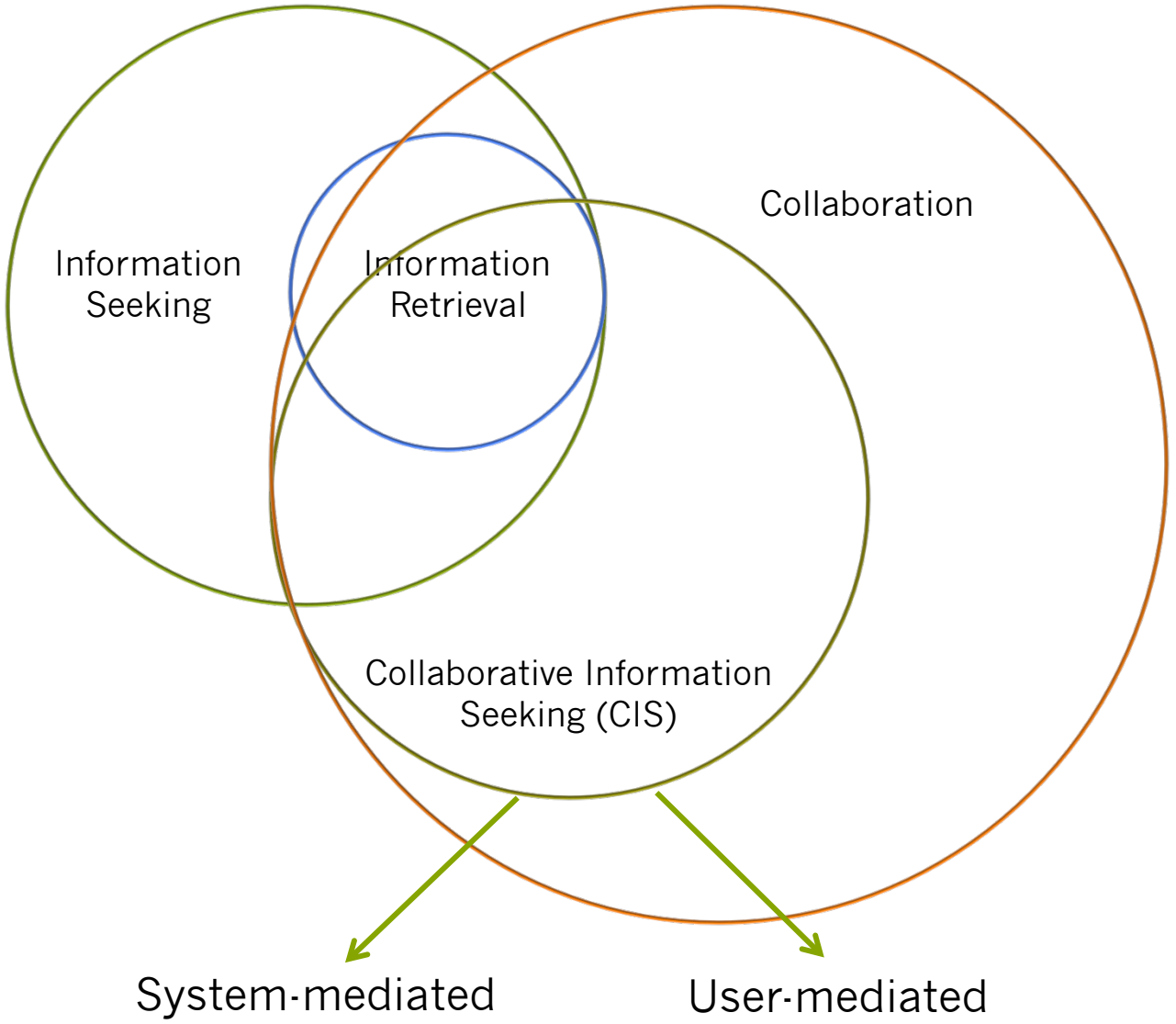
Division of labor

e.g., class assignment

Diversity of skills

e.g., co-authorship

Providing support for collaboration

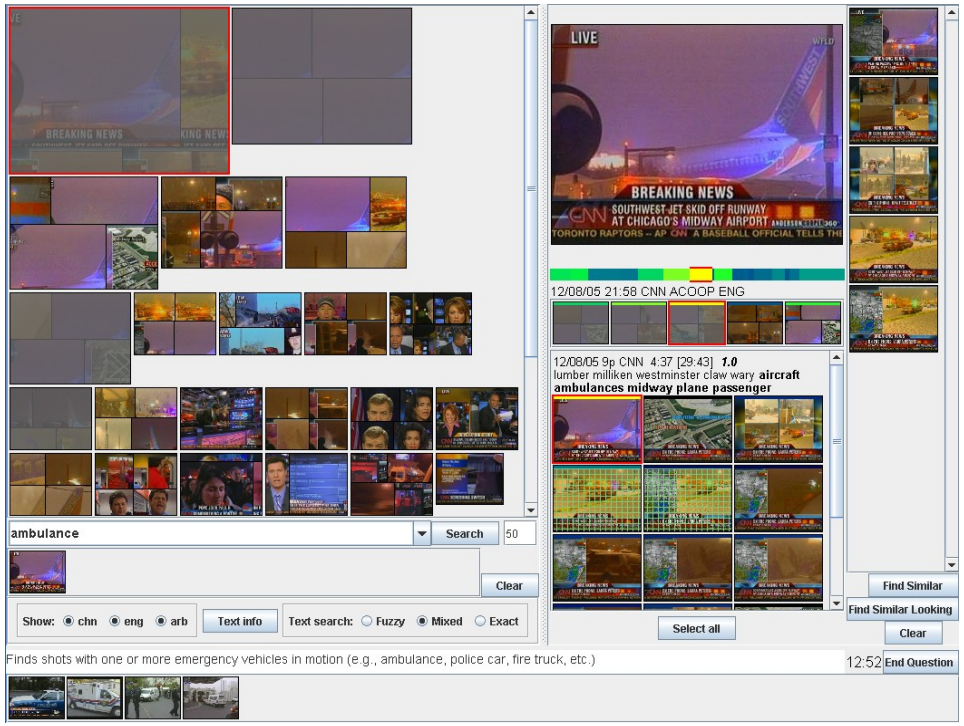


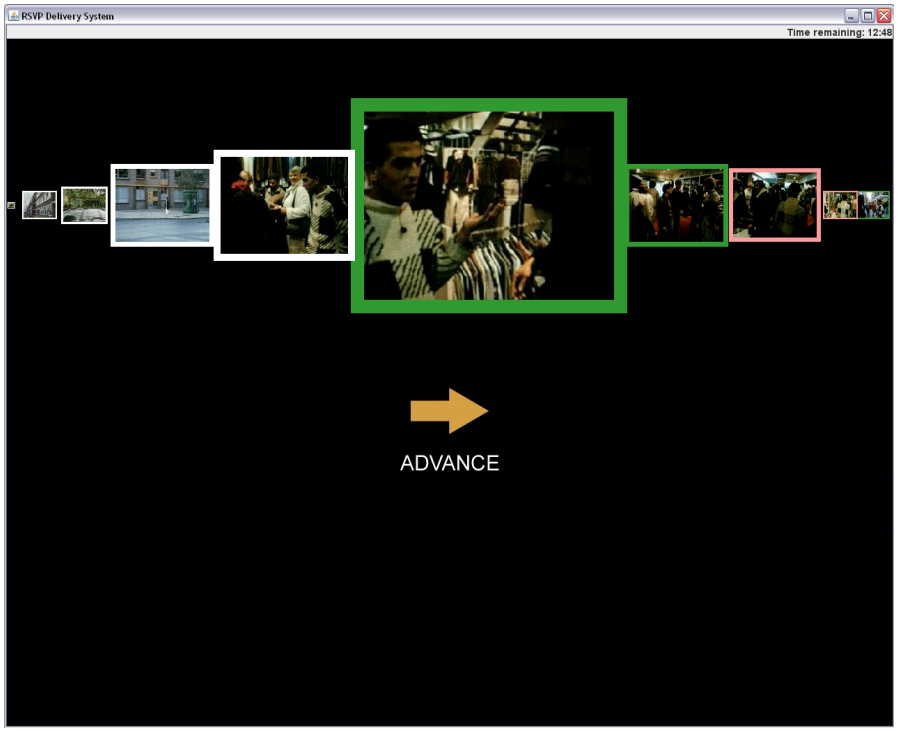
System-mediated collaboration

Jeremy Pickens
Fuji Xerox Palo Alto Lab (FXPAL), California

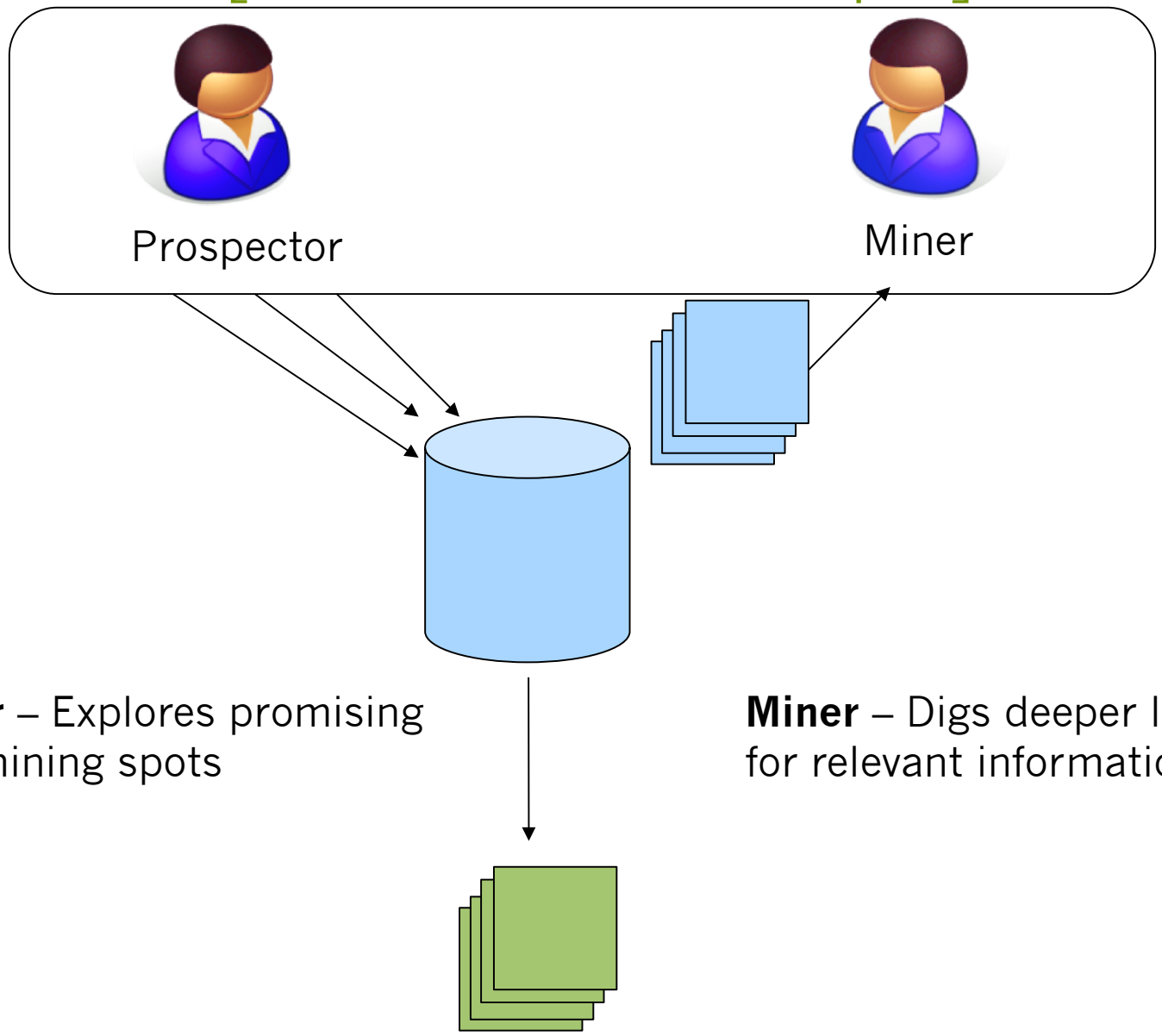
Gene Golovchinsky
Fuji Xerox Palo Alto Lab (FXPAL), California

Internship at FXPAL during summer 2007.





System-mediated asymmetric-role collaboration [SIGIR 2008 Best Paper]



Prospector – Explores promising areas for mining spots

