Going Farther Together: The Impact of Social Capital on Sustained Participation in Open Source

Huilian Sophie Qiu, Alexander Nolte, Anita Brown, Alexander Serebrenik, and Bogdan Vasilescu

@sophiehsqq @alexander_nolte @aserebrenik @b_vasilescu

STRUDEL
SOCIO-TECHNICAL RESEARCH USING DATA EXCAVATION LAB

Carnegie Mellon University
School of Computer Science

isr institute for SOFTWARE RESEARCH

Maintain, or suffer

https://w3techs.com/technologies/history_overview/web_server

(Greenstein and Nagel, 2016)

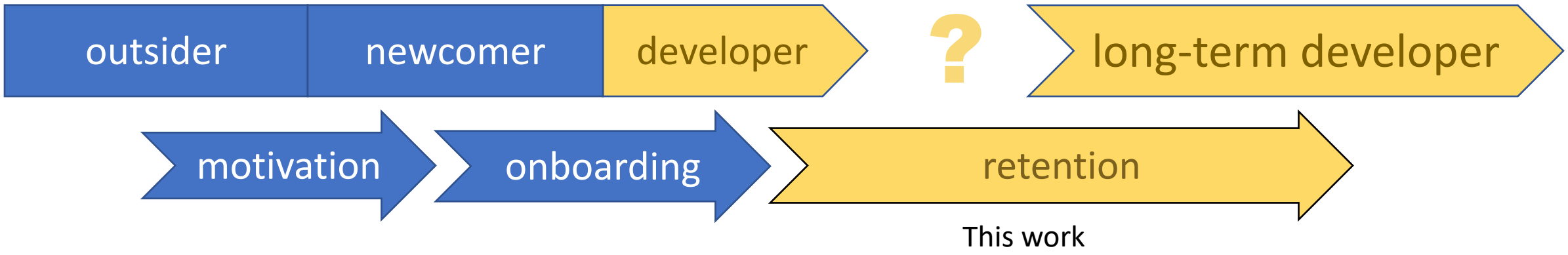
Open-source experience boosts resumes

Employers (and job seekers) use open-source experience to make inferences (or form impressions) about a candidate's technical skills.

(Marlow et al., 2013)

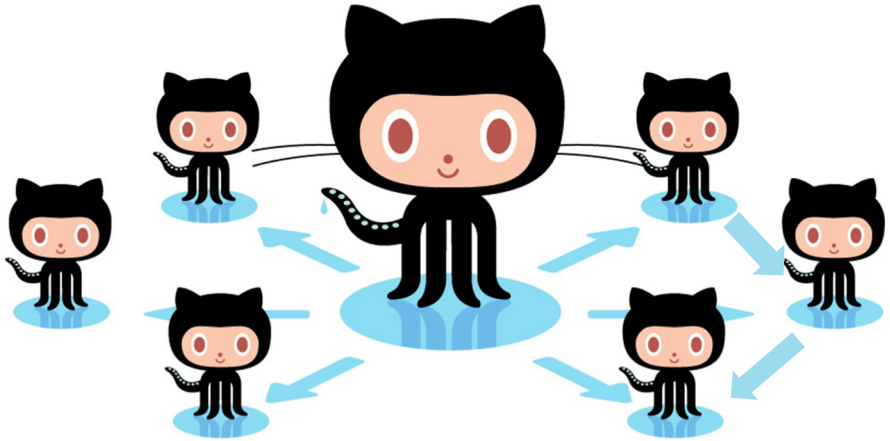
https://codeformore.com/how-to-write-up-open-source-experience-when-you-dont-have-any/

