Open Source Barriers to Entry: A SocioTechnical Perspective

Anita Sarma, Associate Professor
Electrical Engineering and Computer Science (EECS)
Oregon State University
Human Factors in Software Development

- Contextualized development editors for exploratory programming
- Coordination in software development teams
- Usable security – implication of privacy configuration in IOT devices
- Removing barriers for newcomers to Open Source Software
Open Source Barriers to Entry: A SocioTechnical Perspective

Margaret Burnett, Igor Steinmacher, Marco Gerosa, Christopher Mendez, Hema Susmita Pedala, Felipe Fronchetti, Zoe Steine-Hanson, Claudia Hilderbrand, Amber Horvath
Lots of people are left behind!

Researchers analyzed: Apache Hadoop (60 month data)\textsuperscript{1}

Absence of response, politeness, usefulness, and type of the author influence the retention of newcomers in an open source project

82\% of dropouts!!!

\textsuperscript{1}Steinmacher, Wiese, Chaves & Gerosa, ”Why do newcomers abandon open source software projects?”, 6th Int. Workshop on Cooperative and Human Aspects of Software Engineering (CHASE 2013)
11.2% women contributors among 2000 OSS projects
[Laura Arjona-Reina Gregorio Robles, and Santiago Dueas 2013 survey]

< 5% owners of top 5000 OSS projects are women
[Alyssa Frazee. 2014. Exploring the data on gender and GitHub repo ownership]

No women contributors among the top 100 contributors or followed
Paul Miller. 2013. Most active GitHub users.

“...women’s work was more likely to be accepted than men’s, unless “their gender is identifiable”, in which case the acceptance rate was worse than men’s”
[Josh Terrel et al. 2016, Gender Differences and Bias in Open Source: Pull Request Acceptance of Women versus Men]
Lot of researchers are looking into OSS community biases…

…but what about the tools?
• Q: Does software support a variety of smart users?
• A: No.
  • Unconscious bias, supporting (mainly) 1 kind of smart user.

• What to do about unconscious bias in software tools?
What **NOT** to do

Shrink it and **Pink it** is not a strategy

Dell’s pink laptops (2009)

BIC for her

Nobody is “typical”

World doesn’t divide into “typical” F vs. M

So, needs to be about inclusiveness

“We want your $. Here’s a surface gesture to show that we’re **not** thinking about you”
What to do: GenderMag to find tool biases

• **Gender Inclusiveness Magnifier**
  • Process: Evaluate tools’ inclusiveness

• GenderMag Personas:
  • "representatives" of a range of users, but only...

• ...from the perspective of **5 facets**:
  • Motivations
  • Information processing style
  • Computer self-efficacy
  • Risk averseness
  • Tech learning style (tinkering)
How GenderMag Works

1. Pick a persona. eg: Abby
2. Pick a use case/scenario in your tool, eg:
in Augmented (Physical) Bookstore
   “Find science fiction books”
3. Walk thru scenario via
   “intended” subgoals & actions
Abby has always liked music. When she is on her way to work in the mornings, she listens to music that spans a wide variety of styles. But when she arrives at work, she turns it off, and begins her day scanning all her emails first to get an overall picture before answering any of them. (This extra pass takes time but seems worth it.) Some nights she exercises or stretches, and sometimes she likes to play computer puzzle games like Sudoku.

**Background and skills**

Abby works as an accountant. She is comfortable with the technologies she uses regularly, but she just moved to this employer 1 week ago, and their software systems are new to her. She’s a “numbers person”, but she has never taken any computer programming or IT classes. She likes Math and knows how to think with numbers. She writes and edits spreadsheet formulas in her work. She, also enjoys working with numbers and logic. She especially likes working out puzzle games, either on paper or on the computer.

**Computer Self-Efficacy:** Abby has low confidence about doing unfamiliar computing tasks. If problems arise with her technology, she often blames herself for these problems. This affects whether and how she will persevere with a task if technology problems have arisen.

**Attitude toward Risk:** Abby’s life is a little complicated and she rarely has spare time. So she is risk averse about using unfamiliar technologies that might need her to spend extra time on them, even if the new features might be relevant. She instead performs tasks using familiar features, because they’re more predictable about what she will get from them and how much time they will take.