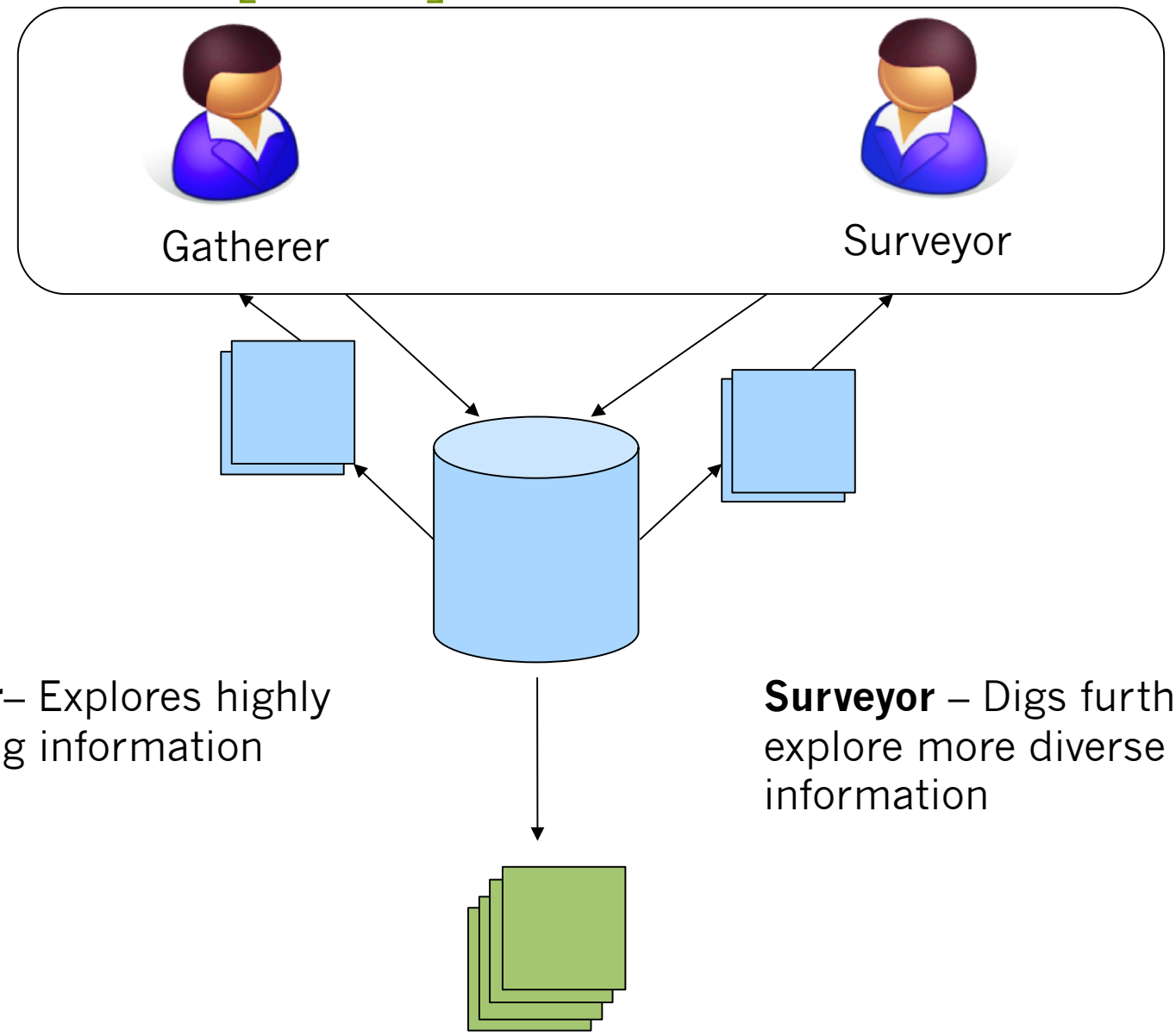
Miner – Digs deeper looking for relevant information

Application to TRECVid

Findings

- Users selected more relevant documents.
- Users found these relevant documents quicker (effects apparent even after 3-4 minutes).
- Users found these relevant documents with less effort (viewed precision).
- Users found relevant documents that otherwise were not found (uniqueness).

System-mediated asymmetric-role collaboration [IP&M]



Gatherer– Explores highly promising information

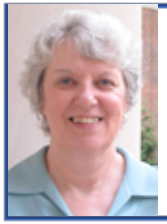
Surveyor – Digs further to explore more diverse information

