Web Science & Online Communities

Amy Bruckman

School of Interactive Computing

Georgia Tech | College of Computing

The Internet Changes Everything

- "It seems passé today to speak of 'the Internet revolution.' In some academic circles, it is positively naïve. But it should not be. The change brought about by the networked information environment is deep. It is structural. It goes to the very foundations of how liberal markets and liberal democracies have coevolved for almost two centuries."
  – Yochai Benkler, *The Wealth of Networks*

Georgia Tech | College of Computing

Outline

- The Web Science Initiative
- Research in the Electronic Learning Communities Group

What is "Web Science"?

- ## The interdisciplinary study of the Internet
  - ### Inter-disciplinarity
  - ### New ways to work with industry
  - ### New educational initiatives
  - ### Fund raising
  - ### Strategic hiring

Georgia Tech | College of Computing

Interdisciplinarity

- Within the College
  - Leads:
    - Interactive Computing:
      - Bruckman, social computing
      - Irfan Essa, computational journalism
    - Computer Science:
      - Milena Mihail, theory
      - Constantine Dovrolis, networking
  - As we grow, we need to stay in touch
- Collaborators across campus
  - Public Policy (Hans Klein)
  - International Affairs (Mike Best)
  - Management (Nicholas Lurie)
  - ECE (George Riley)

Georgia Tech | College of Computing

Working with Industry

- The online communities dissertation, late 1990s
  - No longer viable for one person to do everything
  - Need access to real systems, real data
- Sharing data
  - AOL gets in big trouble for releasing data publicly
  - But can still release anonymized data to researchers
- Student internships to work on research
  - Maintain right to publish?
- Intellectual property challenges
  - New Microsoft model: option set up in advance to buy IP

Georgia Tech | College of Computing

Educational Initiatives

- Goal: to make our grads the most sought after by industry
  - "I wanted to write and tell you that while online communities was my favorite class in grad school, I never got a chance to use anything I learned in that class until now. I just recently accepted a job with Amazon.com to work on Askville.com, a fairly young little online community we're trying to grow into a larger, happier online community."

- What kind of preparation is needed?
  - Google's answer: two kinds
    - Classic: as rigorous as possible technical training, especially theory
    - Applied: like our Human Centered Computing program
  - Other answers?

Georgia Tech | College of Computing

Educational Initiatives

- Initial small steps:
  - HCC PhD area in Social Computing (fall 2007)
  - New courses:
    - Undergrad version of Design of Online Communities (Bruckman, spring 2008)
    - Networks and the WWW (Mihail, fall 2008)
- Longer term:
  - Certificate?
  - Minor?
  - Thread?
    - Or combination of threads?
      - Proposal: Networking + People + Web industry internship -> Web Science certificate
  - Degree program?

Georgia Tech | **College of Computing**

Fund Raising

- NSF Opportunities to Consider
  - Integrative Graduate Education and Research Traineeship (IGERT) program
    - $3 million over five years to support interdisciplinary graduate studies
  - Cyber-enabled Discovery and Innovation (CDI)
    - Three areas:
      - From data to knowledge
      - Understanding complexity in natural, built, and social systems
      - Building virtual organizations
    - Emphasis on bold, interdisciplinary work
- Industrial Funding?
  - How do we make that happen on a larger scale?

Strategic Hiring

- Are there key people who would help us take this effort to the next level?
  - Senior hires
  - Junior hires

User-Generated Content Online

- ## The World in 1995:
  - The Internet can help individuals become creators, not merely recipients, of content
  - Democratizing force
  - Educational opportunity
- ## The World in 2000:
  - Lots of commercially published (one to many) content
  - Maybe it's business as usual after all

The World in 2007

- User-generated content is happening!
  - The Blogosphere
    - As predicted by science fiction writer Orson Scott Card
  - Wikipedia
  - MySpace
  - YouTube
  - Etc.
- The results:
  - Citizen journalists, artists, activists, scientists
  - Gossip, copyright violations, really bad poetry

Georgia Tech | College of Computing

Outline of the Rest of the Talk

- Why is the Internet an interesting learning environment?
  - Constructionist learning & online communities
  - What is a "community"?
- Electronic Learning Communities (ELC) research:
  - Understanding Wikipedia
  - Science Online
  - GameLog and The Game Ontology Project
  - Splat!
  - Computer-Supported Collaborative Innovation (CSCI)
- Conclusion: designing for educational opportunity

