Transparency in open collaboration environments

Laura Dabbish
Associate Professor
Human-Computer Interaction Institute & Heinz College
Carnegie Mellon University
“If there were such a thing as complete reciprocal transparency, the relationships of human beings to each other would be modified in a quite unimaginable fashion.”

Simmel, 1906
Dirk Stoop

It's a great day for flight SAN-SFO.

Joseph

Really, United! Good call, guy.

Liz Law

What on earth that ended at 6:30am flight that ended at 10pm? order of business Orleans tomorrow.

#Obama #2012 #Election #Vote

Delete this now
Transparency

Accurate observability of an organization’s low-level activities, routines, behaviors, output, and performance (Bernstein, 2012)

Can be thought of as an attribute of a work environment
Digital transparency

Users can follow the activities of other people and their actions on artifacts.
Social Media as Transparency Tool

Users articulate an interest network of people and artifacts

Receive updates on actions by those people or on those artifacts

Affords unprecedented level of transparency in form of visibility of actions by others
Transparency affords broader information access for both work process and product
Social Challenges to Transparency

- Multiple stakeholders with disparate interests and desires
- Social and political reasons to hide or share information in a conflict ridden environment
- Manipulate action to communicate certain information (signaling – Donath; impression management – Goffman)
Appropriating Transparency

How is the technology of transparency coopted to support the conflicting, emergent and dynamic needs of contributors in open collaboration settings?

– What practices emerge to increase or limit openness?

– How does the technology of transparency interact with a culture of openness?

– What social, political and cultural factors support cross-project coordination and collaboration in a transparent open collaboration environment?
What does coordination look like in a transparent open collaboration environment?
- 5.9 million+ public code repositories
- 3.4 million registered users
- 10,000 new users and 20,000 new repos a day
- 80,000+ code commits per day
# Contributions

## Popular repositories

<table>
<thead>
<tr>
<th>Repository</th>
<th>Stars</th>
</tr>
</thead>
<tbody>
<tr>
<td>rbenv</td>
<td>3,109</td>
</tr>
<tr>
<td>prototype</td>
<td>2,651</td>
</tr>
<tr>
<td>sprockets</td>
<td>1,824</td>
</tr>
<tr>
<td>eco</td>
<td>1,266</td>
</tr>
<tr>
<td>ruby-build</td>
<td>1,127</td>
</tr>
</tbody>
</table>

## Repositories contributed to

<table>
<thead>
<tr>
<th>Repository</th>
<th>Stars</th>
</tr>
</thead>
<tbody>
<tr>
<td>37signals/pow</td>
<td>2,115</td>
</tr>
<tr>
<td>prototypejs/prototypejs.github.com</td>
<td>6</td>
</tr>
<tr>
<td>documentcloud/underscore</td>
<td>7,684</td>
</tr>
<tr>
<td>rails/turbolinks</td>
<td>1,603</td>
</tr>
<tr>
<td>37signals/highrise-api</td>
<td>44</td>
</tr>
</tbody>
</table>

## sstephenson's Open Source Contributions

![Timeline of contributions](image-url)
<table>
<thead>
<tr>
<th>Repository Name</th>
<th>Language</th>
<th>Stars</th>
<th>Forks</th>
<th>Shelves</th>
</tr>
</thead>
<tbody>
<tr>
<td>rbenv</td>
<td>Shell</td>
<td>2,652</td>
<td>159</td>
<td></td>
</tr>
<tr>
<td>ruby-build</td>
<td>Shell</td>
<td>855</td>
<td>162</td>
<td></td>
</tr>
<tr>
<td>sprockets</td>
<td>Ruby</td>
<td>1,699</td>
<td>200</td>
<td></td>
</tr>
<tr>
<td>kinect</td>
<td>C</td>
<td>85</td>
<td>9</td>
<td></td>
</tr>
<tr>
<td>bats</td>
<td>Shell</td>
<td>129</td>
<td>8</td>
<td></td>
</tr>
<tr>
<td>ruby-build-download-mirror</td>
<td></td>
<td>1</td>
<td>0</td>
<td></td>
</tr>
<tr>
<td>prototype</td>
<td>JavaScript</td>
<td>2,541</td>
<td>327</td>
<td></td>
</tr>
<tr>
<td>stitch</td>
<td>CoffeeScript</td>
<td>607</td>
<td>86</td>
<td></td>
</tr>
</tbody>
</table>
Public Activity 📢

👩‍💻 sstephenson opened issue 1 on josh/nack 2 days ago
Fire an event when an app fails to boot

👩‍💻 sstephenson opened pull request 2 on skampler/ndns 2 days ago
Add package.json
1 commit with 14 additions and 0 deletions