Knowledge gap

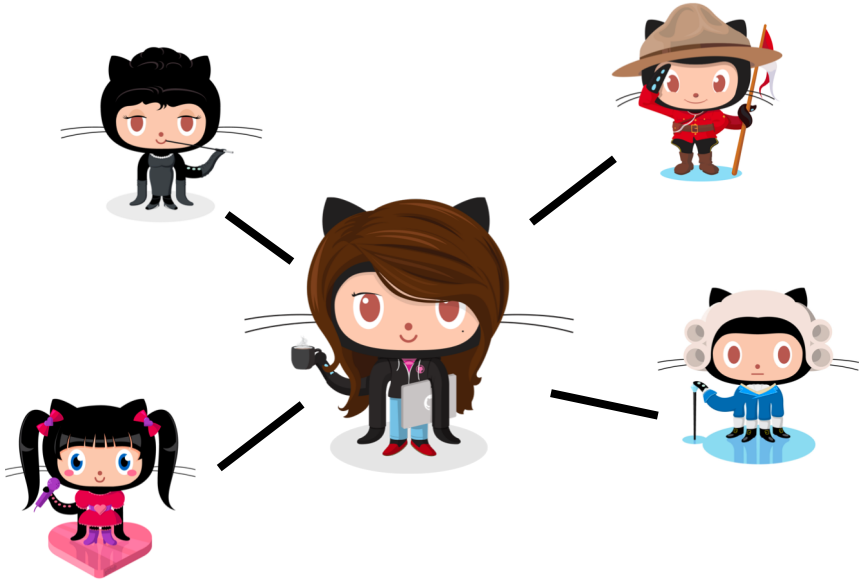


RESULT HIGHLIGHTS: Social capital explains sustained participation

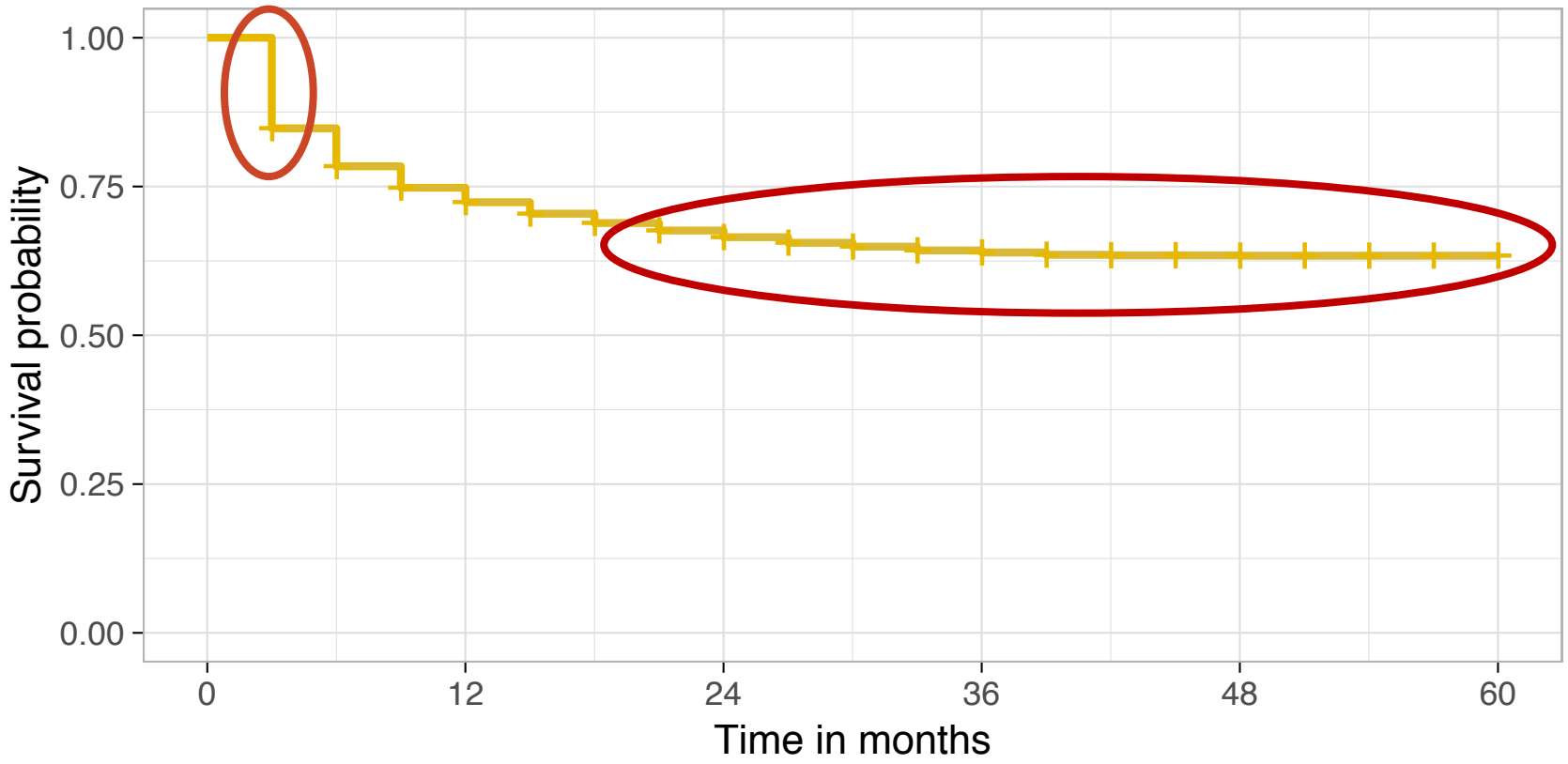
Higher social capital



Being part of teams with more diverse information, especially for women

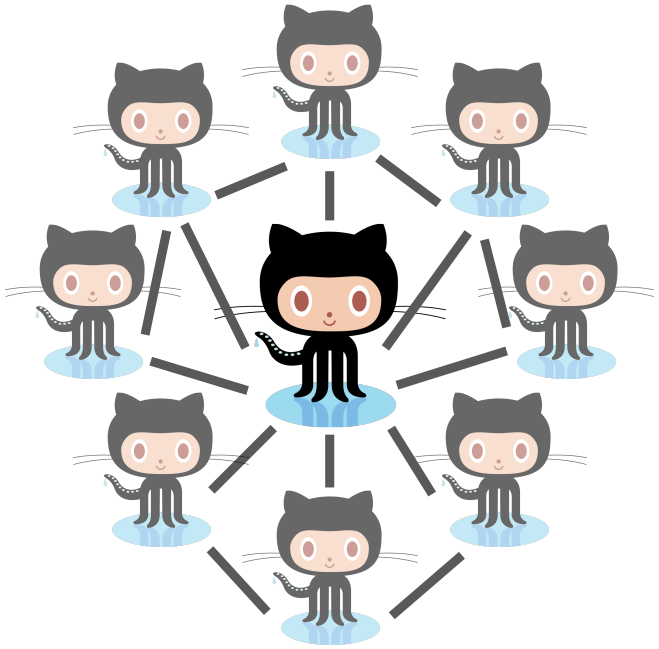


~40% of GitHub contributors disengage



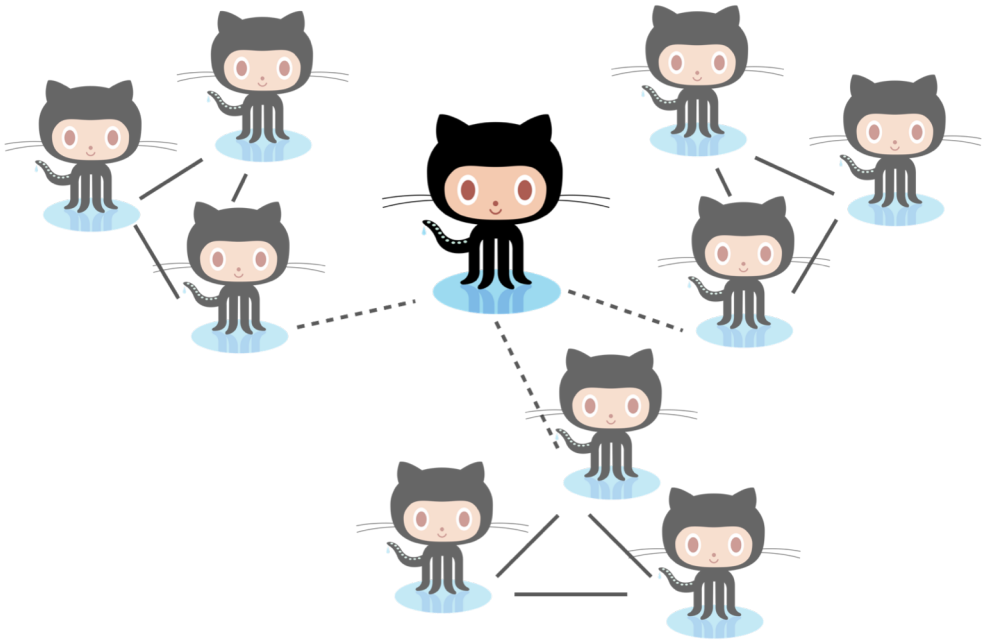
Social capital theory for sustainable participation

