



# **From Openness to Transparency: The Role of Social Media in Open Source Ecosystems**

Jim Herbsleb

School of Computer Science

[jdh@cs.cmu.edu](mailto:jdh@cs.cmu.edu)

# Agenda

- Why are ecosystems becoming so important?
- What about Conway's Law?
- A role for social media?
- Takeaways

# Socio-Technical Ecosystem

- “Open” Platform
  - Common functionality everyone needs
  - Different forms: standards to implementations
- Innovative engineering on top
  - Diversity of contributors
- Co-opetition
  - Ecosystems compete
  - Participants in an ecosystem mostly complement

# Ecosystem Design Parameters

- Platform architecture
- Governance
- Incentives
- Collaborative infrastructure

# Why Ecosystems? The Problem

- Convergence of several trends:
  - Everything is connected to everything else
  - Greater and greater need for specialized knowledge
  - Functionality is moving to software

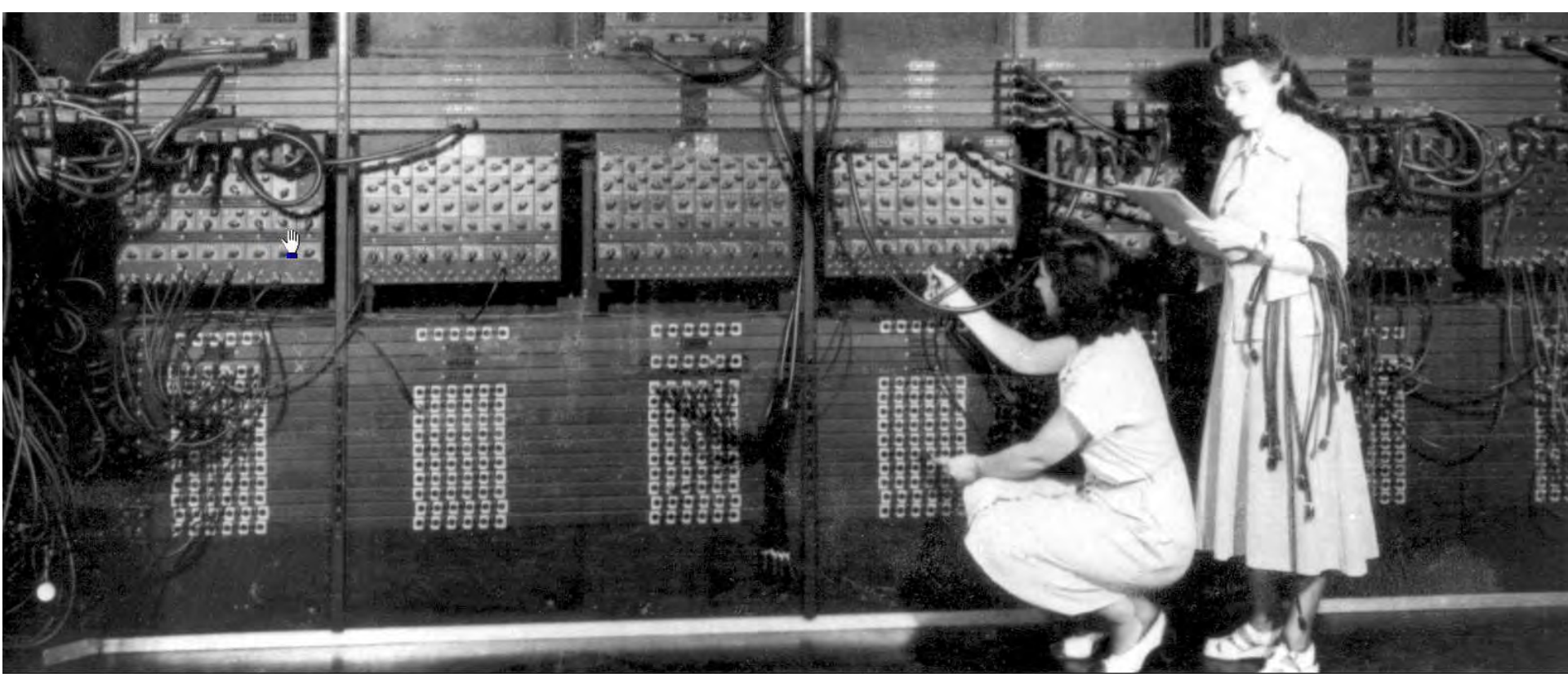


Marc Andreessen, *Wall Street Journal*, August 20, 2011:

**SOFTWARE IS EATING THE  
WORLD!**

→ Ecosystems important?

# Individual Computations



→ Ecosystems important?

