Mentoring at scale: Acknowledging Implicit Mentoring

Anita Sarma, Oregon State University

@asarma
Mentoring in OSS

Formal mentoring

Programs where mentors and mentees are formally connected

Informal mentoring

Explicit mentoring

Mentees seek out/formally paired with mentors

Implicit mentoring

Occurs in everyday development activities (e.g., code reviews)
Implicit Mentoring
Implicit mentoring can be defined as “mentoring that occurs in everyday development activities such as code reviews, where a mentor provides an underlying explanation when providing suggestions, instructions, or mechanisms to address errors”
We define, we mine, we classify

37 repositories

97,444 Pull Request (PR) comments
11,634 contributors

Machine learning classifier

Training data through manual labeling
## Training the ML classifier

**Explanations accompanying GitHub PR Comments**

<table>
<thead>
<tr>
<th>Mentoring Action</th>
<th>PR_comment sample</th>
</tr>
</thead>
<tbody>
<tr>
<td>Instruction</td>
<td>M-284: “...run [tool] on the project before creating a PR. You would have noticed [problem]...”</td>
</tr>
<tr>
<td>Suggestion</td>
<td>B-553: “ I would still duplicate [action] like I did in [certain PR] because it’s widely used in [tests]. Maybe this could be removed after [situation].”</td>
</tr>
<tr>
<td>Mechanisms to fix errors</td>
<td>I-1376: “ Would you mind just doing [action] again to kick off [framework]? I think [framework] is just not happy when it has a lot of loads.”</td>
</tr>
</tbody>
</table>
Implicit mentoring is widespread

- 30.27% of Pull-Request included implicit mentoring.
- 25.24% of contributors served as implicit mentors.
Features of implicit mentoring

- More than dyadic
- More than top to down
- Is topical

Of the 29,502 PRs:
- dyadic (65.15%),
- triads (22.79%)
- > quadrads (12.06%).

“it takes a village...”

It is not just top to down, also including bottom up (13.08%), and peer to peer (<6 months diff, 34.14%).

Implicit mentoring is interest driven; can it be drive-by mentoring?
Does gender play a role?
Identifying gender

Used NameSorML API
- First & last names
- Geographical location
- >90% confidence
Gender and implicit mentoring?

- Men perform implicit mentoring more often. 
- Proportionality test (7% more, p-value<0.001), but low effect size
Homophily preferences of implicit mentoring

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Top -&gt; Down</th>
<th>Peer -&gt; Peer</th>
<th>Bottom -&gt; Up</th>
</tr>
</thead>
<tbody>
<tr>
<td>W -&gt; W</td>
<td>316</td>
<td>55</td>
<td>209</td>
<td>52</td>
</tr>
<tr>
<td>W -&gt; M</td>
<td>453</td>
<td>125</td>
<td>179</td>
<td>149</td>
</tr>
<tr>
<td>M -&gt; W</td>
<td>1642</td>
<td>839</td>
<td>735</td>
<td>68</td>
</tr>
<tr>
<td>M -&gt; M</td>
<td>48982</td>
<td>26256</td>
<td>17162</td>
<td>5564</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>51393</td>
<td>27275</td>
<td>18285</td>
<td>5833</td>
</tr>
</tbody>
</table>

Homophilic mentoring **surpassed** cross-gender mentoring for top->down, peer -> peer, and bottom -> up mentoring with large effect size.
Do mentors “reach across the aisle”? 

In the few cases of cross-gender implicit mentoring, women tend to cross gender boundaries more often (56%) than men (p-value<0.001, d=1.39).
Discussion
Discussion

• Why is there homophily in implicit mentoring via code review?
  — Does this mean fewer women get mentored?

• How to acknowledge implicit mentoring?
  — Would we retain more mentors if they are acknowledged formally?

• What other “invisible work” are we missing?
  — Would acknowledging invisible /non-coding work attract and retain more women in OSS?
Thank you!
anita.sarma@oregonstate.edu
@asarma

Thank you to Interviewees and team and