Constructionist Learning

- Pedagogy:
  - Piaget's constructivism
  - Papert's constructionism
    - Learning by working on personally meaningful projects
- Examples:
  - Logo
  - Microworlds
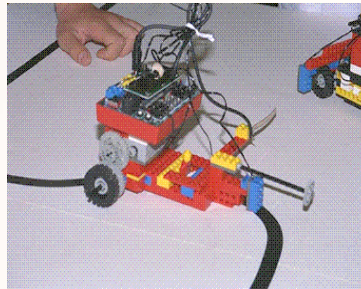
Constructionist Online Communities

- People creating something together online
  - Stricter sense: artifact
  - Looser sense: shared understanding
- Community provides both motivation and support:
  - Technical support
  - Emotional support
  - Role models
  - An appreciative audience

What is a "Community" Anyway?

- What is a "community" has always been hotly debated (Schnore 1967)
- Cognitive science can help
  - Community is a category
  - Prototype based

Bruckman, "A New Perspective on 'Community' and its Implications for Computer-Mediated Communication Systems", CHI 2006 WIP

Georgia Tech | College of Computing

Rosch's Prototype-Based Theory of Categories

- Categories have a set of "best members" (prototypes)

- Members of a category often have degrees of membership

  - Example: a robin is a better example of a bird than an emu or penguin

- Categories can have fuzzy boundaries

"Community" as a Prototype-Based Category

- What are our prototypes for "community"?
  - Idealized 1950s America that never existed
- "Pundits worry that virtual community may not truly be community. These worriers are confusing the pastoralist myth of community for reality. Community ties are already geographically dispersed, sparsely knit, connected heavily by telecommunications (phone and fax), and specialized in content." (Wellman & Gulia, 1999)
  - In other words, our common prototypes are idealized

Georgia Tech | College of Computing

New Salient Questions

- ## What are our prototypes?
  - Possible prototypes for a learning community:
    - Traditional schooling
    - Scouting
    - Samba schools
    - Tailors in West Africa

- ## What are the key characteristics of those prototypes?
  - How can we learn from them?
  - What features of each should we keep?

Georgia Tech | College of Computing

Understanding Wikipedia

- How many people have ever used the Wikipedia?

- How many people have ever edited the Wikipedia?

  – Did you learn something while you were doing it?

- "The problem with Wikipedia is that it only works in practice. In theory, it can never work." (*New York Times*, 4/23/07)

Georgia Tech | College of Computing

What Makes Wikis Unique?

- Fits constructionist paradigm:
  - Low barrier to entry
  - Easy learning curve
  - No ceiling
- Extremely light weight
  - Small differences in accessibility change user behavior
    - Example: salary database
- Collaboration on a large-scale
  - Doesn't work with 14 people in "The Bakeoff" (Gladwell 05)
- Open Source and Open Content are different
  - Open source has more centralized authority

Georgia Tech | College of Computing

Becoming Wikipedian: Transformation of Participation

- Interviews with 21 "Wikipedians"

- Becoming a part of Wikipedia is a process of:

  - Legitimate peripheral participation (Lave & Wenger), in a

  - Knowledge-building community (Scardamalia & Bereiter)

- Andrea Forte, Susan Bryant et. al. (Group 2005)
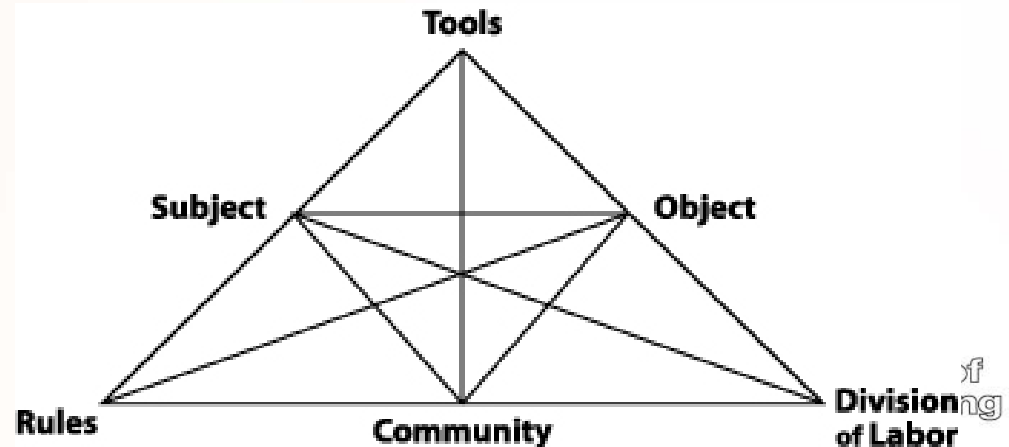
Power and Authority on Wikipedia