# Individual Work Applications



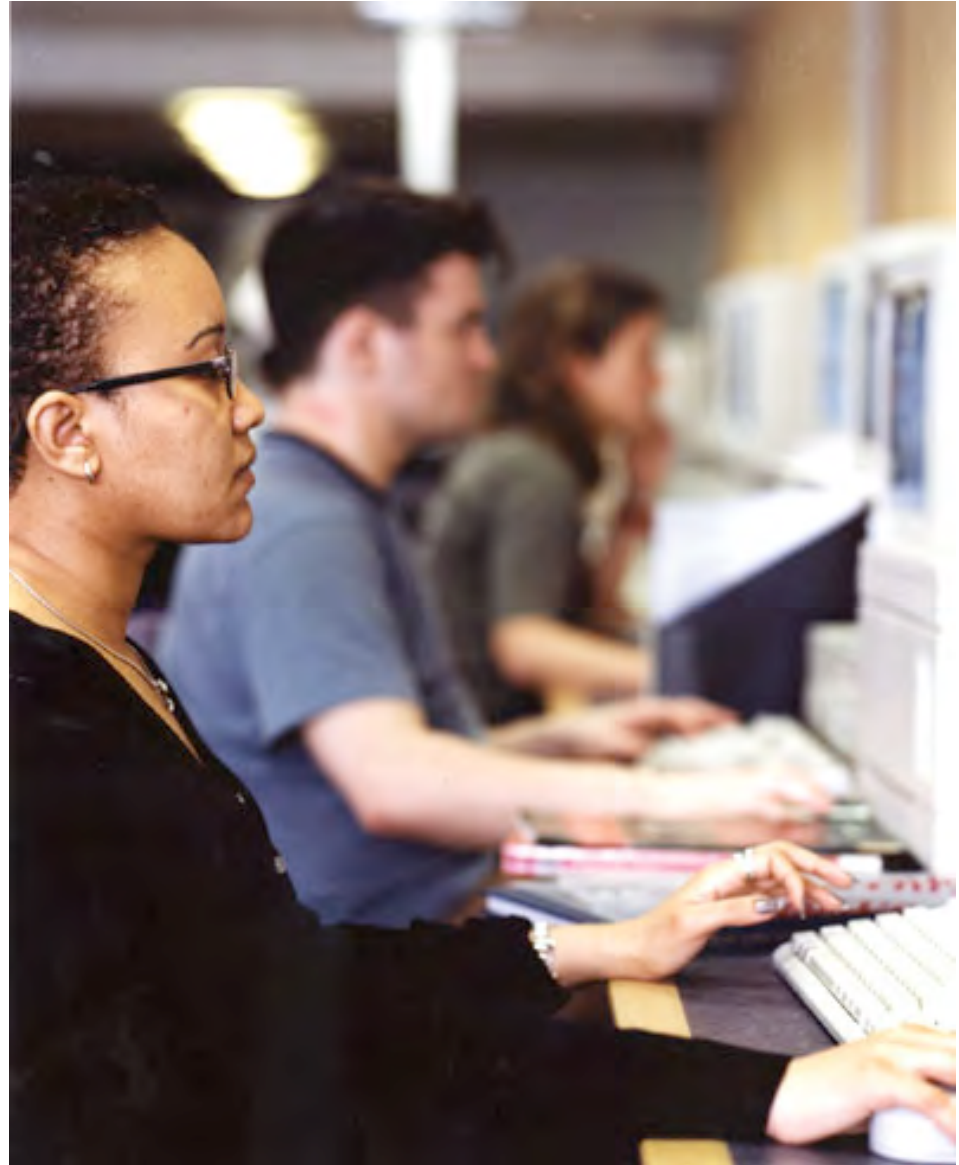


→ Ecosystems important?

Networks

Communication

Group  
applications



→ Ecosystems important?



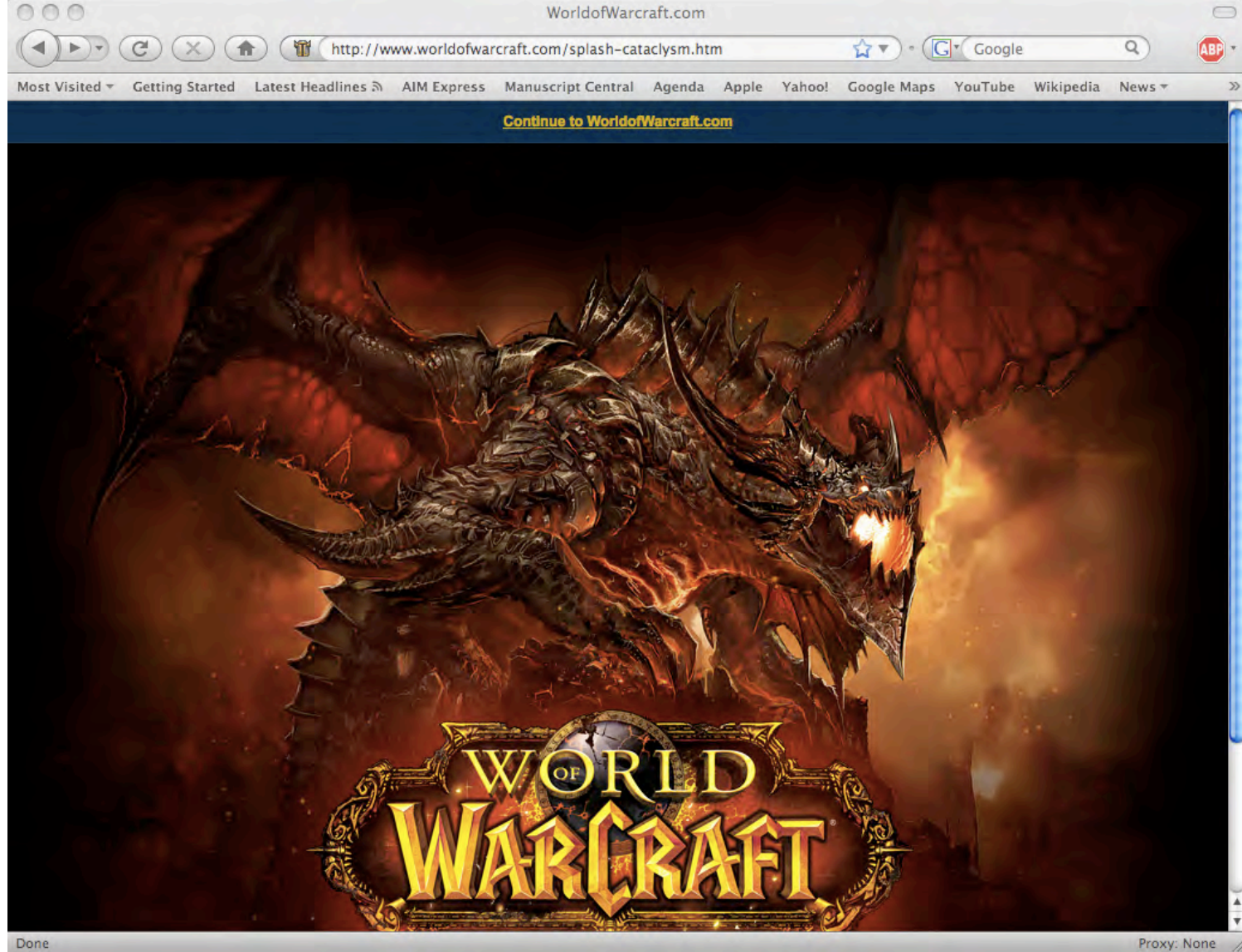
Talking to friends

Texting

Sharing photos

Listening to music

On and on . . .