**How Abby Works with Information and Learns:**

- **Information Processing Style:** Abby tends towards a comprehensive information processing style when she needs to more information. So, instead of acting upon the first option that seems promising, she gathers information comprehensively to try to form a complete understanding of the problem before trying to solve it. Thus, her style is “burst-y”; first she reads a lot, then she acts on it in a batch of activity.

- **Learning: by Process vs. by Tinkering:** When learning new technology, Abby leans toward process-oriented learning, e.g., tutorials, step-by-step processes, wizards, online how-to videos, etc. She doesn’t particularly like learning by tinkering with software (i.e., just trying out new features or commands to see what they do), but when she does tinker, it has positive effects on her understanding of the software.

1 Abby represents users with motivations/attitudes and information/learning styles similar to hers. For data on females and males similar to and different from Abby, see [http://eusesconsortium.org/gender/gender.php](http://eusesconsortium.org/gender/gender.php)
Abby Jones

- 28 years old
- Employed as an Accountant
- Lives in Cardiff, Wales

---

**Background and skills**

Abby works as an accountant. She is comfortable with the technologies she uses regularly, but she just moved to this employer 1 week ago, and their software systems are new to her.

Abby says she’s a “numbers person”, but she has never taken any computer programming or IT systems classes. She likes Math and knows how to think with numbers. She writes and edits spreadsheet formulas in her work.

---

**Motivations and Attitudes**

- **Motivations**: Abby uses technologies to accomplish her tasks. She learns new technologies if & when she needs to, but prefers to use methods she is already familiar and comfortable with. She likes to work with music that spans a wide variety of styles.

- **Computer Self-Efficacy**: Abby has low confidence about doing unfamiliar computing tasks. If problems arise with her technology, she often blames herself for these problems.

- **Attitude toward Risk**: Abby’s life is a little complicated and she rarely has spare time. So she is risk averse about using unfamiliar technologies that might need her to spend extra time on them, even if the new features might be relevant.

---

**Facet #1**

- **Information Processing Style**: Abby tends towards a comprehensive information processing style when she needs to more information. So, instead of acting upon the first option that seems promising, she gathers information comprehensively to try to form a complete understanding of the problem before trying to solve it. Thus, her style is “burst-y”; first she reads a lot, then she acts on it in a batch of activity.

- **Learning by Process vs. by Tinkering**: When learning new technology, Abby leans toward process-oriented learning, e.g., tutorials, step-by-step processes, wizards, online how-to videos, etc. She doesn’t particularly like learning by tinkering with software (i.e., just trying out new features or commands to see what they do), but when she does tinker, it has positive effects on her understanding of the software.

---

1 Abby represents users with motivations/attitudes and information/learning styles similar to hers. For data on females and males similar to and different from Abby, see [http://eusesconsortium.org/gender/gender.php](http://eusesconsortium.org/gender/gender.php)
Facet #1: Risk

Most tolerant

Middle 1/3

Most averse

42%  25%

26%  33%

29%  38%
Study design

- Five OSS Teams used the method – GenderMag – on their own projects
- Three from large software development company
- Two from Open Source Lab at OSU
GenderMag field study: teams

<table>
<thead>
<tr>
<th>Teams’ Gender Make-Up</th>
<th>Open Source Project</th>
<th>Use Cases</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>A cloud</td>
<td>1. Use GitHub issue tracker to find an issue</td>
</tr>
<tr>
<td>Team Y</td>
<td>All female team</td>
<td>3. Find help with pull request</td>
</tr>
<tr>
<td></td>
<td>A graph database</td>
<td>4. Review submitted pull request</td>
</tr>
<tr>
<td>Team Z</td>
<td>Mixed gender team</td>
<td>5. Set up the environment</td>
</tr>
<tr>
<td></td>
<td>A graph database</td>
<td>3. Reviewing submitted pull request</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Use Cases**
1. **Get familiar** with the project
2. Use GitHub issue tracker to **find an issue**
3. **Find help** with pull request
4. **Review** submitted pull request
5. **Set up** the environment
Tools implicated in gender issues

Analyzed the tools and barriers with respect to each of the GenderMag facets:
- Motivations
- Computer self-efficacy
- Risk aversion
- Information processing style
- Learning style
OSS tools deeply create Glass Floors