Bonding social capital:
benefiting from strongly connected network



Willingness to continue
(Coleman, 1990)

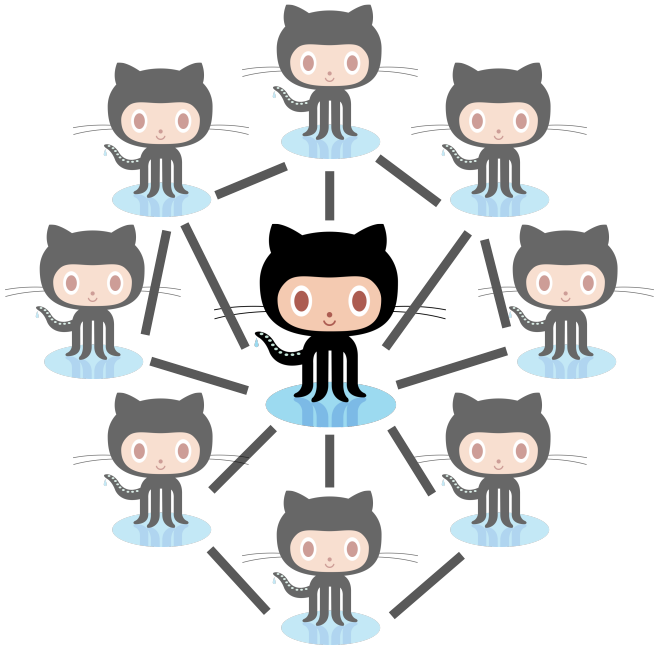
Bridging social capital:
benefiting from network with diverse info



Opportunity to continue
(Burt, 1998, 2001)

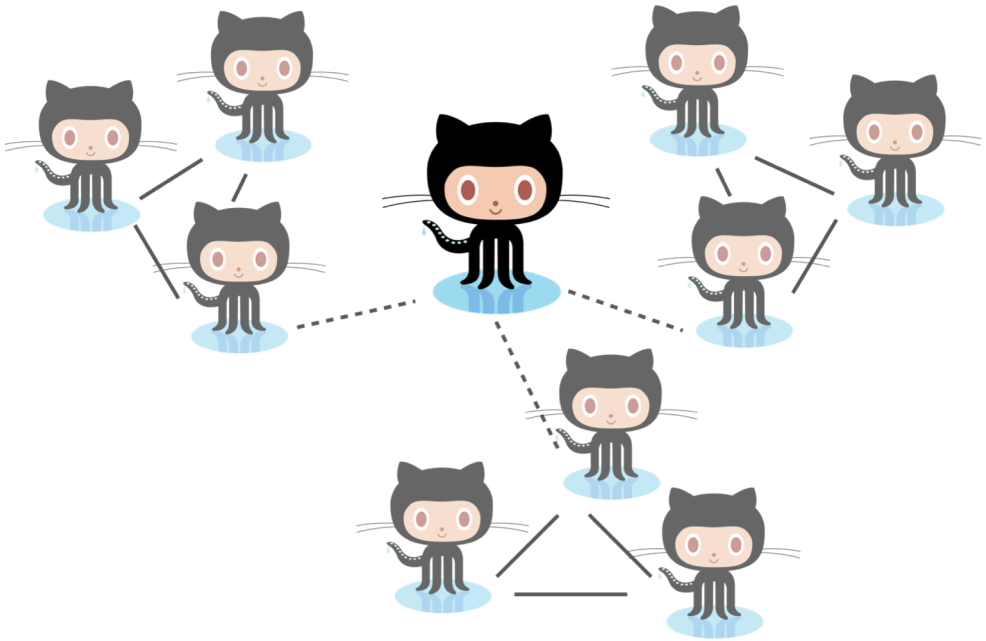
H1: more social capital ~ more prolonged engagement