User-mediated collaboration

Gary Marchionini
Information & Library Science
UNC Chapel Hill

Diane Kelly
Information & Library Science
UNC Chapel Hill

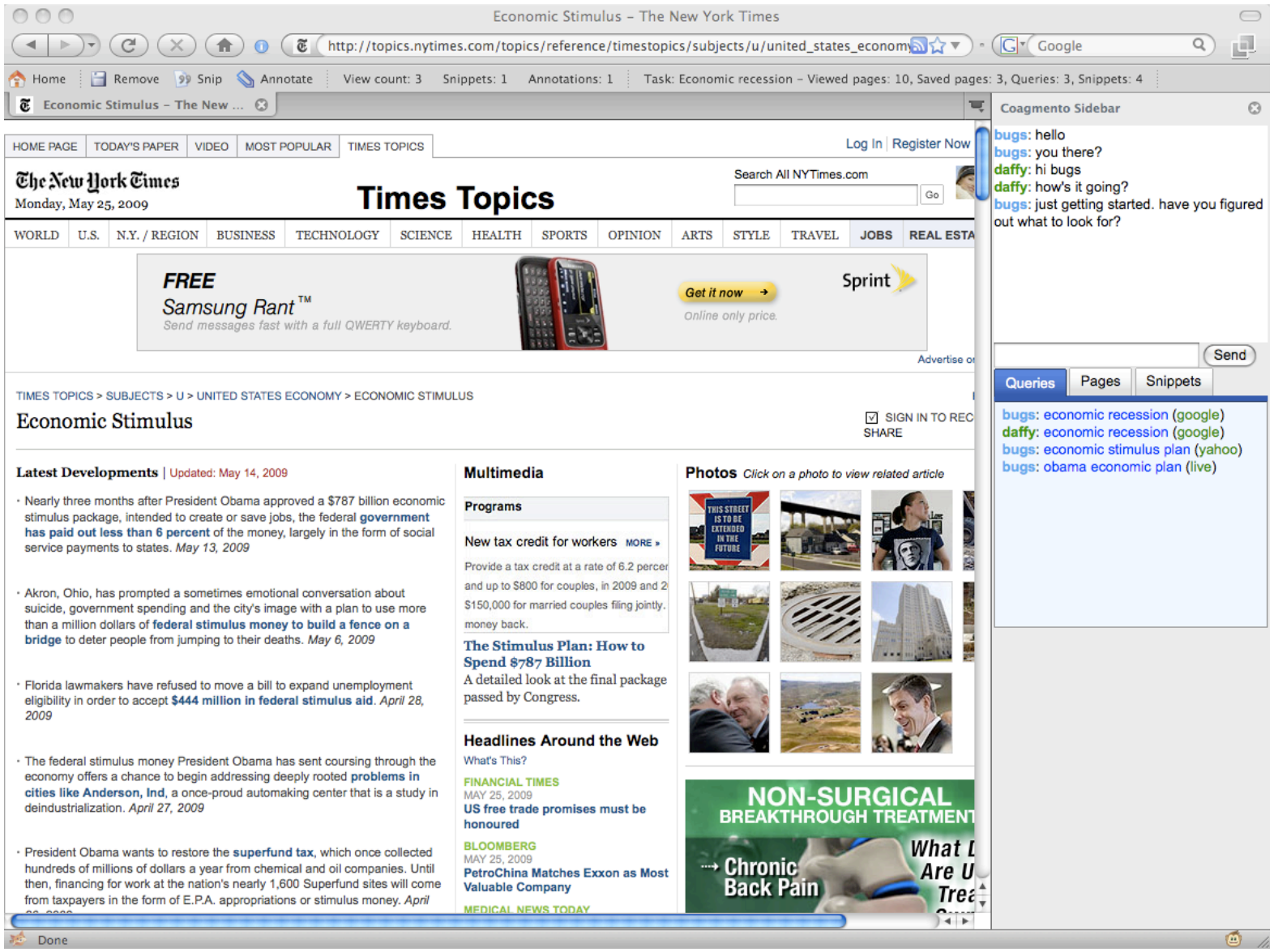
Barbara Wildemuth
Information & Library Science
UNC Chapel Hill



Deborah Barreau
Information & Library Science
UNC Chapel Hill

Susan Dumais
Principle Researcher
Microsoft Research

Coagmento

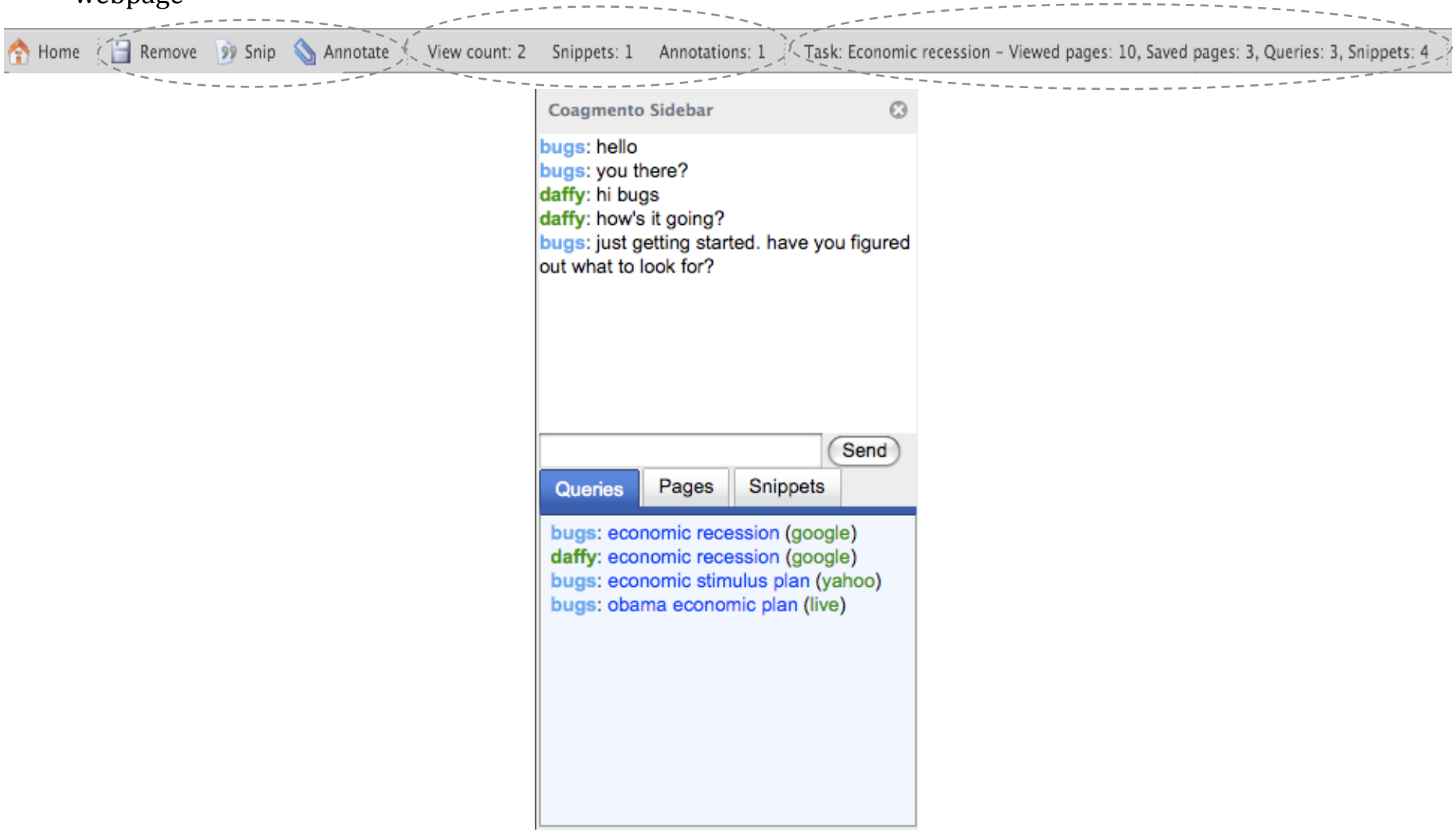

Economic Stimulus – The New York Times
http://topics.nytimes.com/topics/reference/timestopics/subjects/u/united_states_economy
Google
Home
Remove
Snip
Annotate
View count: 3
Snippets: 1
Annotations: 1
Task: Economic recession – Viewed pages: 10, Saved pages: 3, Queries: 3, Snippets: 4
Economic Stimulus – The New ...
HOME PAGE
TODAY'S PAPER
VIDEO
MOST POPULAR
TIMES TOPICS
Log In | Register Now
The New York Times
Monday, May 25, 2009
Times Topics
Search All NYTimes.com
Go
WORLD
U.S.
N.Y. / REGION
BUSINESS
TECHNOLOGY
SCIENCE
HEALTH
SPORTS
OPINION
ARTS
STYLE
TRAVEL
JOBS
REAL ESTA
FREE
Samsung Rant™
Send messages fast with a full QWERTY keyboard.
Get it now
Online only price.
Sprint
Advertise on
TIMES TOPICS > SUBJECTS > U > UNITED STATES ECONOMY > ECONOMIC STIMULUS
Economic Stimulus
SIGN IN TO REC
SHARE
Latest Developments | Updated: May 14, 2009
Nearly three months after President Obama approved a $787 billion economic stimulus package, intended to create or save jobs, the federal government has paid out less than 6 percent of the money, largely in the form of social service payments to states. May 13, 2009
Akron, Ohio, has prompted a sometimes emotional conversation about suicide, government spending and the city's image with a plan to use more than a million dollars of federal stimulus money to build a fence on a bridge to deter people from jumping to their deaths. May 6, 2009
Florida lawmakers have refused to move a bill to expand unemployment eligibility in order to accept $444 million in federal stimulus aid. April 28, 2009
The federal stimulus money President Obama has sent coursing through the economy offers a chance to begin addressing deeply rooted problems in cities like Anderson, Ind, a once-proud automaking center that is a study in deindustrialization. April 27, 2009
President Obama wants to restore the superfund tax, which once collected hundreds of millions of dollars a year from chemical and oil companies. Until then, financing for work at the nation's nearly 1,600 Superfund sites will come from taxpayers in the form of E.P.A. appropriations or stimulus money. April
Multimedia
Programs
New tax credit for workers MORE »
Provide a tax credit at a rate of 6.2 percer
and up to $800 for couples, in 2009 and 2
$150,000 for married couples filing jointly.
money back.
The Stimulus Plan: How to Spend $787 Billion
A detailed look at the final package passed by Congress.
Headlines Around the Web
What's This?
FINANCIAL TIMES
MAY 25, 2009
US free trade promises must be honoured
BLOOMBERG
MAY 25, 2009
PetroChina Matches Exxon as Most Valuable Company
MEDICAL NEWS TODAY
Photos Click on a photo to view related article
THIS STREET IS TO BE EXTENDED IN THE FUTURE
NON-SURGICAL
BREAKTHROUGH TREATMENT
Chronic Back Pain
Coagmento Sidebar
bugs: hello
bugs: you there?
daffy: hi bugs
daffy: how's it going?
bugs: just getting started. have you figured out what to look for?
Send
Queries
Pages
Snippets
bugs: economic recession (google)
daffy: economic recession (google)
bugs: economic stimulus plan (yahoo)
bugs: obama economic plan (live)
Done

Coagmento

Actions that one can take on the displayed webpage

Statistics relating to the current webpage

Statistics relating to the current task



Coagmento workspace

Coagmento

'cause two (or more) heads are better than one!