👩‍💻 sstephenson pushed to master at sstephenson/ndns 2 days ago
9be5368 Add package.json

👩‍💻 sstephenson forked skampler/ndns 2 days ago
Forked repository is at sstephenson/ndns

👩‍💻 sstephenson created gist: 771090 3 days ago
$ sudo su - # mkdir /etc/resolver # cat > /etc/resolver/test

👩‍💻 sstephenson started watching jamis/csmazes 5 days ago
csmazes’s description:
Maze algorithms implemented in CoffeeScript, with an eye toward demonstrating how the algorithms work by animating them.
Network view

The stitch network graph

All branches in the network using sstephenson/stitch as the reference point. Read our blog post about how it works.
Qualitative Methodology

Semi-structured interviews w/ 50 GitHub members
- Peripheral and heavy users (project with > 80 watchers)
- OSS hobbyists and paid contributors

Focused on recent site usage, project management, and social activity

Open coded their responses to look at how transparent information influenced dependency management
Coordination in a Transparent Open Collaboration Environment

• Temporal process where projects are constantly in motion
• Information and accountability
  – Key decision points or major changes were immediately and broadly visible
  – Audience motivated expressive action
• Transparency allowed members to more readily view these changes and find right instance / moment to coordinate their work
Pre-coordination behaviors

• Lots of activity happens even before collaboration – and its precisely because of the potential for anyone to engage with you at any point
  – Anyone can join and leave at any time
  – Fluidity in open environment (Faraj & Majachzrak)

• Standardization & pre-coordination behaviors help manage fluidity
Pre-coordination

Technology
Public copies and edits
Feeds
Network view

Information seeking practices
- Scouting
- Monitoring
Scouting

Developers used aggregated traces of project activity to assess credibility and future potential.

“Are there changes going in? How recent are changes? Like, weeks, months, years? Another kind of metric is what’s the mix of [contributors]. Is it all one person? Is it all people in a single organization? Is it a wide group of people? The wider the net there’s more likelihood any issue that I might run into has already been identified.” (P22)
Monitoring

The stitch network graph
All branches in the network using sstephenson/stitch as the reference point. Read our blog post about how it works.
Assessing user needs

“I saw somebody trying to use [My Project] with the 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” (P9)
Watching dependencies

Developers watched dependent projects to notice potentially problematic changes in the feed.

“If there was something [in the feed] that would preclude a feature that I would want it would give me a chance to add input to it.” (P4)
Pre-coordination

Technology
Public copies and edits
Feeds
Network view

Information seeking practices
- Scouting
- Monitoring
Pre-coordination

Technology
- Public copies and edits
- Feeds
- Network view

Flexible change granularity
- Public commenting
- Audience feedback

Information seeking practices
- Scouting
- Monitoring

Expressive action practices
- Action Legibility
- Process Legibility
Expressive action and audience
‘Being onstage’

Commit level broadcasts create detailed process visibility
“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, you’ve got your Hollywood smile” (P4)
‘Being onstage’

Norm that commits should be granular and ‘tell a story’

Process visibility created desire to communicate why you were doing what you doing with the sequence of actions
Pre-coordination

Technology
- Public copies and edits
- Feeds
- Network view

Flexible change granularity
- Public commenting
- Audience feedback

Information seeking practices
- Scouting
- Monitoring

Expressive action practices
- Action Legibility
- Process Legibility
Common language of contribution

Fixed structure supported by the infrastructure
- How one changes the code
- Talks about the code
- Thinks about changing the code

Within that people do their work to make things understandable

You can see the physical structure, and the way people use it is to use the language provided
Practices around

Project management and participation
(Dabbish, Stuart, Tsay, & Herbsleb, 2012)

Change management and acceptance
(Tsay, Dabbish, & Herbsleb, 2012; 2014)

Impression formation
(Marlow & Dabbish, 2013; Marlow, Dabbish, & Herbsleb, 2013)
Transparency in open collaboration environments

Laura Dabbish
Associate Professor
Human-Computer Interaction Institute & Heinz College
Carnegie Mellon University

Collaboration platforms
Transparency in open collaboration environments

Laura Dabbish
Associate Professor
Human-Computer Interaction Institute & Heinz College
Carnegie Mellon University
Is peer production the future of work?

Collective vs Crowd
Alternative futures...
Post-currency collective

Open collaboration

Multiple individuals contributing effort towards a collective goal
Is peer production the future of work?