Team-oriented collaborative  
virtual mass murder



# Smart Cars







# Smarter Cars

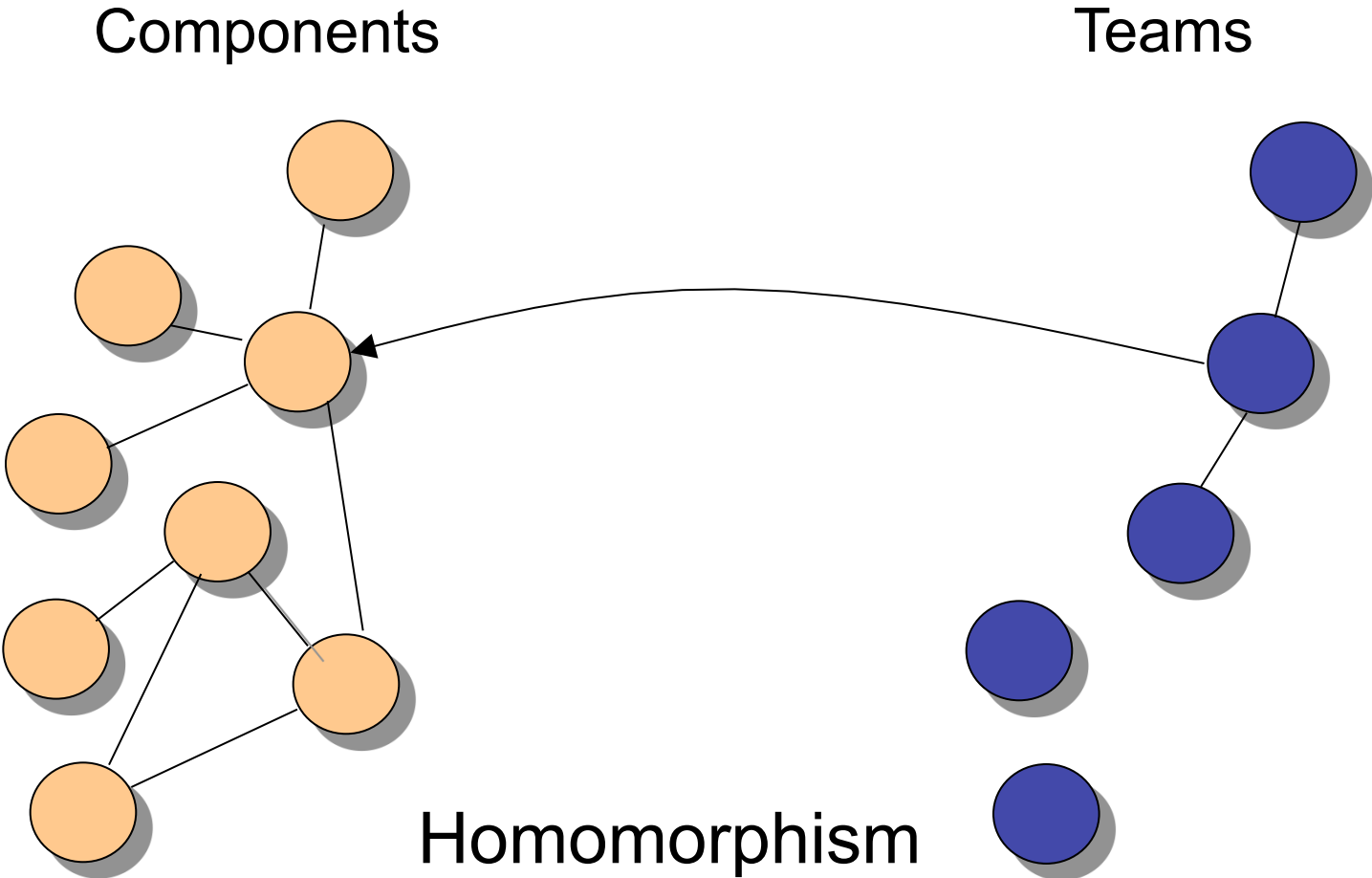
# Why Ecosystems? A Solution

- Ecosystems resolve the tension between accelerating specialization and interconnection
- Support dynamism and evolution
  - Gradually evolving platforms
  - Rapid creation and evolution of niche technologies
  - Keep pace with technology/user co-evolution