Collected Snippets

daffy

What Is Recession? Economic Recession Definition | RECESSION.ORG

Many professionals and experts around the world believe that a true economic recession can only be confirmed if GDP (Gross Domestic Product) growth is negative for a period of two or more consecutive quarters.

another definition. this one is from recession.org.

Move to: 1 ->

bugs

Organizing for America | BarackObama.com | Economy

President Obama inherited a terrible mess: a $1.3 trillion deficit, two wars, rising unemployment and unprecedented crises in our banking system.

Move to: 1 ->

Task

A leading newspaper has hired your team to create a comprehensive report on the causes and consequences of the current economic recession in the US. As a part of your contract, you are required to collect all the relevant information from any available online sources that you can find.

1. **Reasons behind this recession.**

   <- Recession - Wikipedia, the free encyclopedia

   In economics, a recession is a general slowdown in economic activity over a sustained period of time, or a business cycle contraction.

   a definition from wikipedia

2. **Effects on some major areas, such as health-care, home ownership, and financial sector (stock market).**
3. **Unemployment statistics over a period of time.**
4. **Proposal, execution, and effects of the economy stimulation plan.**

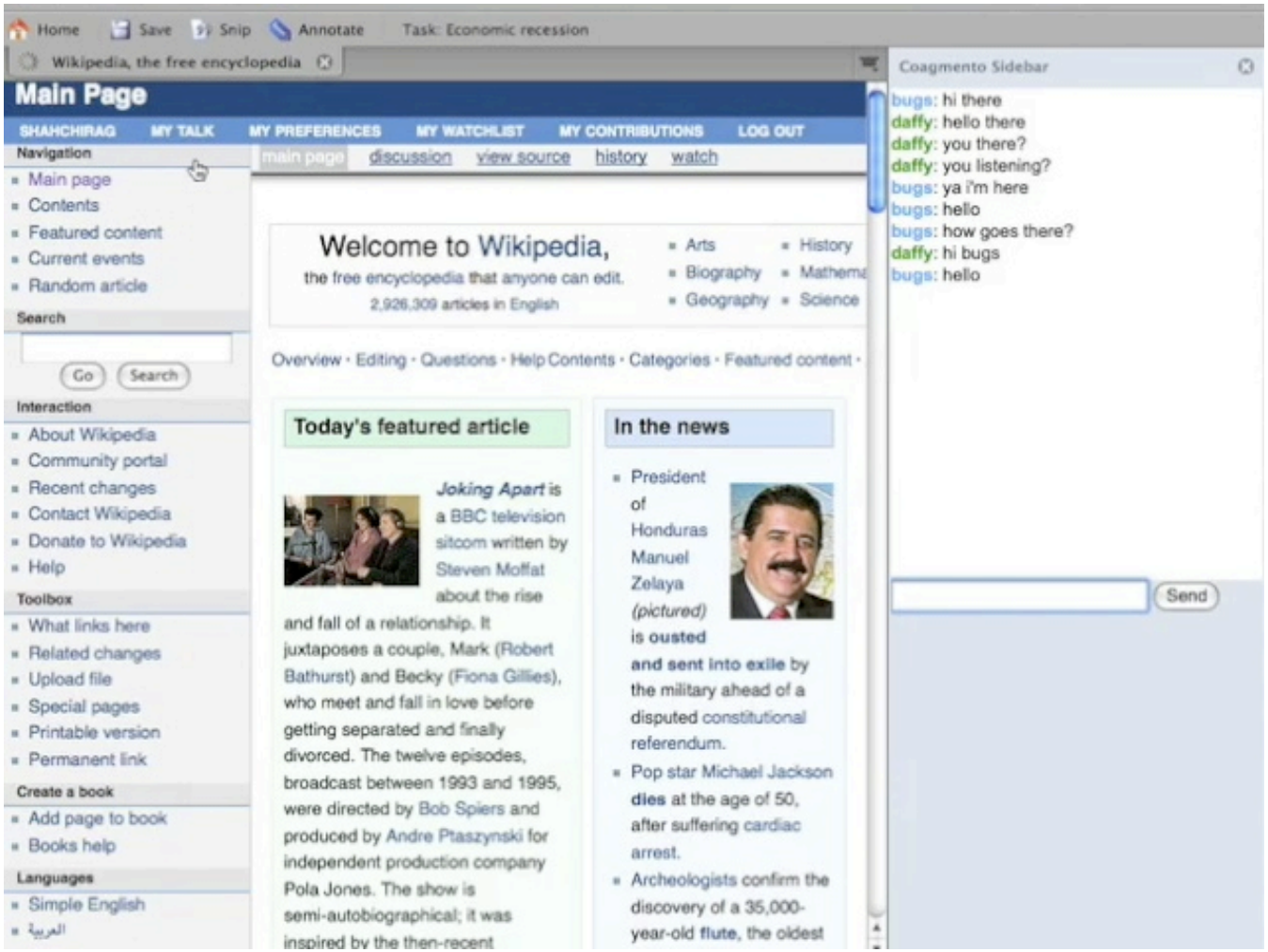
   <- Economic Stimulus - The New York Times

   Nearly three months after President Obama approved a $787 billion economic stimulus package, intended to create or save jobs, the federal government has paid out less than 6 percent of the money, largely in the form of social service payments to states. May 13, 2009

   some stats.

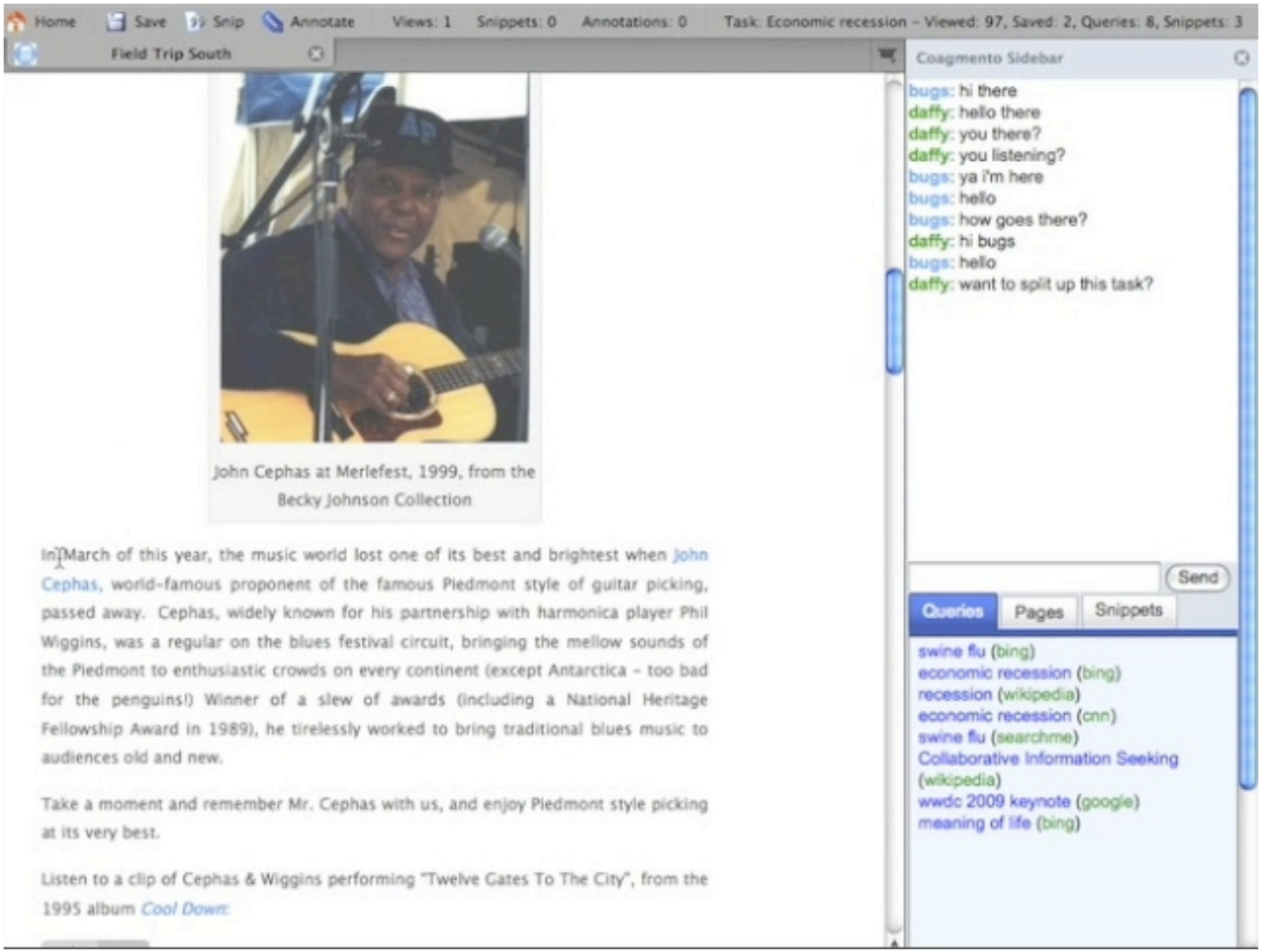
5. **People's opinions and reactions on economy's downfall.**