Bonding social capital:
benefiting from strongly connected network



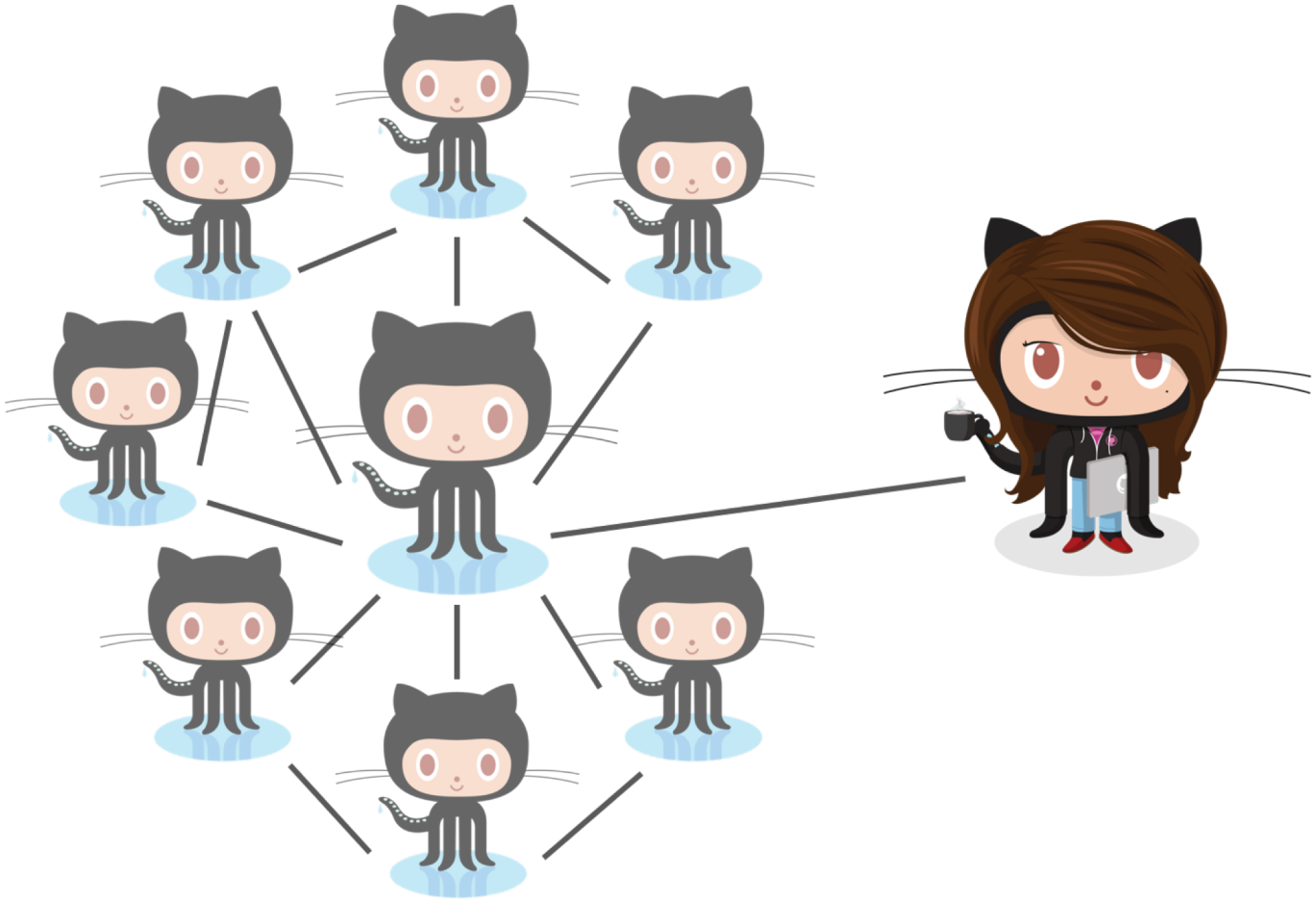
Willingness to continue
(Coleman, 1990)

Bridging social capital:
benefiting from network with diverse info



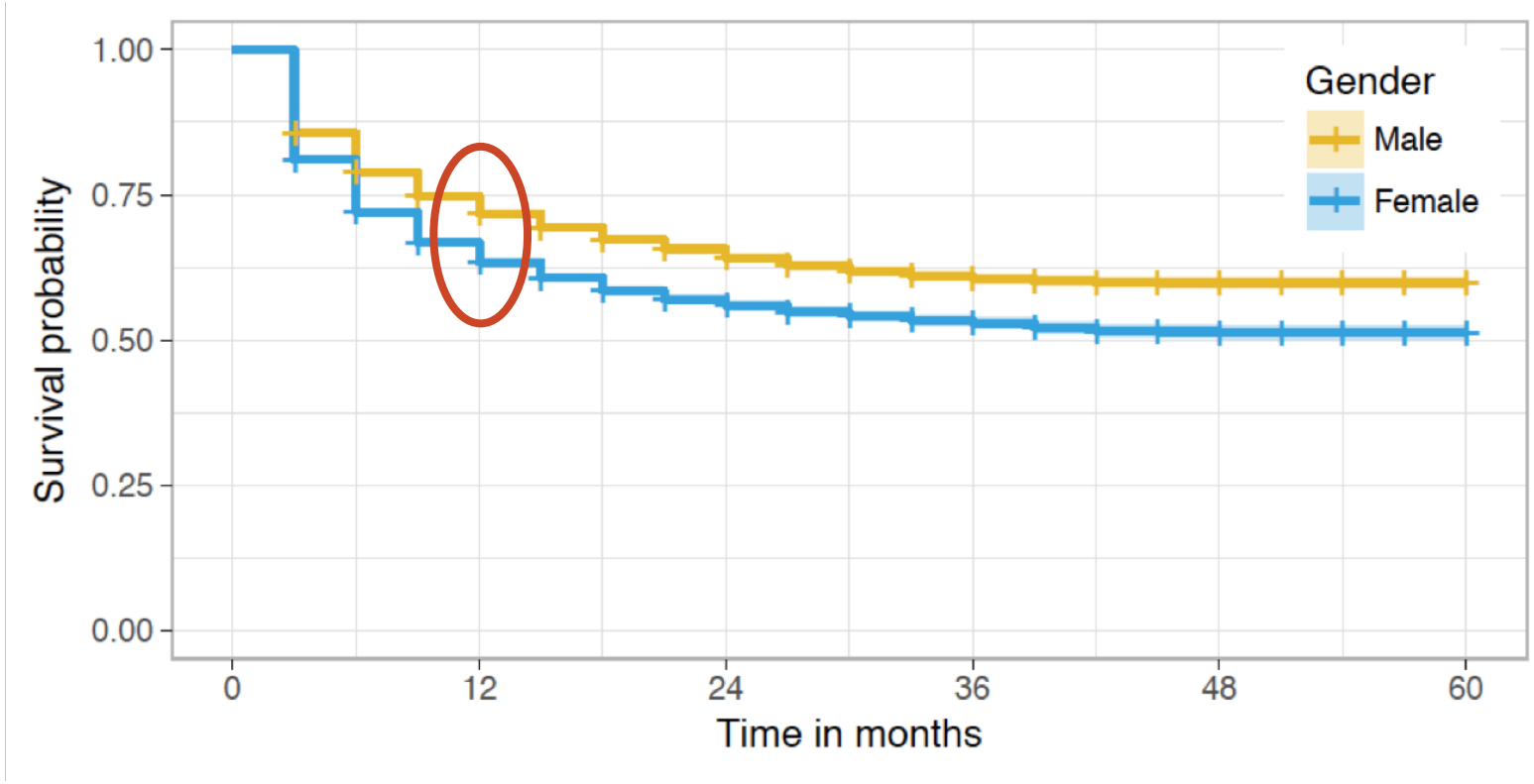
Opportunity to continue
(Burt, 1998, 2001)