<table>
<thead>
<tr>
<th>Barrier</th>
<th># in Tools</th>
<th># with gender bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical barriers</td>
<td>56</td>
<td>37 (66%)</td>
</tr>
<tr>
<td>Documentation barriers</td>
<td>36</td>
<td>23 (64%)</td>
</tr>
<tr>
<td>Cultural barriers</td>
<td>7</td>
<td>6 (86%)</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Tools implicated in newcomer barriers

Newcomers' characteristics
- Lack of proactivity
- Lack of commitment
- Underestimating the challenge
- Lack of patience
- Not sending a meaningful/correct message
- English level
- Shyness
- Making useless comments...
- Low responsiveness
- Not acknowledging/flaking answers
- Lack of domain expertise
- Lack of knowledge in... practices
- Knowledge on technologies... used
- Proper knowledge... programming language
- Knowledge on versioning...
- Choosing the right development tools
- Experience on unit testing

Newcomers' orientation
- Finding a task to start with
- Finding a mentor
- Finding the correct artifacts to fix an issue
- Poor "How to contribute"... NO3
- Outdated list of bugs
- Reproducing issues
- Newcomers don't know... flow

Reception issues
- Not receiving an answer
- Delayed answers
- Impolite answers
- Receiving answers... too... complex...

Cultural differences
- Newcomers need... a real person
- Message is considered rude

Documentation problems
- Outdated documentation
- Information overload
- Unclear documentation
- Spread documentation
- Code comments not clear
- Lack of documentation in general
- Design documents
- Documentation on setting up workspace
- Code comments
- Code documentation
- Documentation on project source

Technical hurdles
- Building workspace locally
- Platform dependency
- Library dependency
- Finding the correct source
- Bad code quality
- Code complexity/insufficiency
- Codebase size
- Bad design quality
- Lack of code standards
- Outdated code
- Understanding the code
- Understanding architecture/code structure
- Understanding flow of information
- Lack of info... to send a contribution
- Delay to get contribution accepted/reviewed
- Getting contribution accepted
- Issue to create a patch
Tools implicated in newcomer barriers

- Newcomers characteristics
- Newcomers orientation
- Reception issues
- Cultural differences
- Documentation problems
- Technical hurdles
### OSS tools deeply create Glass Floors

<table>
<thead>
<tr>
<th>Barrier</th>
<th># in Tools</th>
<th># with gender bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Technical barriers</td>
<td>56</td>
<td>37 (66%)</td>
</tr>
<tr>
<td>Documentation barriers</td>
<td>36</td>
<td>23 (64%)</td>
</tr>
<tr>
<td>Cultural barriers</td>
<td>7</td>
<td>6 (86%)</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Team X-P62: “... take a while because she has comprehensive information processing”

<table>
<thead>
<tr>
<th>Activity</th>
<th>Instr. Motivation</th>
<th>Info. Processing</th>
<th>Self-efficacy</th>
<th>Attitude to Risk</th>
<th>Learning style</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find a task</td>
<td></td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>Learn more about the task</td>
<td></td>
<td>![Checkmark]</td>
<td></td>
<td>![Checkmark]</td>
<td>![Checkmark]</td>
</tr>
<tr>
<td>Continue with the task</td>
<td></td>
<td>![Checkmark]</td>
<td></td>
<td>![Checkmark]</td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Team Z-P57: “...she wants to follow the steps [but instructions not there]”

<table>
<thead>
<tr>
<th>Activity</th>
<th>Motivation</th>
<th>Info. processing</th>
<th>Self-efficacy</th>
<th>Style</th>
<th>Learning style</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find a task</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td>✔</td>
<td></td>
</tr>
<tr>
<td>Learn more about the task</td>
<td>✔</td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>Continue with the task</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>✔</td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Team V-P60: “...she might blame herself right now”

<table>
<thead>
<tr>
<th>Activity</th>
<th>Motivation</th>
<th>Info. processing</th>
<th>Self-efficacy</th>
<th>Attitude to Risk</th>
<th>Learning style</th>
</tr>
</thead>
<tbody>
<tr>
<td>Find a task</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Learn more about the task</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
</tr>
<tr>
<td>Continue with the task</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Results corroborated by data from newcomers

- Diary studies from newcomers (students learning OSS)
- 9 women – 9 men diaries
## Tools need to match people’s cognitive styles