Study design: conditions



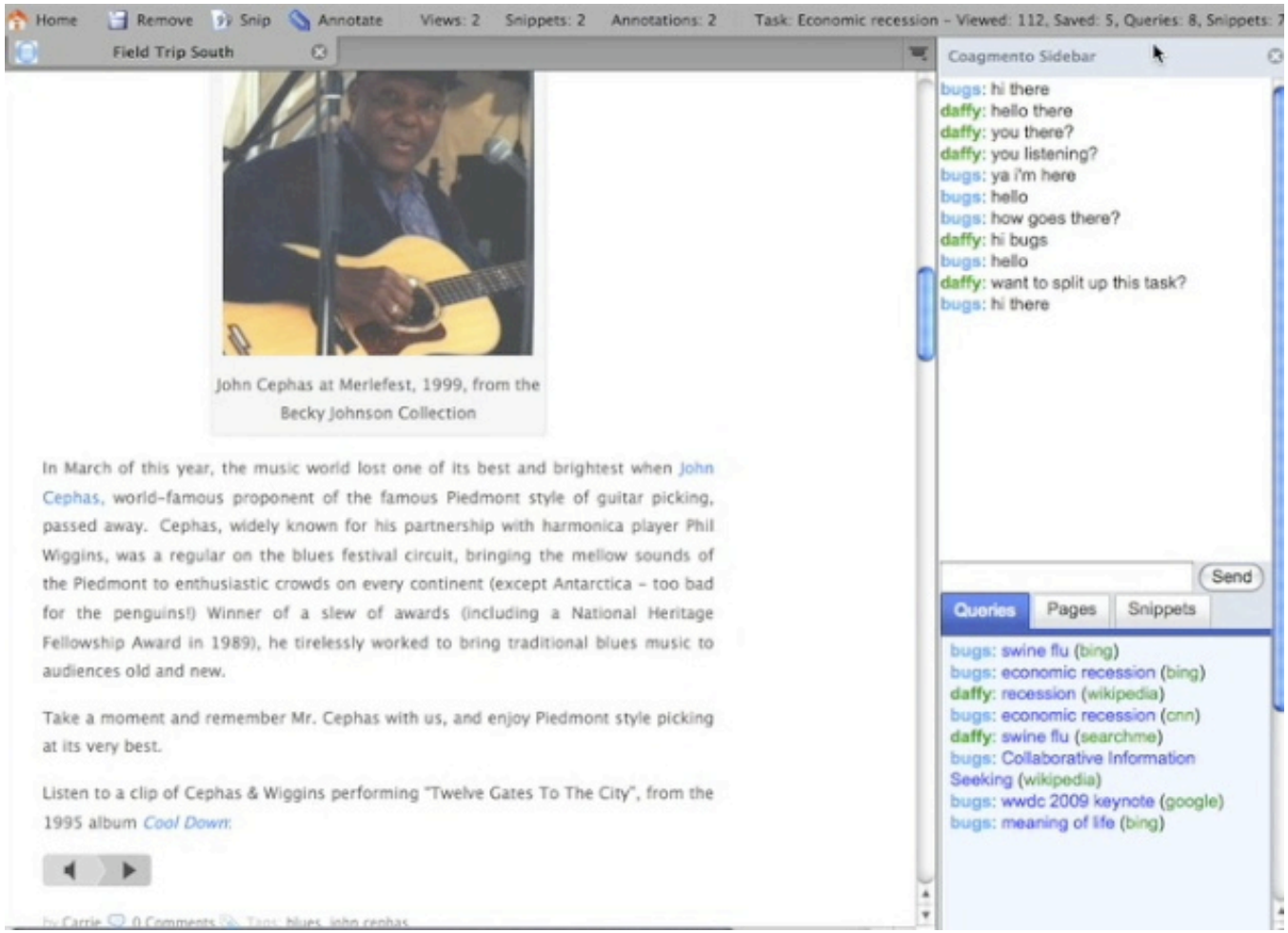
Condition-1
Communication
Workspace awareness
Contextual awareness

Study design: conditions



Condition-2
Communication
Workspace awareness
Contextual awareness
Personal peripheral awareness

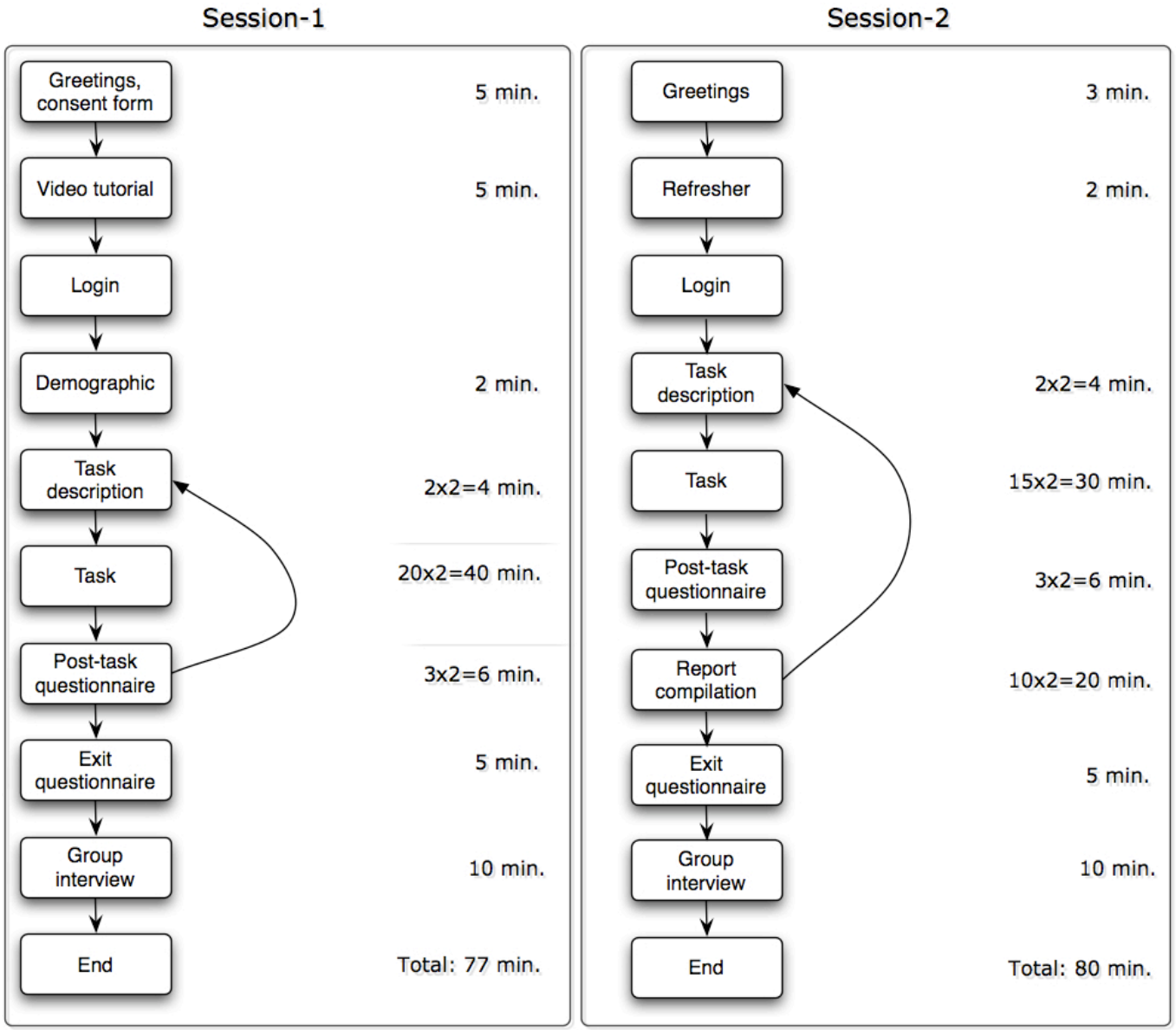
Study design: conditions



Condition-3

Communication
Workspace awareness
Contextual awareness
Group peripheral awareness

Study design: sessions



Study design: subjects and tasks

84 subjects (42 pairs)
Two sessions
14 pairs per condition
27 males, 57 females
Age 17 to 50

Tasks:
1. Economic recession
2. Social networking

Questions and findings

Is there a significant difference in .....

between conditions?	Hypotheses
1. Ease of use and satisfaction	1~2~3
2. Productivity	1<2<3
3. Communication	1>2>3, 1>3, 2>3
4. Engagement	1<2<3
5. Awareness	1<2<3, 1<3, 2<3