Cohesive networks might foster discrimination / exclusion



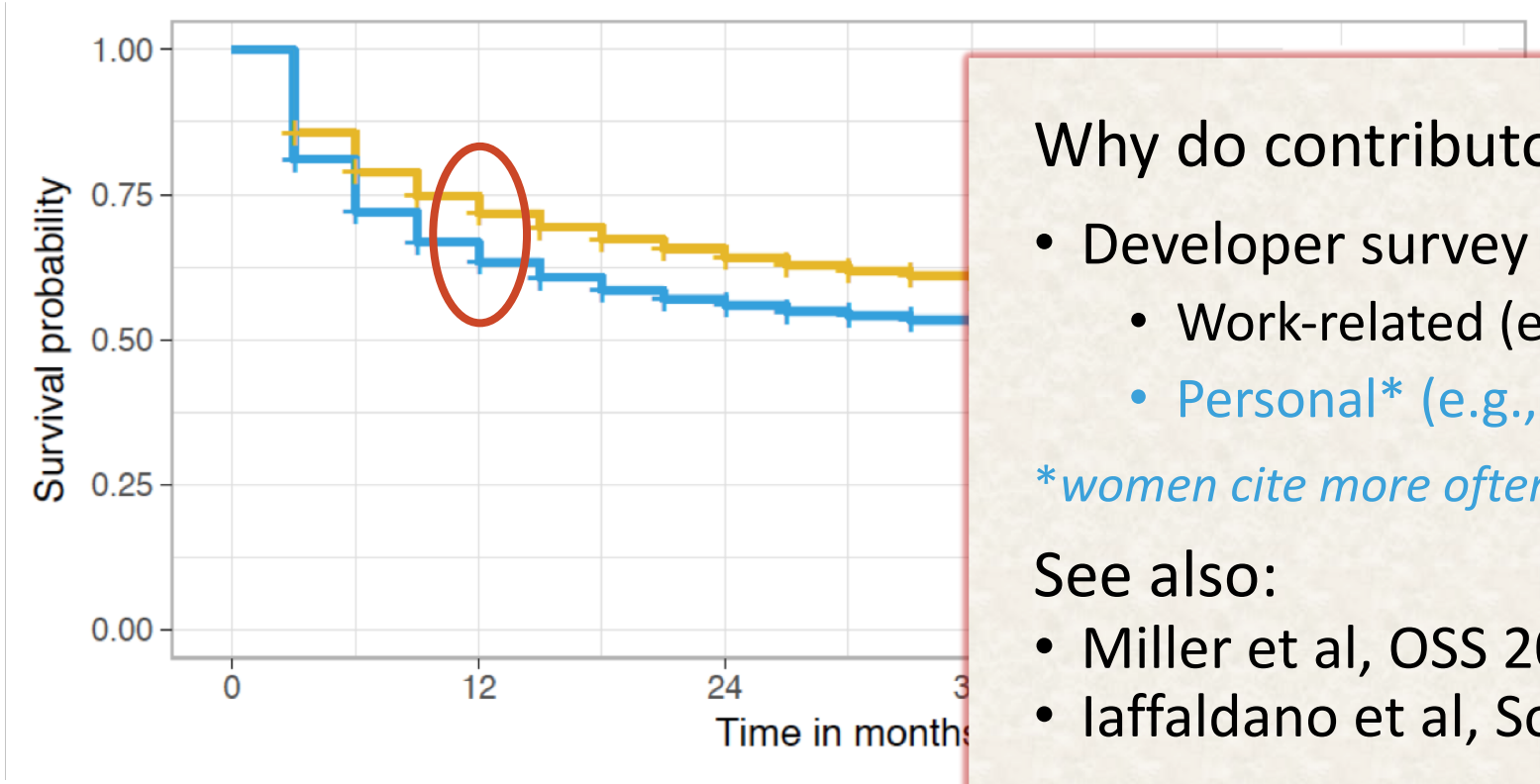
On GitHub, women disengage earlier than men

After one year ca. 70% of men are still active but only ca. 60% of women



On GitHub, women disengage earlier than men

After one year ca. 70% of men are still active but only ca. 60% of women



Why do contributors disengage?

- Developer survey in our paper:
  - Work-related (e.g., new job)
  - Personal* (e.g., different hobby)