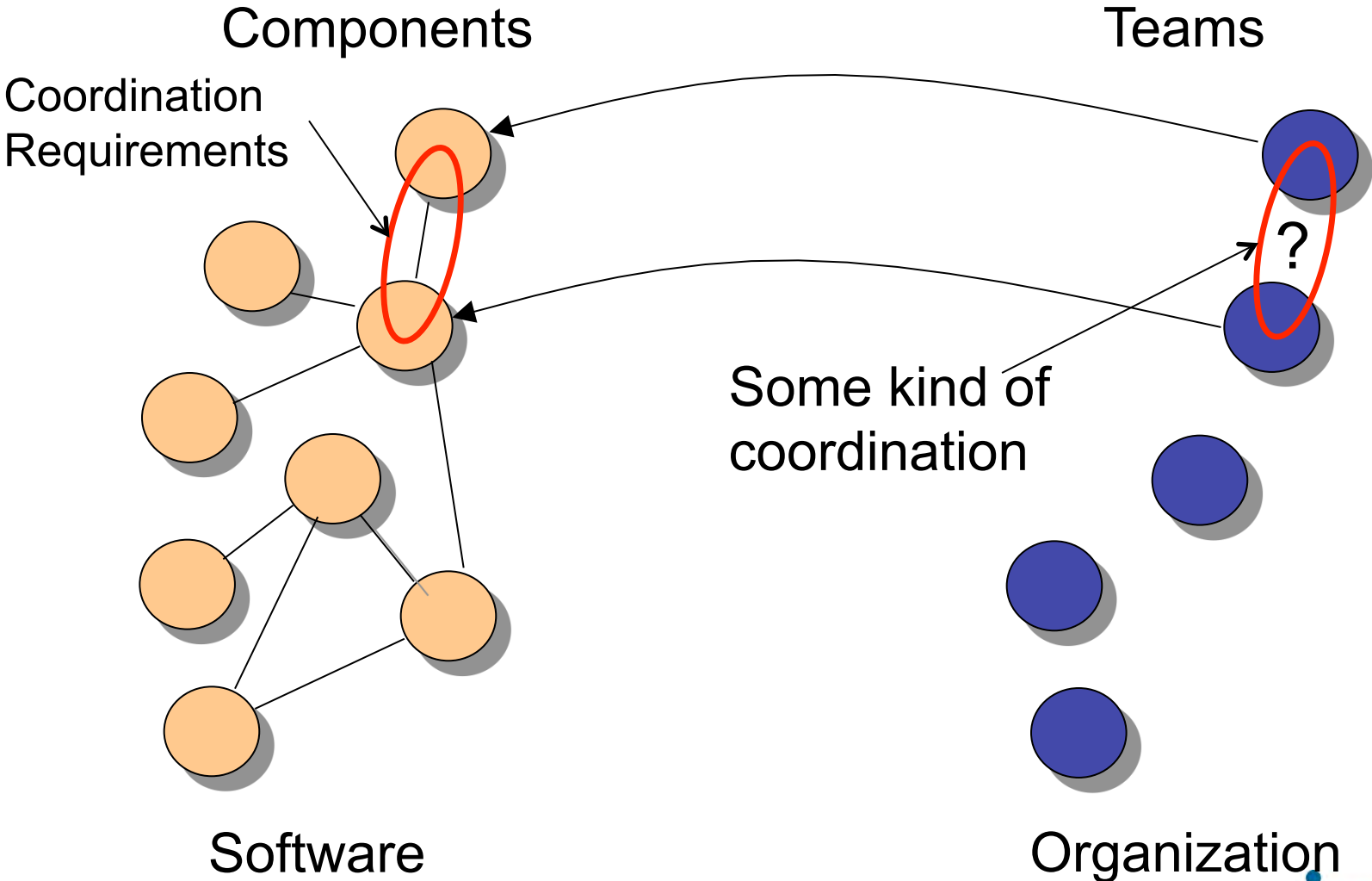
# Conway's Law

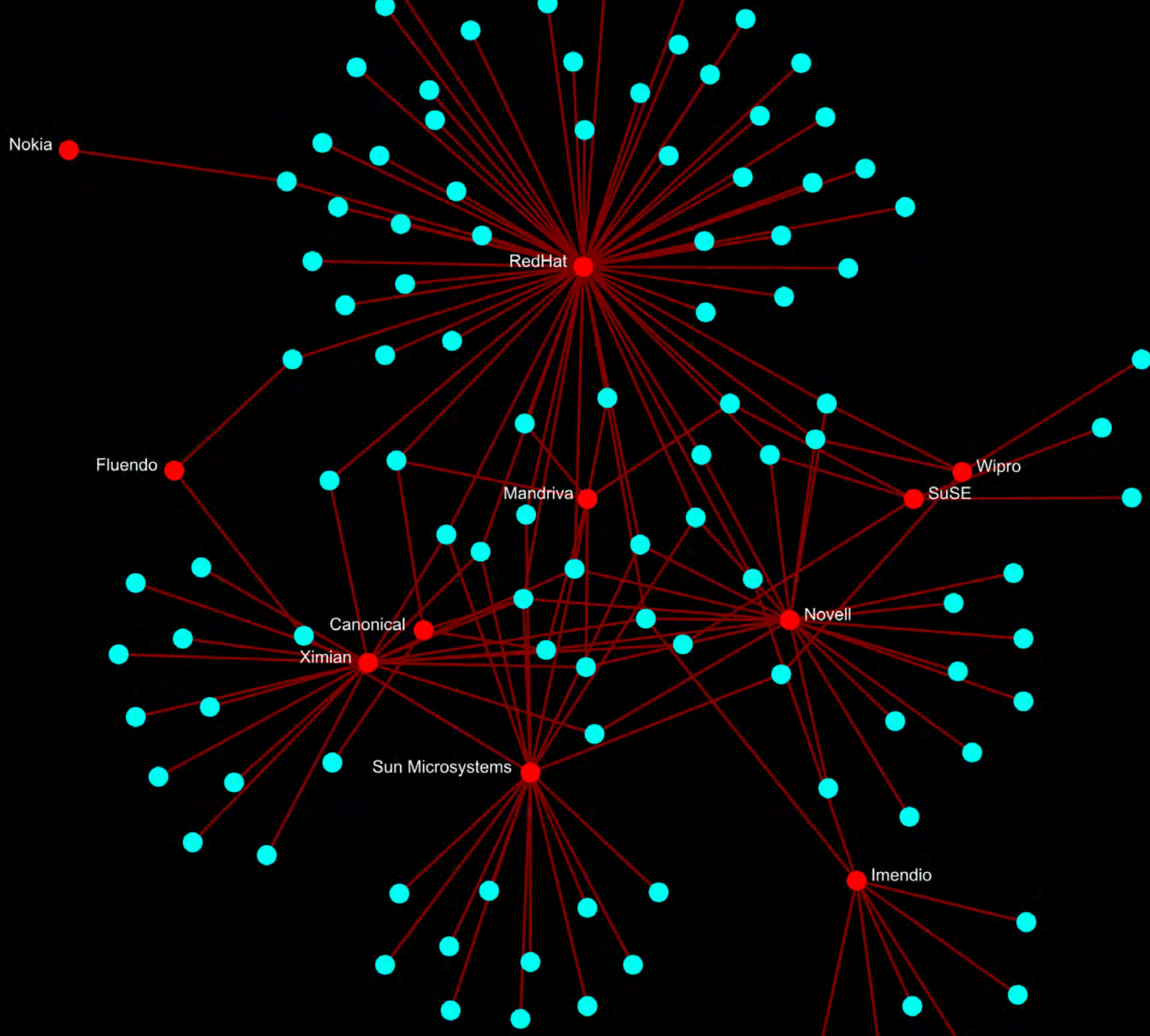
- “Any organization that designs a system will inevitably produce a design whose structure is a copy of the organization's communication structure.”\*