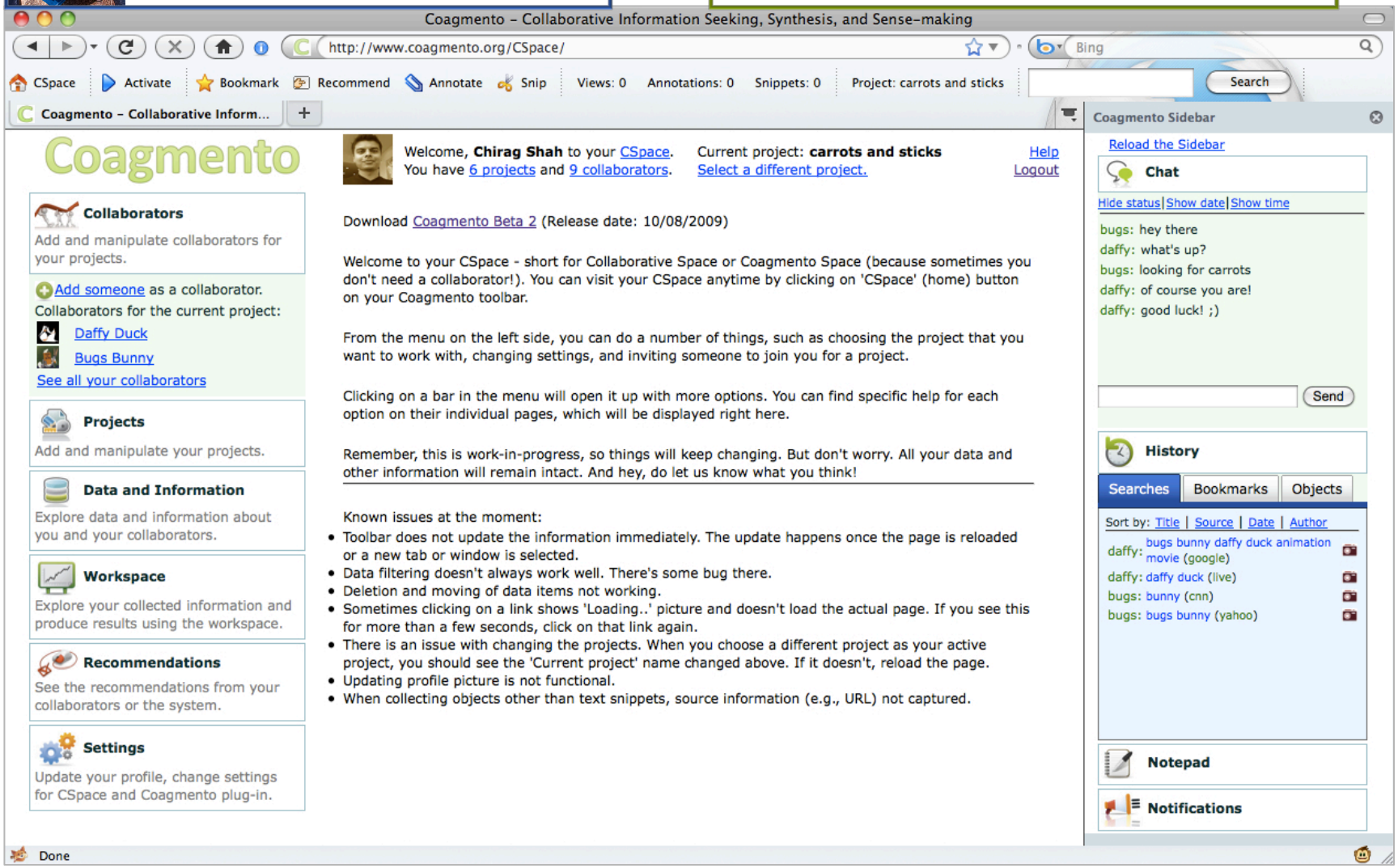
Finding

Basic group awareness components offer substantial advantages to collaborative information seeking without adding new burdens to users.

Coagmento 2.0 coagmento.org

Lori Mears
Fiancé, toughest critic
Duke Lemur Center

Coagmento signups this week:
40+ regular users
10+ subjects



Issues and challenges

- Extend single-user models to collaborative users.
- Identify situations appropriate for collaboration.
- Provide specialized tools and support for collaboration.
- Design integrated system solutions.

Shah, Chirag, Marchionini, Gary, and Kelly, Diane (2009). Learning Design Principles for a Collaborative Information Seeking System. Proceedings of ACM SIGCHI 2009. Boston, MA: April 4-9, 2009.

Shah, Chirag. Designs for Systems to Support Collaborative Information Behavior (in press). Book chapter to be published in Foster, Jonathan (Ed.) *Collaborative Information Behavior*. IGI Golbal, Hershey PA.

Social Information Seeking

Seeking for the wisdom of crowds

Based on your own family's experience, what do you think we should do to improve health care in America?

4962 ☆ In Other - General Health Care - Asked by Hillary Clinton - 38022 answers - 3 years ago

How can we engage more people in the democratic process?

4466 ☆ In Civic Participation - Asked by Barack Obama - 17377 answers - 3 years ago

What should we do to free our planet from terrorism?

4146 ☆ In Civic Participation - Asked by Dr. APJ Abdul Kalam - 29727 answers - 3 years ago

How will Sanjaya's continued success on American Idol impact the show's credibility?

2622 ☆ In Reality Television - Asked by Constantine Maroulis - 24754 answers - 3 years ago

What would you do to stop wasteful government spending in Washington?

2071 ☆ In Government - Asked by John McCain - 16709 answers - 3 years ago

What will you remember most about Michael Jackson?

1862 ☆ In Celebrities - Asked by Jack - 6416 answers - 5 months ago

Our collection (as of today)

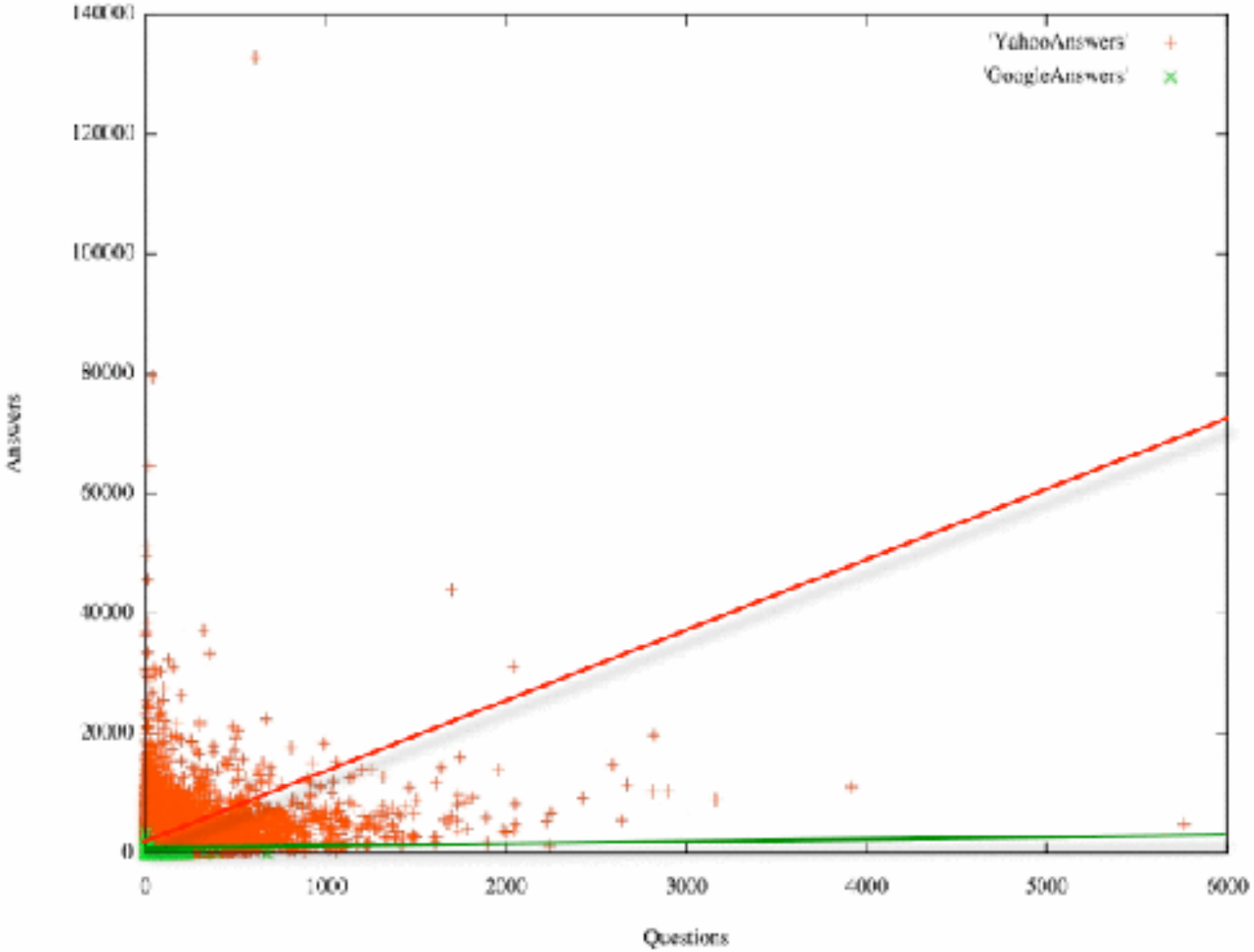
Yahoo! Answers

- Questions: 2 million+
- Answers: 13 million+
- Avg. answers per question: 6+
- Users: ~1 million