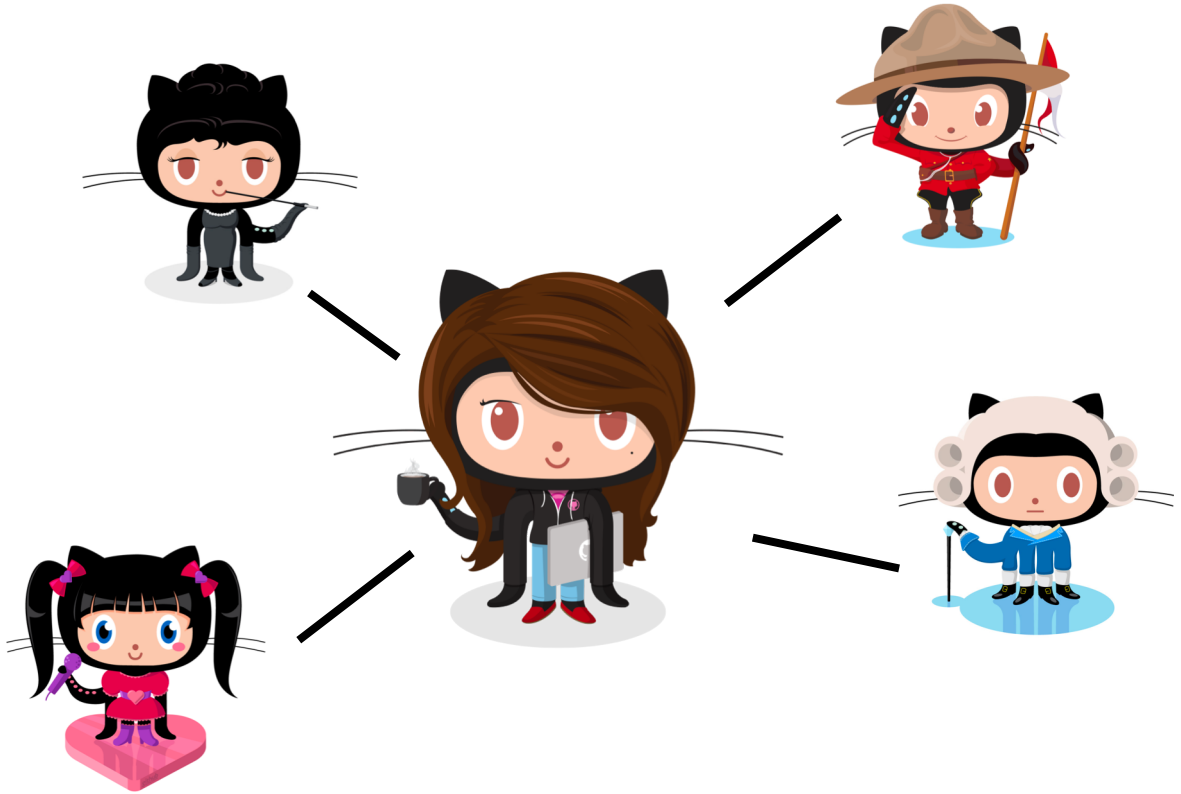
**women cite more often than men*

See also:

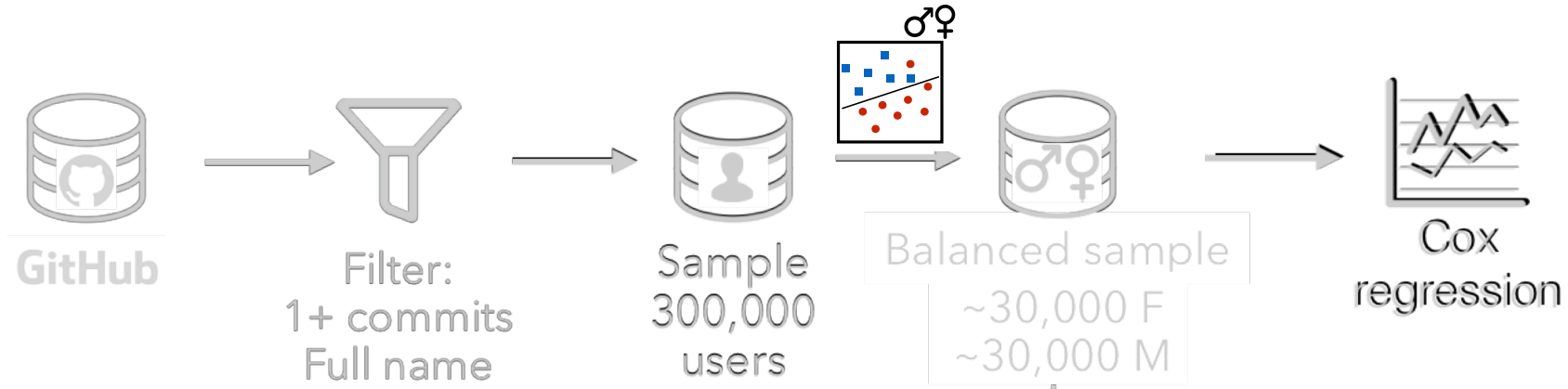
- Miller et al, OSS 2019
- Iaffaldano et al, SoHeal 2019

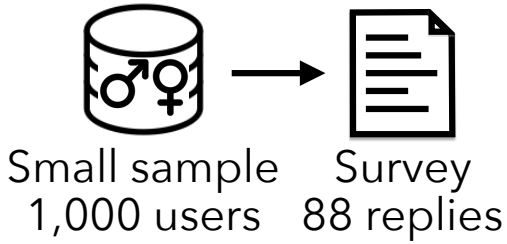
H2: Teams with more diverse information ~ more prolonged engagement, esp. for women

Information diversity should reduce the risk of demographic-based echo chambers.