<table>
<thead>
<tr>
<th>Barrier</th>
<th>Men Diaries</th>
<th></th>
<th>Women Diaries</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># barriers in tools</td>
<td># involving facets</td>
<td># barriers in tools</td>
<td># involving facets</td>
</tr>
<tr>
<td>Technical barriers</td>
<td>19</td>
<td>5 (26.32%)</td>
<td>40</td>
<td>23 (57.5%)</td>
</tr>
<tr>
<td>Documentation barriers</td>
<td>22</td>
<td>13 (59.09%)</td>
<td>79</td>
<td>54 (68.35%)</td>
</tr>
<tr>
<td>Cultural barriers</td>
<td>3</td>
<td>0 (0%)</td>
<td>7</td>
<td>3 (42.86%)</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td>Total</td>
<td>83</td>
<td>32 (38.55%)</td>
<td>251</td>
<td>153 (60.96%)</td>
</tr>
</tbody>
</table>
Tools need to match people’s cognitive styles

<table>
<thead>
<tr>
<th>Barrier</th>
<th>Men Diaries</th>
<th>Women Diaries</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># barriers in tools</td>
<td># involving facets</td>
</tr>
<tr>
<td>Technical barriers</td>
<td>19</td>
<td>5 (26.32 %)</td>
</tr>
<tr>
<td>Documentation barriers</td>
<td>3</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>Cultural barriers</td>
<td>83</td>
<td>32 (38.55%)</td>
</tr>
</tbody>
</table>

Significant differences in number of facets implicated in barriers faced by women as compared to men (Fishers exact test <0.05)
Gender inclusive software rests ...

Not on “gender bucketing”,

Rather on supporting diverse ways of thinking & problem-solving.

One gender at a time -facet-
Follow-ups & Resources

@GenderMag, #GenderMag
gendermag.method

Resources: gendermag.org
Flyer, papers, personas, foundations, ...
Download the kit!

Make it happen @ your department/ project/company!
anita.sarma@oregonstate.edu
margaret.burnett@oregonstate.edu
New GenderMag+OSS Research! Mentors in OSS...

Challenges faced by women mentors:

• Differences in motivation: “…feel a sense of attachment…”
• Self-efficacy: “…communication styles...awkward”
• Lack of Peer Parity: “not seeing a lot of people like oneself in a community”
• Stereotyping:
  • (warmth) “…my gender makes me more approachable”
  • (competence): “…mentor …community organizer, and not see me as an engineer”
Tools need to match people’s cognitive styles

<table>
<thead>
<tr>
<th>Use Cases</th>
<th># of Issues</th>
<th>% of issues per step</th>
</tr>
</thead>
<tbody>
<tr>
<td>Using GitHub issue tracker</td>
<td>12</td>
<td>71%</td>
</tr>
<tr>
<td>Finding help with pull requests</td>
<td>13</td>
<td>54%</td>
</tr>
<tr>
<td>Setting up the environment</td>
<td>40</td>
<td>44%</td>
</tr>
<tr>
<td>…</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
## OSS tools deeply create Glass Floors

<table>
<thead>
<tr>
<th>Barrier</th>
<th>Men Diaries</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># barriers in tools</td>
<td># involving facets</td>
</tr>
<tr>
<td><strong>Technical barriers</strong></td>
<td>19</td>
<td>5 (26.32%)</td>
</tr>
<tr>
<td><strong>Documentation barriers</strong></td>
<td>22</td>
<td>13 (59.09%)</td>
</tr>
<tr>
<td><strong>Cultural barriers</strong></td>
<td>3</td>
<td>0 (0%)</td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>83</td>
<td>32 (38.55%)</td>
</tr>
</tbody>
</table>

- Technical barriers: 19 barriers, 5 barriers involving facets (26.32%)
- Documentation barriers: 22 barriers, 13 barriers involving facets (59.09%)
- Cultural barriers: 3 barriers, 0 barriers involving facets (0%)
- Total: 83 barriers, 32 barriers involving facets (38.55%)
## Newcomer barriers reported in diaries

<table>
<thead>
<tr>
<th>Barrier</th>
<th>Men Diaries</th>
<th></th>
<th>Women Diaries</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td># barriers in tools</td>
<td># barriers in tools</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Technical barriers</td>
<td>19</td>
<td>40</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Documentation barriers</td>
<td>22</td>
<td>79</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cultural barriers</td>
<td>3</td>
<td>7</td>
<td></td>
<td></td>
</tr>
<tr>
<td>...</td>
<td>...</td>
<td>...</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>83</td>
<td>251</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>