- It is NOT a free for all
  - How it really works matters
- Interview study with 11 people in administrative roles on Wikipedia
  - Nature of power and authority
  - How conflicts are resolved

Brett Favre: Profootballreference.com lists Farve as having 8224 career passing attempts, while the official Packers website and NFL.com list him as having 8223. An edit war ensues over the 1 attempt leading to an editor getting indefinitely banned. Sockpuppeting followed, including "aging" accounts to circumvent semi-protection. All over 1 passing attempt... In a 16 year hall of fame career. His name is still spelled weird.

Forte & Bruckman "Scaling Consensus: Increasing Decentralization in Wikipedia Governance" (HICSS 2008)

Georgia Tech | College of Computing

Increasing Decentralization

- Policy
  - Creation
    - Main policy creation slowing
    - Moving into WikiProjects
  - Interpretation & enforcement
    - Jimmy --> ArbCom --> Admins
    - ~1300 admins
    - Complicated process
      - Example: British climatologist William Connelley
        » Broke rules
        » Penalty from ArbCom: limited to one revert per day
        » Penalty not enforced by Admins

- Becoming an Admin
  - Differs by language
  - Criteria getting harder
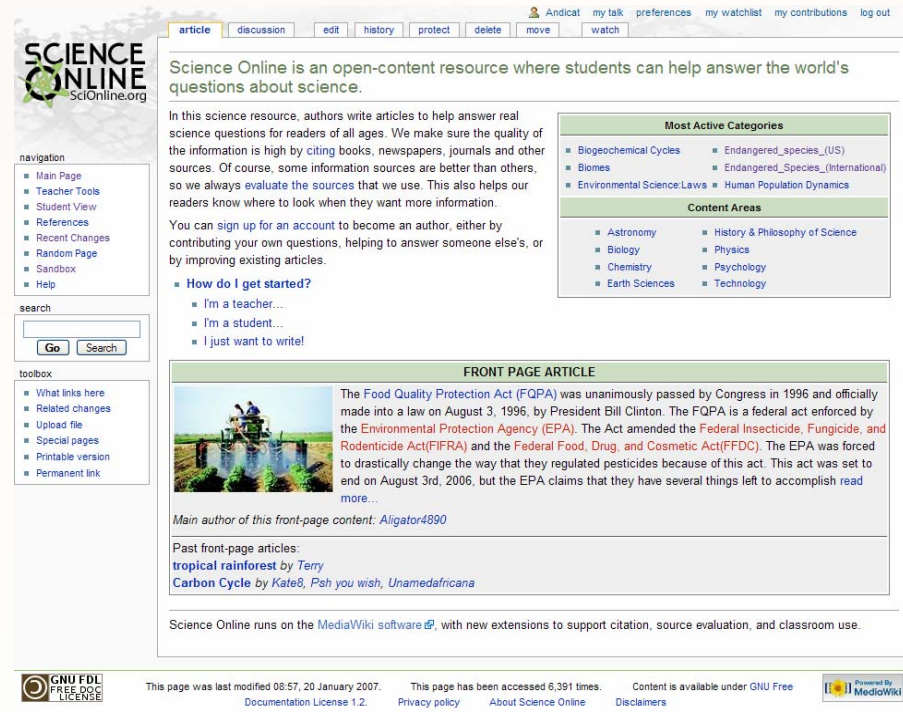
Decentralization: WikiProjects

- Allow local groups to establish editorial guidelines
    - Example: WikiProject Medicine
        - "Medical Collaboration of the Week"
- Function as small group
- Challenge: lack local enforcement mechanisms

- Decentralization happening as a necessity of scale

Georgia Tech | College of Computing

Science Online: Motivation

- What if we created a version of Wikipedia written by high-school students?
  – Focus is on science
- PhD work of Andrea Forte
  – MLIS UT Austin '98