\*M.E. Conway, “How Do Committees Invent?” *Datamation*, 14 (4), 1968, pp. 28–31.









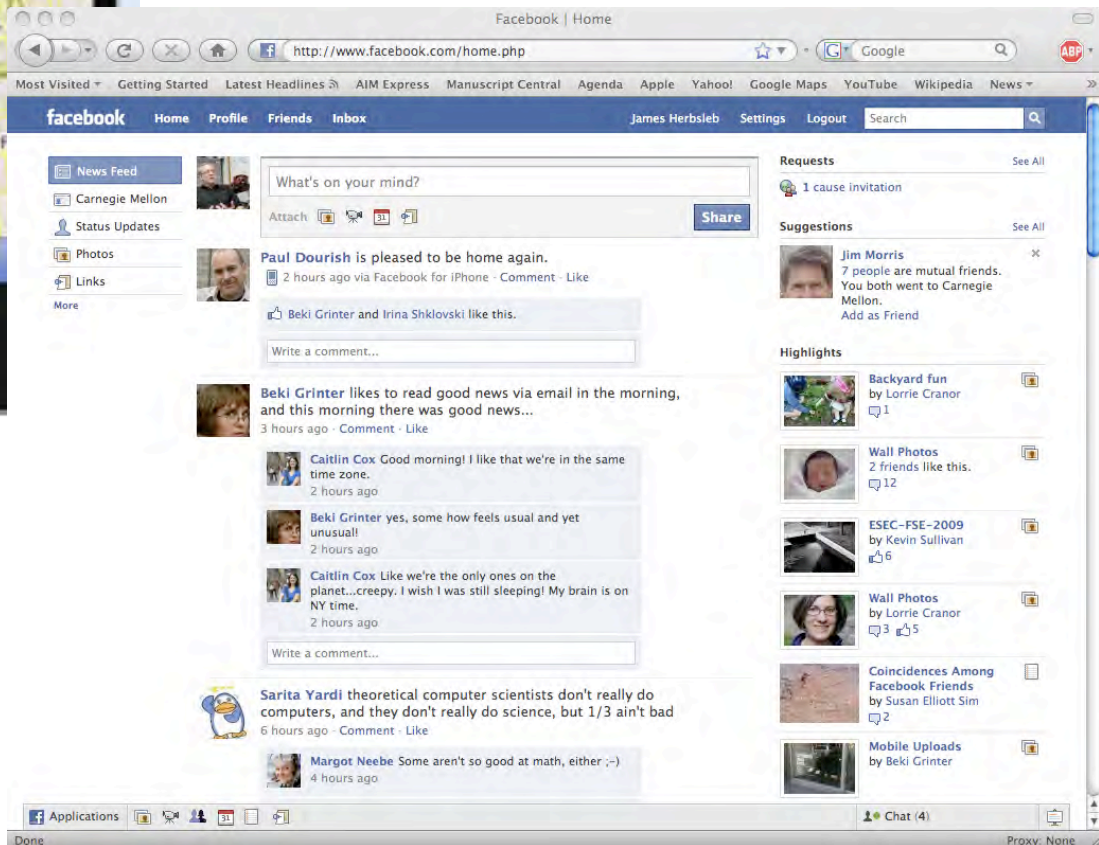


facebook

December 2010