Large-scale mixed-methods study



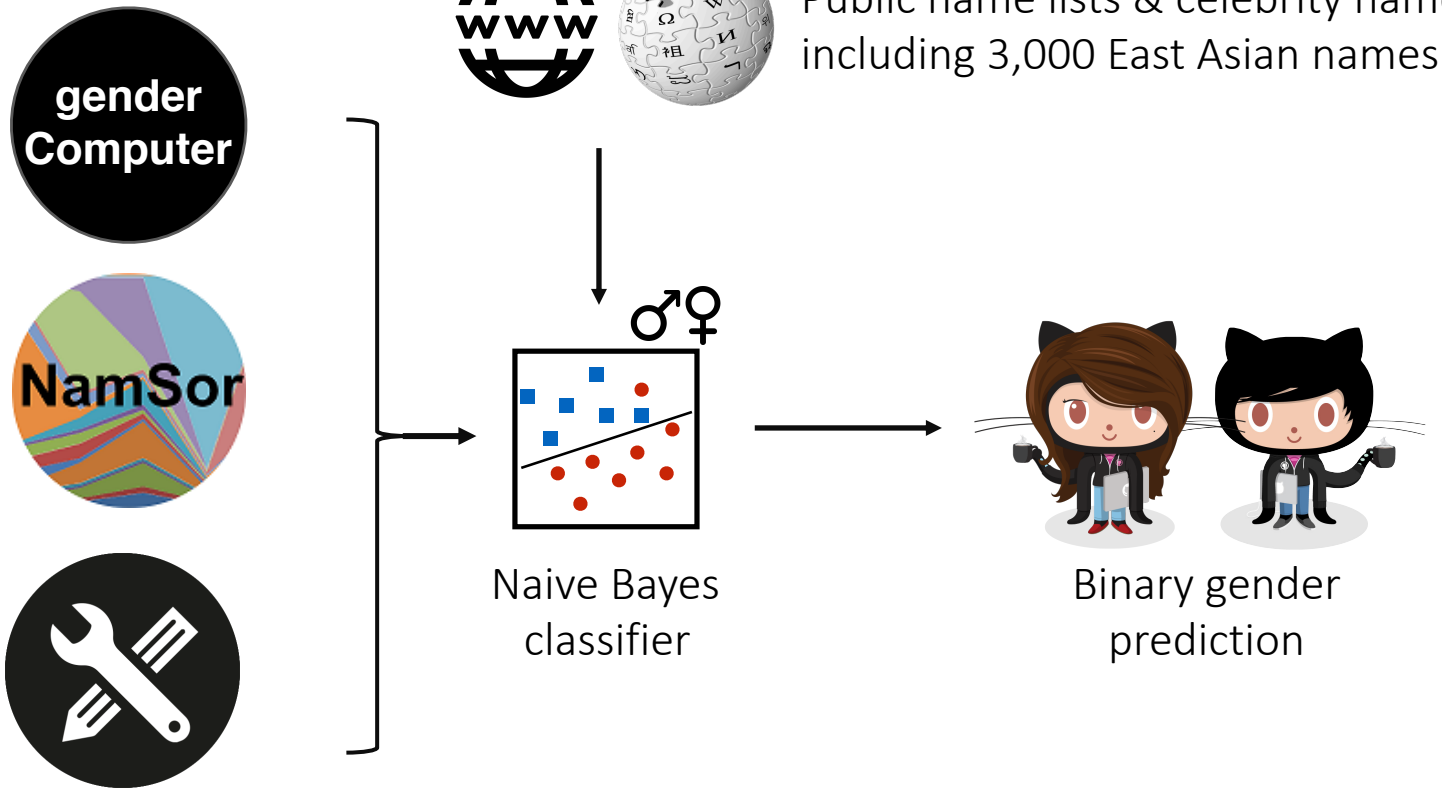


Inferring gender from names

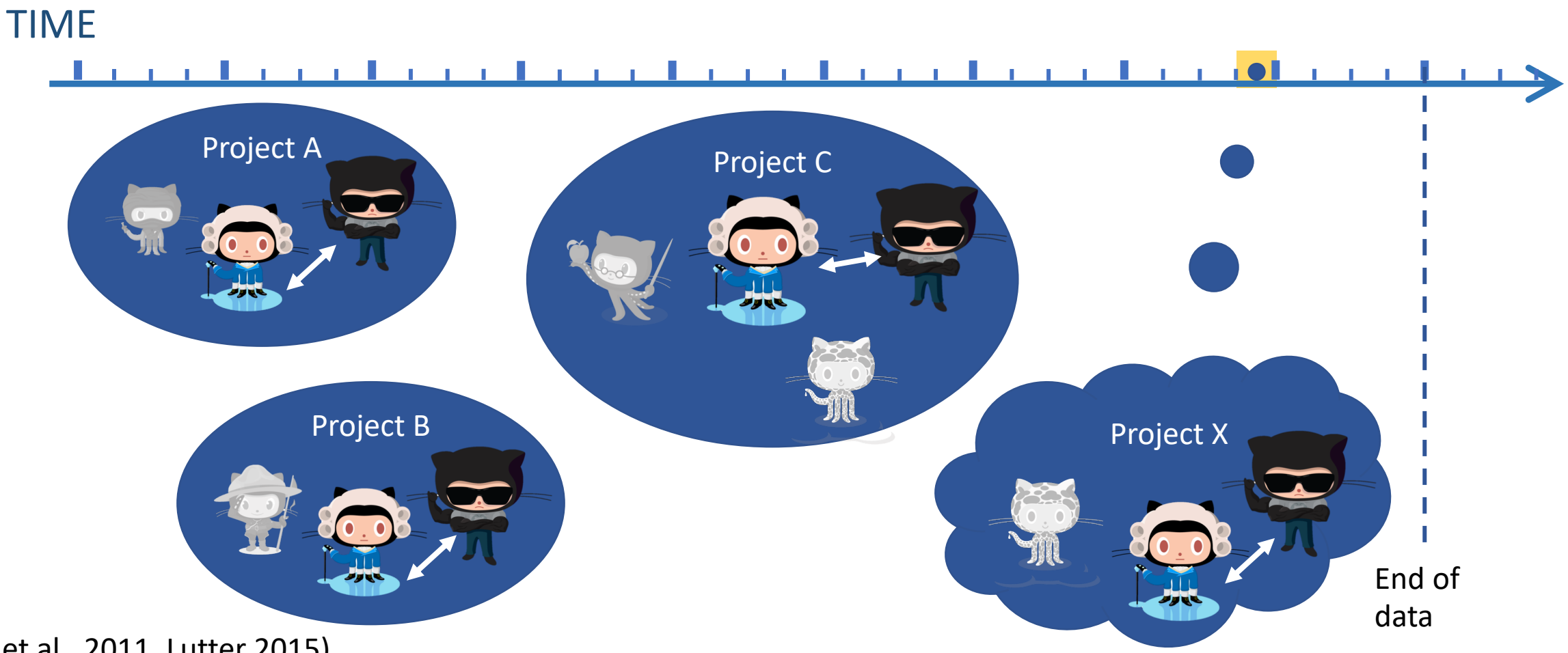
https://github.com/tue-mdse/genderComputer