Wiki as a Construction Kit

- Constructing text is a powerful learning activity
  - Writing-to-learn
    (Scardamalia and Bereiter, Emig, Britton, etc.)
- We can design environments that support specific writing activities
- Design challenges
  - Support critical citation – media literacy skills
  - Make it fit in the classroom

Science Online: Pilot Study

- Pilot study (spring 2005)
  - Students in an American government class
  - Write a position paper on a current issue on a wiki
  - Comment on other people's contributions
- Findings:
  - Instructor sees improvement in writing (subjective)
  - Sense of audience motivates students writing
    - Students don't realize work is world readable
    - But awareness of audience of their peers is enough
      - Try to convince students who might not agree with them
        » Student writing about Title IX wants to make sure "the guys" don't dismiss this in the first two sentences
        » Student citing CNN realizes this is perceived as a liberal source by some

Year-Long Classroom Study

- September 2006–May 2007
- Two high-school AP environmental science classes
- Software improvements, based on pilot study findings:
  - Move to MediaWiki software
  - Add support for citations
  - Support for finding one another's work
  - Teacher tools

Classroom Appropriation

- Participants
  - Teacher: former scientist (ABD), excited about wiki
  - 19 students
    - Juniors and seniors
    - AP students, not science fanatics
    - 11 female/8 male
- Data Collection
  - ~50 observation days
  - Interviews
  - Pre/post test
  - Wiki artifacts

Georgia Tech | College of Computing

Wiki Activity

Names*	Avg. Edits per Week**	Total Edits	Unique Pages Edited***
Reagan	5	97	18
John	6	36	6
Sylvia	7	67	9
Heather	7	221	31
Paige	9	111	13
Ella	13	116	9
Amanda	13	231	18
Jill	14	382	27
Carrie	16	204	13
Ed	17	271	16
Carl	18	72	4
David	18	256	14
Anne	19	194	10
Brian	23	180	8
Larry	26	153	6
Gary	29	234	8
Kelly	30	384	13
Alex	45	404	9
Susan	49	779	16
Avg.	19	231	13
St. Dev.	12	171	7

* All names have been changed.

** Avg edits per week is used to control for the fact that some students participated for 10 weeks, some 17, and some 27.

*** Number includes articles, userpages and talk pages.

- 7 wiki assignments
- Wide range of engagement
- Case: Ed
  – Average in terms of wiki engagement
  – Above average student

Constructionist Learning at Work

-"NAFTA Coalition" mindprod.com/money/nafta.html

-"Genetically Engineered Joke" www.non-gm-farmers.com/news_details.asp?ID=2361

-"Dolphin, The Issue of the Case" http://www.hitech-dolphin.com/bottlenose-dolphin-pictures-2.html

References   By Author   By Reference Type

Book References
Balassa, Bela (1991). *Comparative Advantage, Trade Policy and Economic Development*. New York: New York University Press.

Cunningham, William and Barbara Woodworth Saigo (2001). *Environmental Science A Global Concern 6th Edition*. New York, NY: McGraw-Hill.

McNeil, Robert (2003). *The Human Web*. W. W. Norton and Co..

Journal References
Ivanova, Antonina and Manuel Angeles (2006). "Trade and environment issues in APEC" *The Social Science Journal*. pp.629-642. (43)1.

Logan, Marty (2007). "North American Leaders Micro-Managing Environmental Deal" *Inter Press Service News Agency*.

MacKenzie, Debora (2004). "WTO Ruling May Spell End of Farmer's Subsidies" *New Scientist*.

Wendland, Joel (2005). "CAFTA: Free Trade vs. Jobs, Environment, and Health" *Political Affairs Magazine*.

Web References
Global Exchange. *Free Trade, the Environment, and Biotech*.

Global Exchange. *Free Trade and the Environment*.

Globalization 101. *Environment and Globalization: Introduction*.

Globalization 101. *The Shrimp-Turtle Case in the New WTO Context*.

Categories: Environmental Issues | Environmental Issues Project (100)

Georgia Tech | College of Computing

The Good News

- Constructionism predicts precisely the kind of engagement and learning that we saw when students wrote on the wiki
  - The creation and sharing of a personally meaningful public artifact as a learning activity leads to deep engagement

The Bad News

- Students learn too deeply
- Students can write collaboratively
- Students can share their work

Georgia Tech | College of Computing

Learning "Too Deeply"?