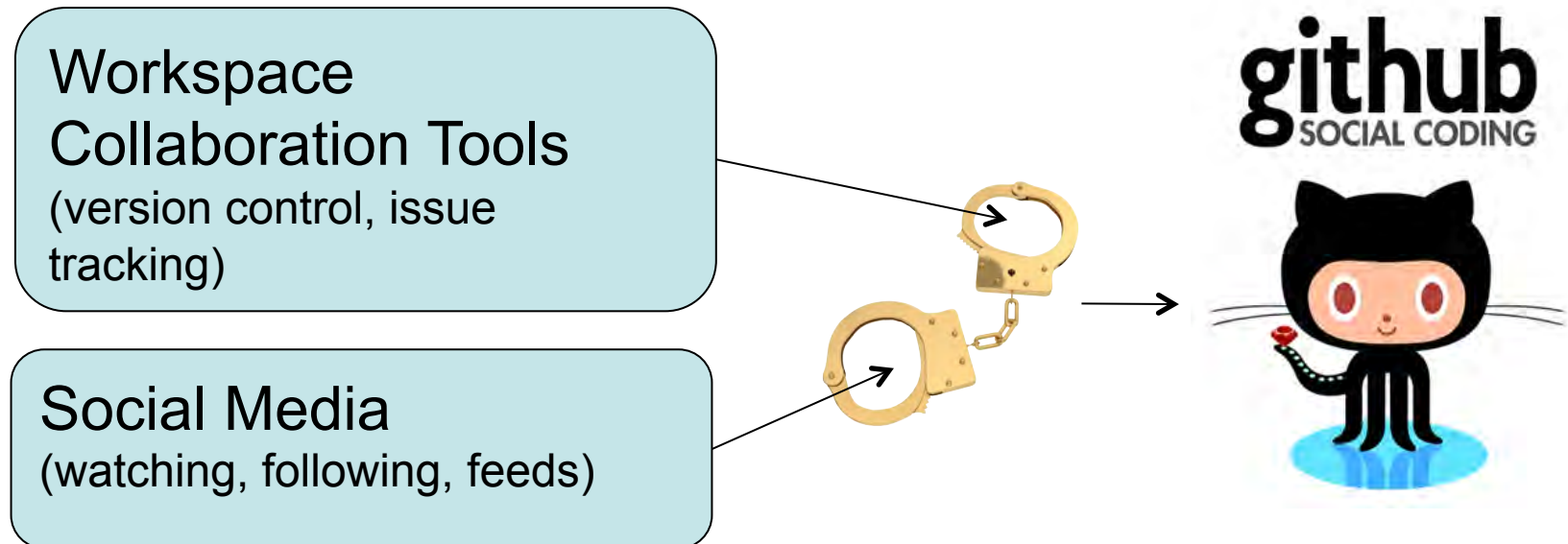
# Social Media



# Social Media Explained . . .

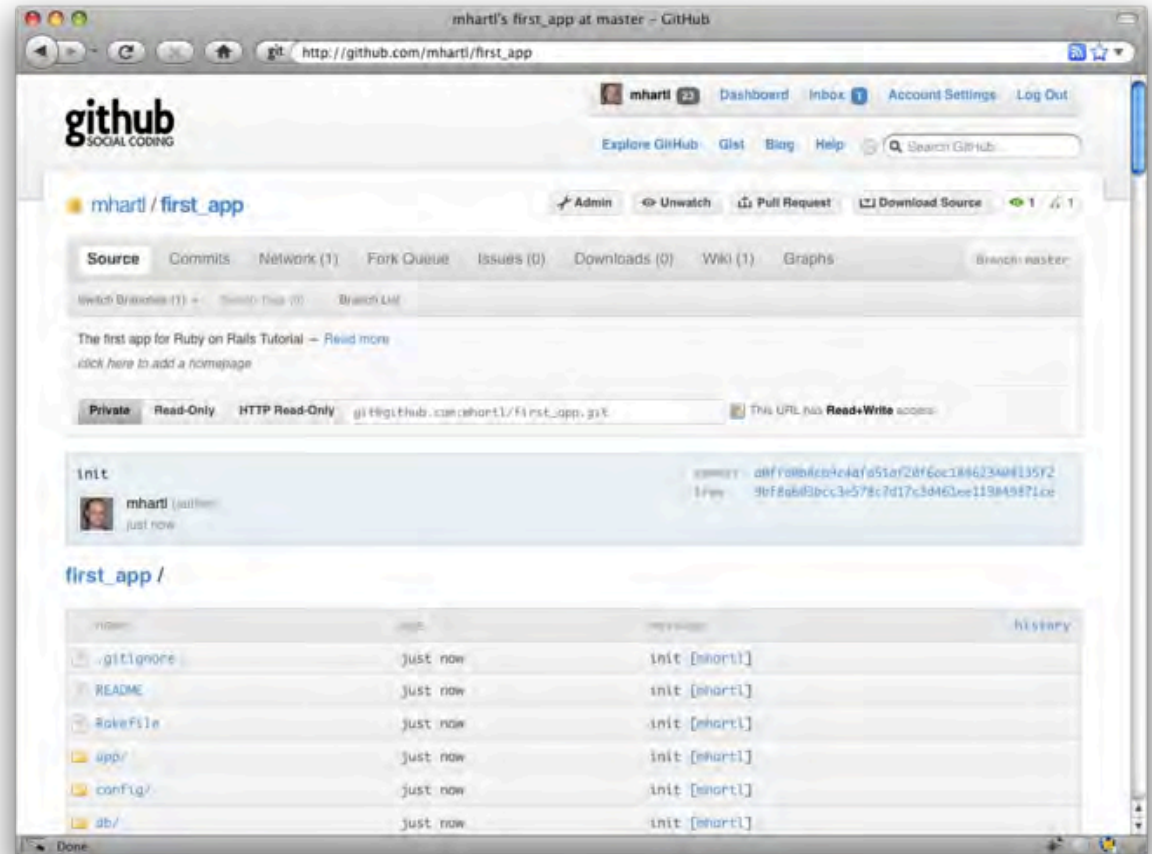


# Social Coding



➡ Social Media?