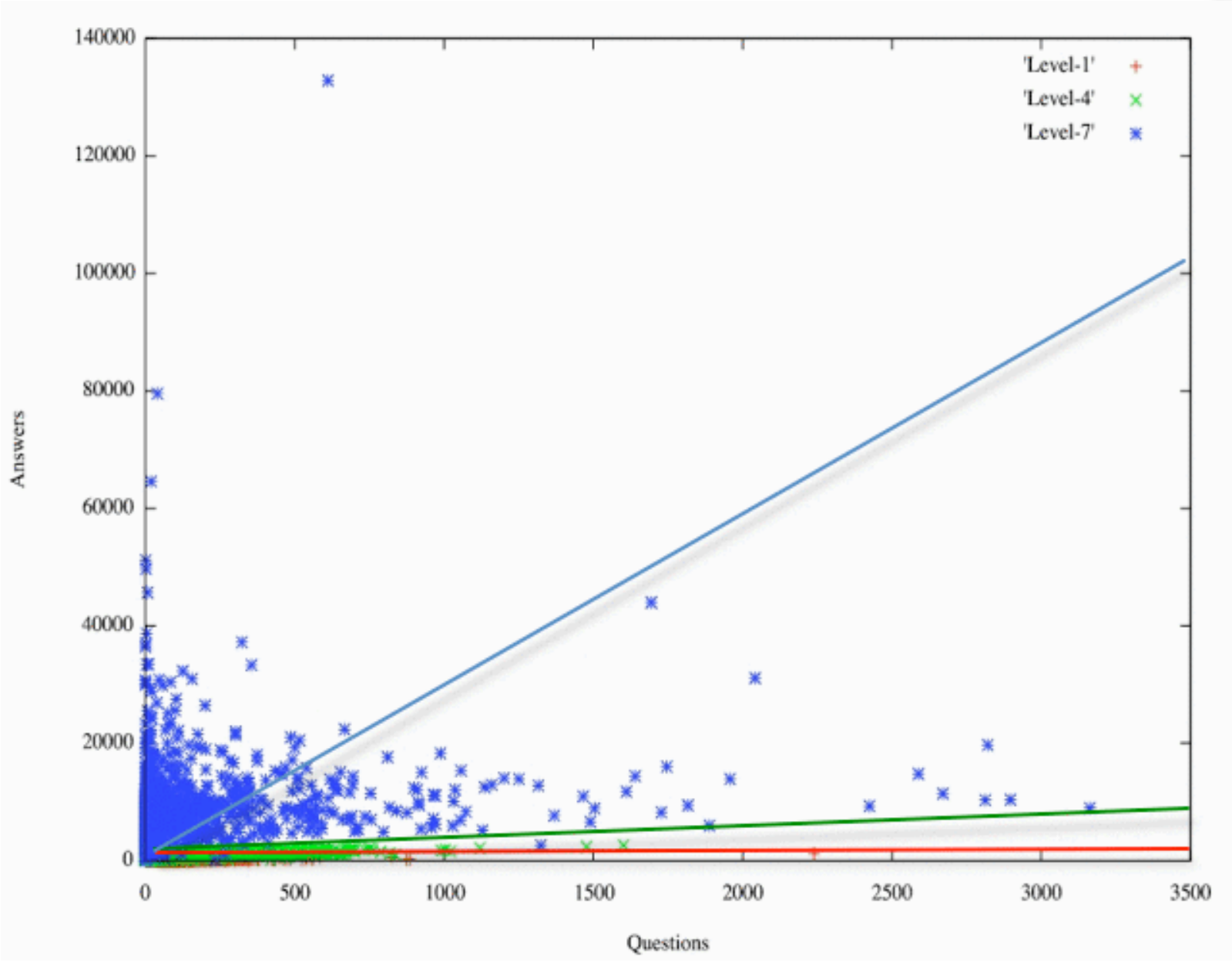
Google Answers

- Questions: 153,343
- Answers: 56,835
- Avg. answers per question: 0.37
- Users: 83,454 + 534

Yahoo! Answers vs. Google Answers



Consumers and contributors



Seeking information in online communities

- Social capital is a big driving force.
- Aim for self-sustainable communities.
- Often, crowd is the smartest “person”!

Ongoing work

Jeffrey Pomerantz
Information & Library Science
UNC Chapel Hill

ipl The Internet Public Library

- Measuring content quality – proposed Turk study.
- Developing models for quality prediction.
- Digital reference vs. social Q&A.

Issues and challenges

- Motivations
- Expectations
- Evaluating content quality and user satisfaction

Shah, Chirag, Oh, Sanghee, and Oh, Jung Sun (2009). Research agenda for Social Q&A. *Library & Information Science Research (LISR), 31*(4), 205-209.

Information Extraction & Discovery

Collecting contextual information

Contextual information helps makes sense of an information object.

Gary Marchionini
Information & Library Science
UNC Chapel Hill

Helen Tibbo
Information & Library Science
UNC Chapel Hill

Cal Lee
Information & Library Science
UNC Chapel Hill

Paul Jones
Journalism & Mass Comm.
UNC Chapel Hill

ContextMiner

VidArch campaigns

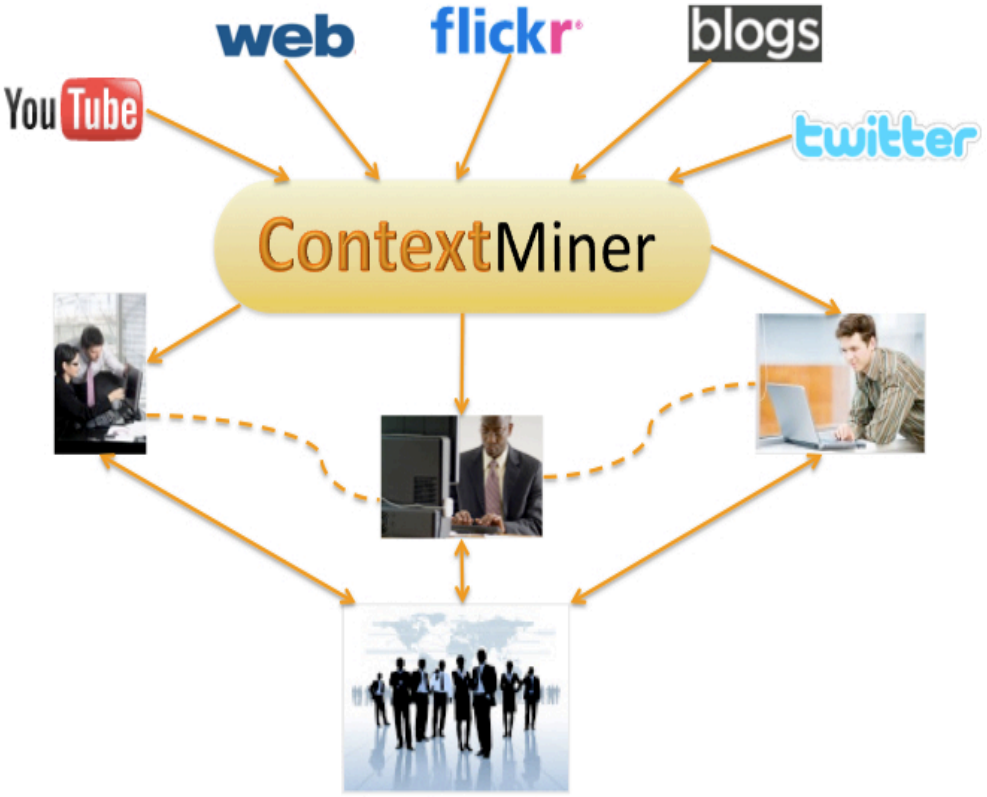
- Elections 2008
- Energy
- Epidemics
- Health
- Natural disasters
- Truth commissions

Data

- Nearly 100,000 videos
- Thousands of user profiles
- Millions of comments
- Contextual attributes for each
- Over two years of collection

ContextMiner today

www.**Context**Miner.org



- 300+ users
- 800+ campaigns
- Millions of objects (videos, blogs, webpages, etc.)

ContextMiner users

US Academic Institutes

- California State University
- Drexel
- Harvard University
- Kent University
- NC State University
- Northwestern University
- Stanford University
- University of California at Santa Barbara
- University of Hawaii
- University of Idaho
- University of Maryland
- University of Massachusetts Amherst
- University of Michigan
- University of North Carolina (Chapel Hill and Greensboro)
- University of Pennsylvania
- University of Puerto Rico
- University of Texas at Austin
- University of Wisconsin (Madison and Milwaukee)
- Yale University

ContextMiner users

Non-US Academic Institutes

- Aarhus University, Denmark
- American University of Beirut, Lebanon
- Chaitanya PG College, India
- Delft University of Technology, The Netherlands
- Education AU, Australia
- Manchester University, UK
- National University of Singapore
- PRRM Engineering College, India
- Simon Fraser University, Canada
- University of Alberta, Canada
- University of Florence, Italy
- University of Hong Kong
- University of London, UK
- University of Paris, France
- University of Seville, Spain
- University of Sydney, Australia
- Université de Versailles-Saint-Quentin, France
- Westminster College, UK

ContextMiner users

Organizations

- Academy of Motion Picture Arts and Sciences (Oscars)
- Bauer Media, UK
- Cruise Addicts
- Elsevier
- European Archive
- Google
- IBM
- Library of Congress
- New South Wales Rural Fire Service, Australia
- Optimedia
- Razorfish
- R/GA
- SACE (Southern Alliance for Clean Energy)
- Soft Experience, France
- State Archive in Lublin, Poland
- US Navy
- WRAL TV Station, Raleigh NC

ContextMiner tomorrow

Paul Jones
Journalism & Mass Comm.
UNC Chapel Hill

Sri Kalyanaraman
Journalism
UNC Chapel Hill

John Conway
Director, Creative Services
WRAL TV Station, Raleigh NC

Kirk Ross
Editor
Carrboro Citizen Newspaper

Michael Tiemann
Vice President
Open Source Affairs, Red Hat

Shah, Chirag (2009). Mining Contextual Information for Ephemeral Digital Video Preservation. *International Journal of Digital Curation (IJDC), 4*(1).
Shah, Chirag (2009). ContextMiner: Supporting the Mining of Contextual Information for Ephemeral Digital Video Preservation. *International Journal of Digital Curation (IJDC), 4*(2).