"They're student-made so there could be a lot more information on the wiki than we actually need to know for the test." – Sylvia

"The level of thinking that I guess I had them do and work on some of those is probably deeper than the curriculum requires for the assessment." – Mr. Grant

Collaboration & Assessment

- Students can write collaboratively
  - Early on:

  "I can just go back and document how little or how much that person contributed." – Mr. Grant

  - End of term:

  "There's a problem with collaboration and then assigning grades… when you come back to the tried and true method of doing things you don't have to worry about all that… one person one grade." – Mr. Grant

- Seven wiki assignments
  - First major assignment was collaborative
  - Mr. Grant changes the rest to be individual

Georgia Tech | College of Computing

Sharing or Cheating?

- Students can share their work
  - Teacher repeatedly suggested features to allow for the option of homework turn-in instead of publication
- Need to think more critically about collaboration in traditional academic settings
  - Avoid the 'paste together and correct the font' problem

Georgia Tech | College of Computing

Proposed Solutions

- Better visualization tools for teachers to assess student contributions
    - We could do more to support existing practices
        - To what extent do we want to?
        - Is there a danger of sublimating the potential of the medium? (Wiki for test preparation.)

- Radical educational reform

GameLog &
The Game Ontology Project

- ## What is game studies?
  - Classes showing up at more and more schools
    - How do we teach this new field?
- ## Two tools:
  - GameLog
  - The Game Ontology
- ## Trials:
  - Fall 2006:
    - Undergrad lecture class, 24 (36) students
    - Mixed grad/undergrad discussion class, 11 (25) students
  - Spring 2007
    - Undergrad lecture class 81 (213) students
- ## PhD work of José Zagal

GameLog

- Online blogging tool for games
  - www.gamelog.cl
  - Multiple, parallel blogs (one per game)
- Helps students:
  - Reflect on their gameplay experience
  - Connect game elements across multiple games
  - Gain insight on how the experience of playing a game changes over time
  - Students find they start noticing things, start playing differently
  - Comment on one another's entries
- Promotes reflection and meta-cognition

The Game Ontology

- Classification of structural elements of games, and their inter-relationships
- Wiki-based
  - Contribute examples
  - Contribute new terms
  - Anyone can contribute
  - Knowledge is always evolving
  - Of use to scholars
- Helps students:
  - Use their experience and knowledge to meaningfully contribute to an ongoing games research project
  - Learn and create vocabulary and concepts for understanding games
- Ontology created by Michael Matteas

Georgia Tech | College of Computing

Splat!

- GA Computes!
  - NSF Broadening Participation in Computing Alliance led by Mark Guzdial
  - Online community component
- Strategy: meet teens where they already are
- Facebook application
  - Share completed media projects from other sites
    - Flash, Scratch, etc.
  - Peer rating system
  - Contests
    - "How I did it" profiles
- PhD work of Sarita Yardi

Georgia Tech | College of Computing

Computer-Supported Collaborative Innovation (CSCI)

- Open-source and open-content work best with a well-defined goal
  - Example: porting UNIX to the PC
  - Example: creating an encyclopedia
- Could these approaches work with a more open-ended goal?
  - Clarifying the goal state is part of the task
- Pilot work: study of animation online
  - How do people collaborate to create animations?
    - Especially on newgrounds.com
    - Found four different collaborative modes:
      - Contest, collection, continuation, collaboration
- Studying how to support new forms of creative collaboration online
- PhD work of Kurt Luther

Dusting Off My Crystal Ball…

- Amateurs as first-class participants
  - This doesn't eliminate professional content
  - But it raises the bar
- Democratization of content creation
  - Richer variety of views
  - Not filtered by what "sells" or what is "acceptable"
- Shifting privacy awareness & norms
- Business models as a driving force
- Growing online/offline integration
  - The compelling "3D world" is the real one!

Georgia Tech | College of Computing

Designing for Educational Opportunity

- The Internet is a natural fit for constructionist learning
  - Support for learning
  - Audience for completed work
- Making this work in real educational settings is a challenge

Georgia Tech | College of Computing

Acknowledgments

- ELC students:
  - Betsy DiSalvo, Andrea Forte, Kurt Luther, Sarita Yardi, José Zagal
- ELC sponsors:
  - IBM, Intel, Microsoft, Pitney Bowes, Ricoh
  - The National Science Foundation
  - US Department of Education
- For more info:
  - http://www.cc.gatech.edu/elc/