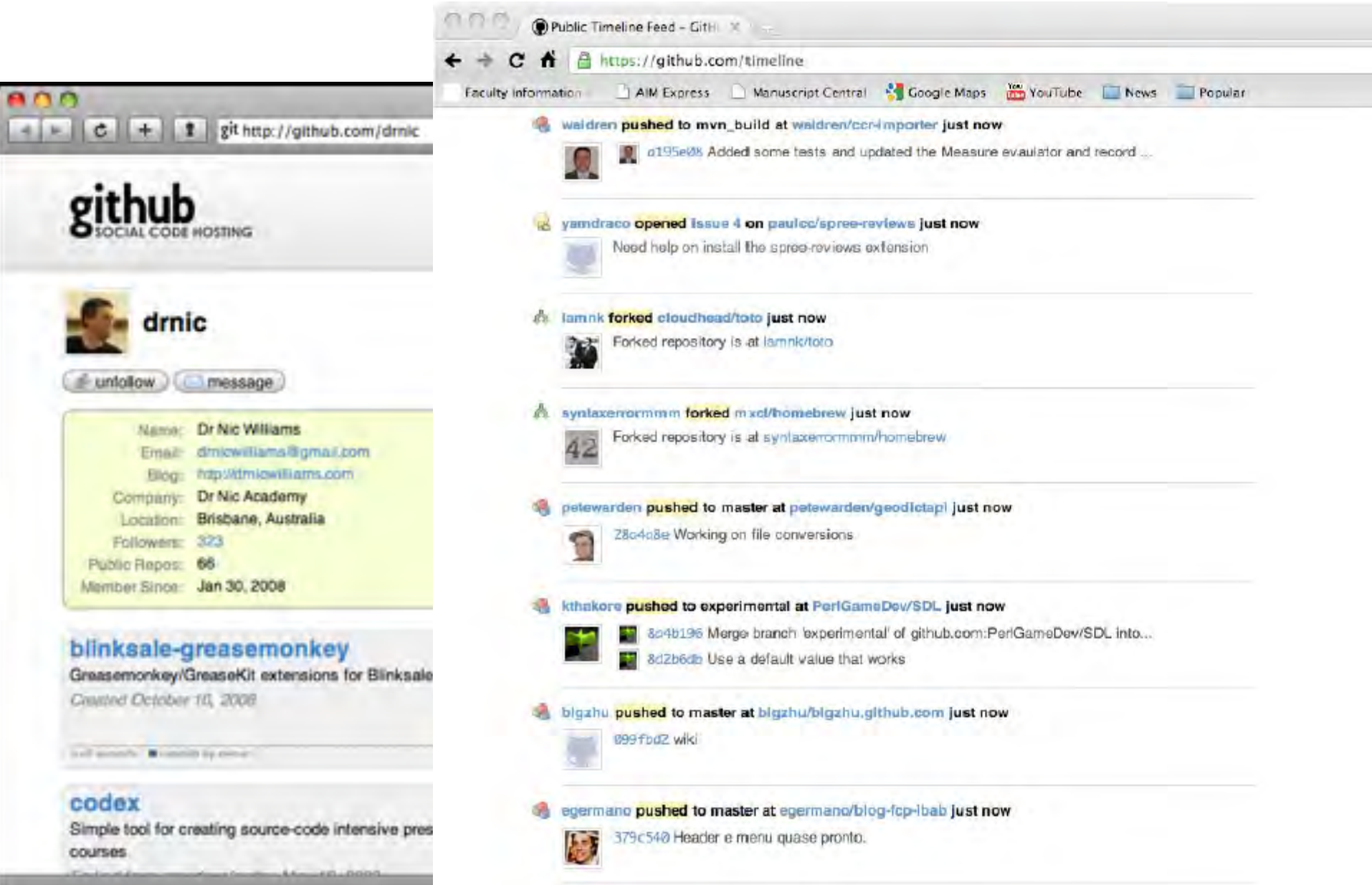
**github**  
SOCIAL CODING



- Over one million public code repositories
- 340,000 registered contributors
- 80,000 code commits per day



# Users + Code + Actions on Code



The image is a screenshot of the GitHub website, split into two main sections. The left section shows a user profile for 'drnic' (Dr Nic Williams), and the right section shows a 'Public Timeline Feed' with various user actions.

**User Profile (drnic):**

- Name:** Dr Nic Williams
- Email:** drnicwilliams@gmail.com
- Blog:** http://drnicwilliams.com
- Company:** Dr Nic Academy
- Location:** Brisbane, Australia
- Followers:** 323
- Public Repos:** 66
- Member Since:** Jan 30, 2008

**Repositories:**

- blinksale-greasemonkey:** Greasemonkey/GreaseKit extensions for Blinksale. Created October 10, 2008.
- codex:** Simple tool for creating source-code intensive pres courses.

**Public Timeline Feed - GitHub:**

- waldren pushed to mvn\_build at waldren/ccr-importer just now**  
a195e08 Added some tests and updated the Measure evaluator and record ...
- yamdraco opened Issue 4 on paulco/spree-reviews just now**  
Need help on install the spree-reviews extension
- lamnk forked cloudhead/toto just now**  
Forked repository is at lamnk/toto
- synlaxerormmm forked mxd/homebrew just now**  
Forked repository is at synlaxerormmm/homebrew
- petewarden pushed to master at petewarden/geodictapi just now**  
28c4c8e Working on file conversions
- kthakore pushed to experimental at PerlGameDev/SDL just now**  
8c4b196 Merge branch 'experimental' of github.com:PerlGameDev/SDL into...  
8d2b6db Use a default value that works
- bigzhu pushed to master at bigzhu/bigzhu.github.com just now**  
099fdb2 wiki
- egermano pushed to master at egermano/blog-fcp-lbab just now**  
379c540 Header e menu quase pronto.



# The Promise of Social Media

- Social applications let users
  - Articulate interest network of people and artifacts
  - Track and follow the activities of these people or actions on these artifacts
- Affords unprecedented level of transparency
  - Visibility of others' actions on artifacts
  - Visibility of others' interactions

# Qualitative Study of Social Coding\*

- Goals
  - Understand how users interpret and use the social functionality on GitHub.
  - Drive design of awareness/collaboration tools
- Semi-structured interviews with 24 users
- Focused on typical site usage, project management, social functionality
- Looked for similarities across the nature of inferences they made based on visible information

\*Dabbish, L., Stuart, C., Tsay, J. and Herbsleb, J. (2012). Social Coding in GitHub: Transparency and Collaboration in an Open Software Repository. To appear in *Proceedings CSCW 2012*.