More power to the people

TubeKit
A YouTube Crawling Toolkit

Helping people with contextual information

Dhavan Shah
Journalism & Mass Comm.
Univ. of Wisconsin-Madison

CHESS
Comprehensive Health Enhancement Support System

Issues being studied:

- Proposition 8 in California
- Question 1 in Maine
- Social information seeking for cancer patients

Thorson, Kjerstin, Borah, Porismita, Ekdale, Brian, Namkoong, Kang, and **Shah, Chirag** (in press). YouTube and Proposition 8: a case study in video activism. To be published in *Journal of Information, Communication, and Society (iCS)*.

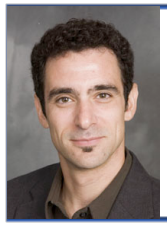
Discovering patterns from contextual information

Stuart Shulman
Political Science
UPitt & UMass Amherst

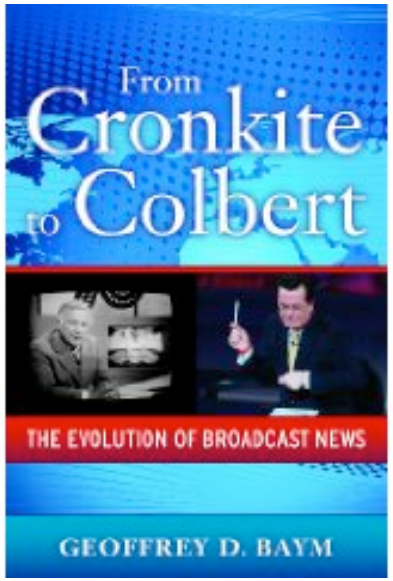
texifter

Shah, Chirag (in press). Supporting Research Data Collection from YouTube with TubeKit. To be published in *Journal of Information Technology & Politics*.

Republic of Web



Geoffrey Baym
Media Studies
UNC Greensboro



- Republic of Web – A Web of, by, and for the people.
- Track the spread of primary information through secondary channels.
- Information derivatives and their interconnections to study the population, their preferences, and behavioral patterns.

Republic of Web

http://www.colbertnation.com/the-colbert-report-videos/227618/may-13-2009/michael-pollan (380)
http://www.colbertnation.com/the-colbert-report-videos/220021/june-03-2009/eric-schlosser (295)
http://www.thedailyshow.com/watch/thu-july-2-2009/robert-kenner (111)
http://www.thedailyshow.com/watch/thu-may-14-2009/lisa-p--jackson (94)
http://www.colbertnation.com/the-colbert-report-videos/230610/june-16-2009/jim-rogers (267)
http://www.colbertnation.com/the-colbert-report-videos/232640/july-01-2009/nicholas-kristof (358)
http://www.thedailyshow.com/watch/thu-july-16-2009/robert-glennon (540)

Michael Pollan Trades Barbs With Stephen Colbert | The Food Section ... (33)
The Food Section - Food News, Recipes, and More | The Food Section ... (999)
The Food Section - Food News, Recipes, and More | The Food ... (0)

Michael Pollan Trades Barbs With Stephen Colbert | The Food Section ...
New York - Fork in the Road - The Blogroll Gazette--Bruni Bombshell ...
le cochon de chef
soupnancy - My Blog: Bits and bites
News | Melgab International CC
Megite Food News: What's Happening Right Now
Michale Pollan - news
Your FeedBlitz Updates
Specialty teas gain foothold - Related Stories - National Restaurant ...
The Food Section - Food News, Recipes, and More | The Food Section ...
Stephen Colbert : Pictures, Videos, Breaking News
The Food Section:
the Food on Kosmix
Megite Food News: What's Happening Right Now
SmartBrief: culinary - culinary arts - culinary competitions -

Thursday, December 03, 2009

The Food Section

DISCOVER BANK

Michael Pollan Trades Barbs With Stephen

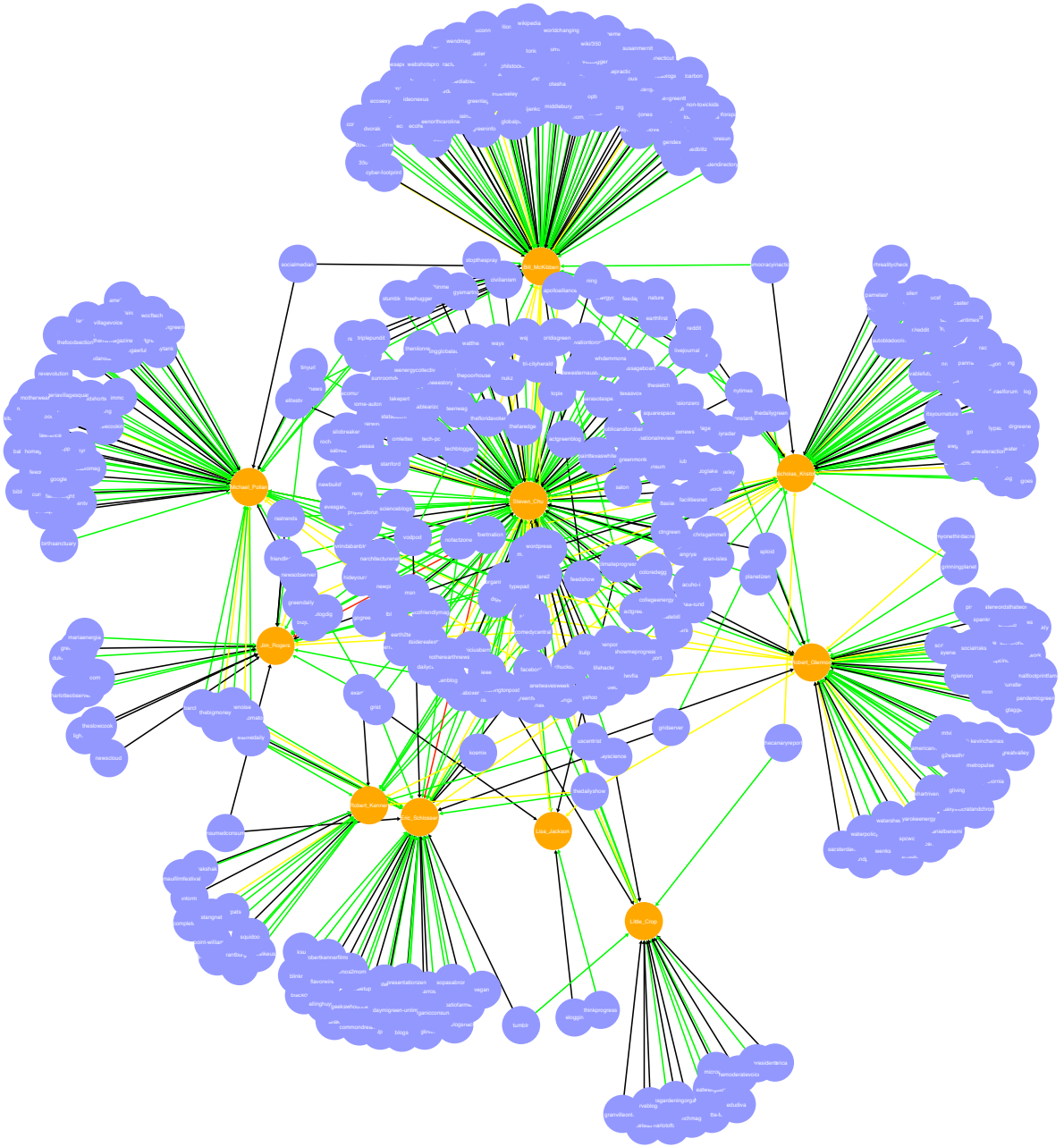
The Colbert Report — Mon - Thurs 11:30pm / 10:30c
Michael Pollan

Colbert Report Full Episodes | Political Humor | Gay Marriage

In this hilarious interview on "The Colbert Report," author Micha Colbert argue about American cheese, food science, crack, soda, a advancements in Cheetos; Pollan discovers live on TV that he was credits high fructose corn syrup for his healthy head of hair.

POSTED BY JOSH FRIEDLAND ON MAY 14, 2009 IN Television, What's Fresh | Li

Republic of Web



Issues and challenges

- What is context?
- How much is enough?
- Evaluation
- Methods for knowledge discovery

Takeaways

- Many situations call for multi-session and/or multi-person information seeking.
- Often we need more information (contextual, related) about an object to make sense of it.
- Social aspect could play a key role in information seeking.
- Need to consider integration, cognitive load, and adoption factors.

Sites and services

Coagmento.org

ContextMiner.org

TubeKit.org

RepublicOfWeb.org

XCreator.org

http://idl.ils.unc.edu/~chirag/**DiscoverInfo**/

http://idl.ils.unc.edu/~chirag/**DIToolkit**/

Many thanks to...

Associates and Collaborators

Carrboro Citizen Newspaper

FXPAL

Internet Archives

Library of Congress

Microsoft Research

Red Hat

University of Massachusetts at Amherst

University of North Carolina at Chapel Hill

University of North Carolina at Greensboro

University of Pittsburgh

University of Wisconsin at Madison

WRAL TV Station

Disciplines of Collaborators

Communications

Computer Science

Engineering

Health Science

Information & Library Science

Journalism

Mass Media, Media Studies

Political Science

Funding Support

ACM

Google

IEEE

Microsoft

NCR

NSF

SIG USE

UNC