Collective vs Crowd
Computational organizations
Crowdsourcing: workers as computation devices
Crowdsourcing: workers as computation devices

- Humans training algorithms
- Contributors as computation devices carrying out tasks computers are learning to do
- Power centralized in platform that mediates connection between workers and requesters
- Little transparency of work assignment, evaluation, workflow
- Monitoring and surveillance
As soon as Winston had dealt with each of the messages, he clipped his speakwritten corrections to the appropriate copy of ‘The Times’ and pushed them into the pneumatic tube. Then, with a movement which was as nearly as possible unconscious, he crumpled up the original message and any notes that he himself had made, and dropped them into the memory hole to be devoured by the flames.
Crowdworkers get few cues about who else is around and what they are doing.
The crowdsourced workplace

<table>
<thead>
<tr>
<th>Requester</th>
<th>HIT Expiration Date</th>
<th>Time Allotted</th>
<th>Reward</th>
<th>HITs Available</th>
</tr>
</thead>
<tbody>
<tr>
<td>SebastianDerr</td>
<td>Aug 8, 2012 (4 weeks 1 day)</td>
<td>30 minutes</td>
<td>$0.02</td>
<td>46340</td>
</tr>
<tr>
<td>WSOYC.COM</td>
<td>Aug 8, 2012 (4 weeks 1 day)</td>
<td>2 hours</td>
<td>$0.00</td>
<td>23517</td>
</tr>
<tr>
<td>rohzi0d</td>
<td>Jul 30, 2012 (2 weeks 6 days)</td>
<td>48 minutes</td>
<td>$0.00</td>
<td>15517</td>
</tr>
<tr>
<td>CrowdSource</td>
<td>Jul 9, 2013 (52 weeks)</td>
<td>32 minutes</td>
<td>$0.16</td>
<td>11947</td>
</tr>
<tr>
<td>Tagasaurus</td>
<td>Aug 8, 2012 (4 weeks 1 day)</td>
<td>60 minutes</td>
<td>$0.02</td>
<td>9846</td>
</tr>
</tbody>
</table>

Find email for given person via web search (~60sec) approved regularly

3 questions about your city UNDER 230,000 population only = $0.17 bonus!!! - qualification instantly granted (no wait)
Tag 5 Images

Accurately Tag 5 Images with Main Subjects

Caption:

[Collapsed building.]
Research directions

- Scaling up the crowd [Kittur et al]
- Moving to more complex tasks
- Growing experts in the crowd
- Crowd innovation [Dow et al]
- Creating quick response teams within the crowd
Pernille Bjorn added 2 new photos — with Micka Berkmann Ag and 2 others at City Hall Square, Copenhagen.

October 29 at 2:10pm · København, Denmark · &

Stop besparelsene på uddannelserne - demonstration fra rådhuspladsen til Christiansborg

See Translation
Tracking and sensing capabilities

Optimizing workers – personal productivity
Optimizing teams & relationships
Presenting new privacy challenges
Leading to automated future: algorithmic design, architecture, self driving cars
Algorithmic management

Lee, Metsky, Kusbit, & Dabbish (2015)

minimize
\[ f(x, y, z) = x^2 + (y - z)^2 + 10\sin(yz - x - y)z \]
such that
\[ x^2 - y^2 + z^2 \leq 1 \]
\[ x^2 + y - z \leq 0 \]
\[ -1 \leq x \leq 2 \]
\[ -3 \leq y \leq 3 \]
\[ -2 \leq z \leq 4 \]
Workflow
Kinnaird, Dabbish & Kiesler, 2013

Work group size
Kinnaird, Dabbish, Faste, & Kiesler, 2013

Results will be reviewed by a computer:

Evaluator identity
Mere presence of activity history increases credibility

Nugyen, Dabbish, & Kiesler 2015

Presentation affects perceived competence

Marlow & Dabbish, 2015; Marlow, Dabbish, & Forlizzi, 2015
What is the role of the network?

What networks are important?
  * Network of people and tasks

Who gets to know the network?

How does knowing the network help? When do you need network information?

Do networks facilitate performance in the same way in these different systems?
Implications

We can't design online work environments from hunches or guesswork. We need to test.

We need to consider the impacts of these systems. Whose interests should they serve? How do we decide?

What can open collaboration learn from crowdsourcing and vice versa?

Sociotechnical view of these systems: social effects of and interactions around platforms.
collaborators

Colleen Stuart  Jim Herbsleb  Jason Tsay  Jennifer Marlow  Min Kyung Lee

Mary Nguyen  Peter Kinnaird  Sara Kiesler  Haakon Faste  Luis von Ahn
Thank you!

This research is supported by:

Carnegie Mellon University
Center for the Future of Work
EMERGING TECHNOLOGIES FOR THE NEXT GENERATION WORKPLACE

The National Science Foundation under awards: IIS-1111750, CNS-1040801

coexlab.com
dabbish@cmu.edu
@dabbish