# What Does GitHub Expose (1)?

- People
  - Profiles
  - Gravatar
  - Contact info
  - Repositories
  - Latest activities
- Code artifacts
  - Project page
  - Source code
  - Commit history
  - Issues
  - Comments
  - Permanent URL at line level

# What Does GitHub Expose (2)?

- Actions on code
  - Commits
  - Forking
  - Pull request
  - Comment
- Subscription actions
  - Following
  - Watching
- Visualizations
  - Network view
  - Compare view

# Network View

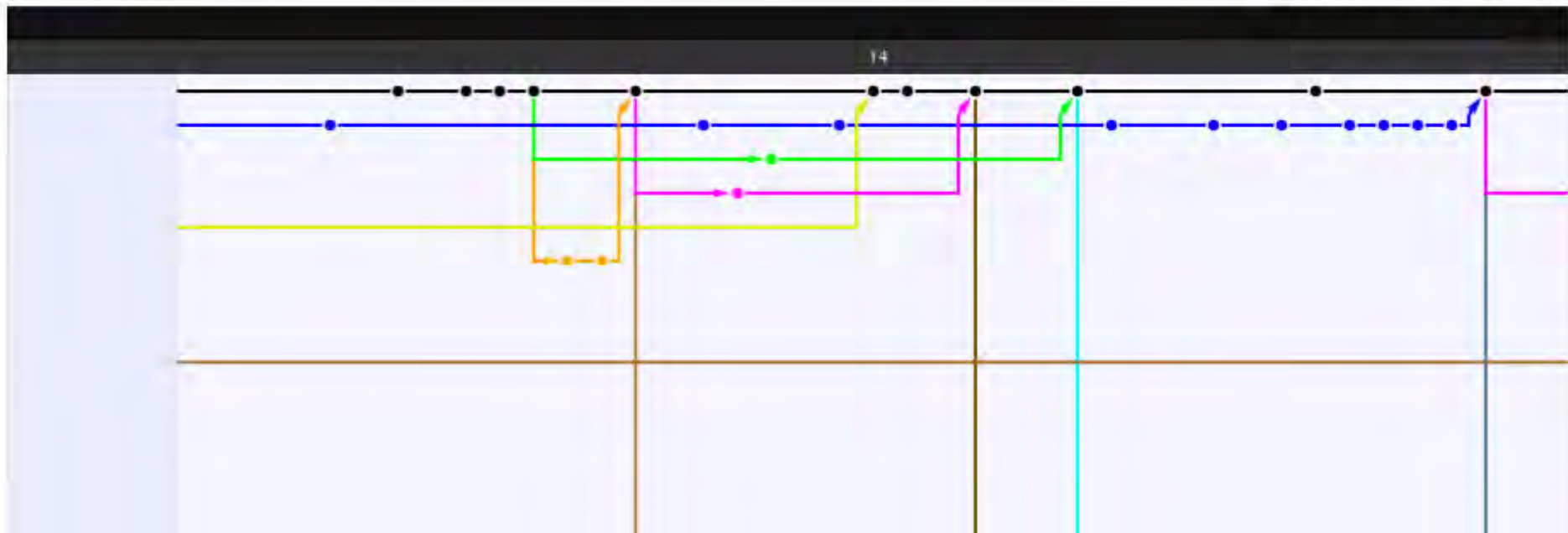
## The rails network graph

Keyboard shortcuts available 

All branches in the network using rails/rails as the reference point. [Read our blog post about how it works.](#)

[Show Help](#)

Last updated: about 5 hours ago



# Inferences:

Recency, volume, and location of actions

## Projects

- Is this project alive?\*
- How much does anyone care about it?
- How well is it managed and maintained?
  - Lots of open pull requests?

## People

- How committed is this developer to this project?
- What is this developer interested in?

\*“Commit activity in the feeds shows that the project is alive, that people are still adding code.”

# Inferences:

Sequence of actions conveys meaning

- History of activity signals developer intention\*
- History of activity signals competence
- History of activity signals project structure and roles

\*“*Your commits tell a story.*”

# Inferences:

Attention signals community support

- Attention signals action or artifact importance\*
- Attention signals developer status
- Attention signals project quality

*\*“The way you know how useful something is, is how much community there is behind it.”*



# Social Inferences Inform Joint Action

- Recruiting developers
- Identifying user needs\*
- Managing incoming code contributions
- Managing dependencies with other projects

*\*“I saw somebody trying to use it with Rails master I'm like well crap I don't know if it works with Rails master so let me check. So that type of stuff has been useful just to get a sense of the kinds of things people might like to see, you know?”*

# Learning from Others

- Following rockstars
- Watching watching\*
- Identifying new technical knowledge
- Direct feedback

*\*“This guy has good taste in projects. He curates for me. Watching him is like watching the best of objective C that GitHub has to offer.”*

# Managing Reputation and Status

- Self-promotion
- Social capital, identity, and recognition
- Being onstage\*

*\*“I try and make sure my commit messages are snappy and my code is clean because I know that a lot of people are watching. ...It’s like being on stage, you don’t want to mess up, you’re giving it your best . . .”*

# Takeaways

- Ecosystems are important because they solve an urgent problem
- Social media has some ability to address coordination at scale, across boundaries, in dynamic environments
- Some research issues:
  - More refined ways to push/pull information
  - Run-time socio-technical ecosystems
  - Technical architectures to support loose coupling in ecosystem contexts

# Questions?

We gratefully acknowledge support from the Future of Work Center, Heinz College, Carnegie Mellon University, and National Science Foundation, Award IIS 1111750.