https://www.namsor.com

name features, e.g., the last two characters

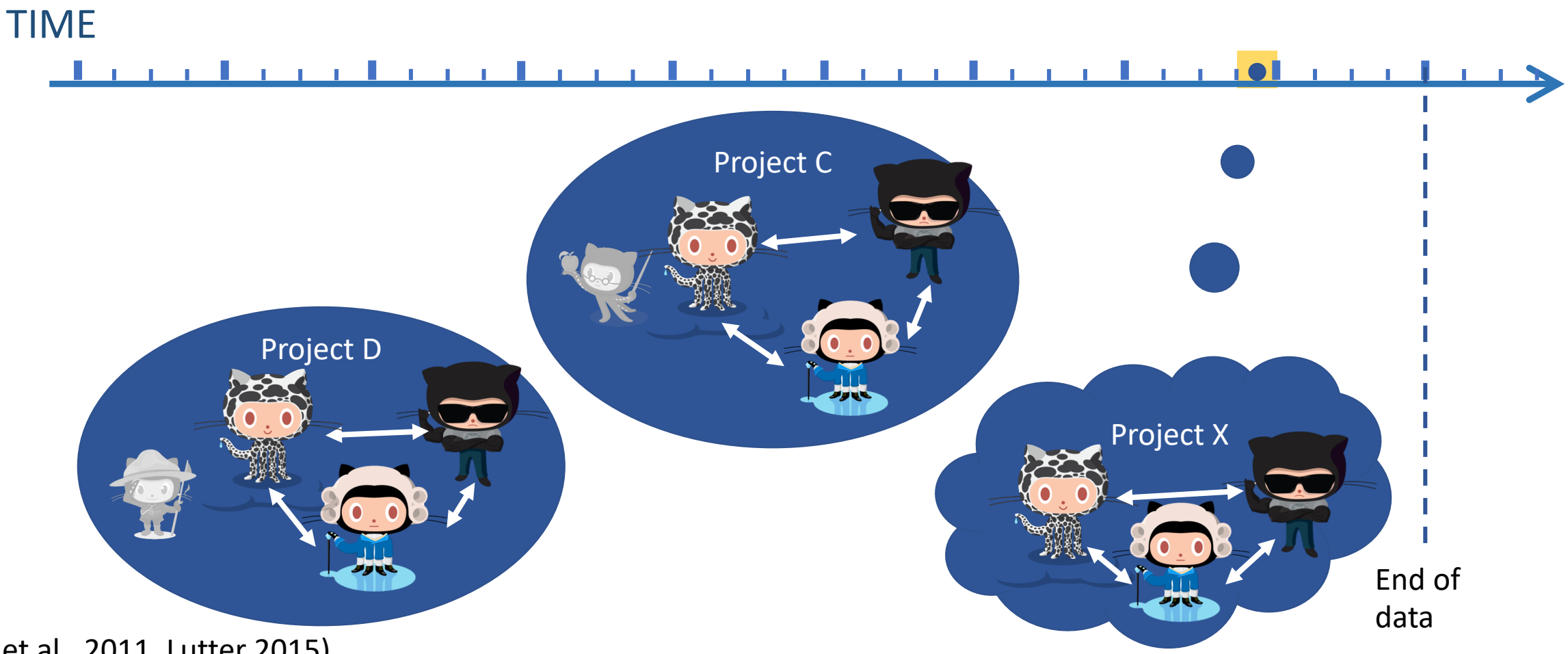


Bonding social capital – Team Familiarity



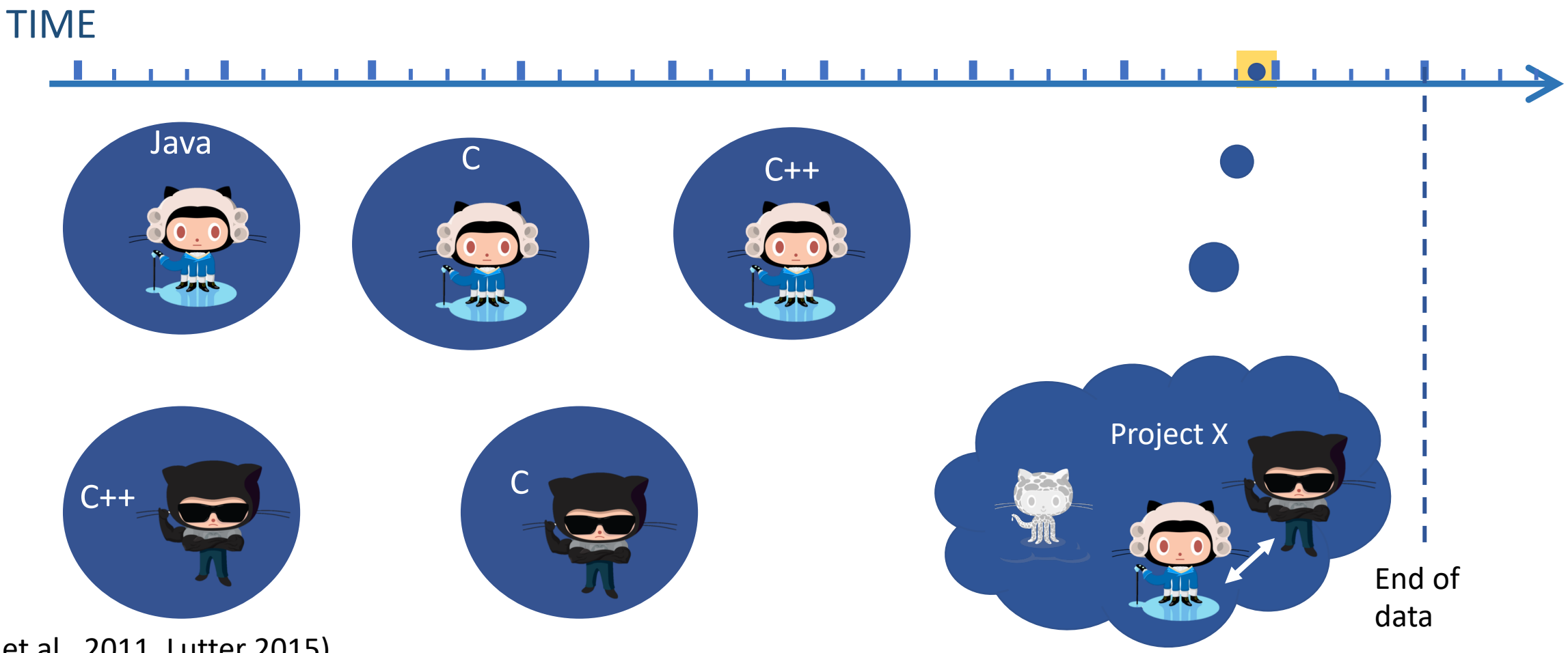
(de Vaan et al., 2011, Lutter 2015)

Bonding social capital – Recurring Cohesion



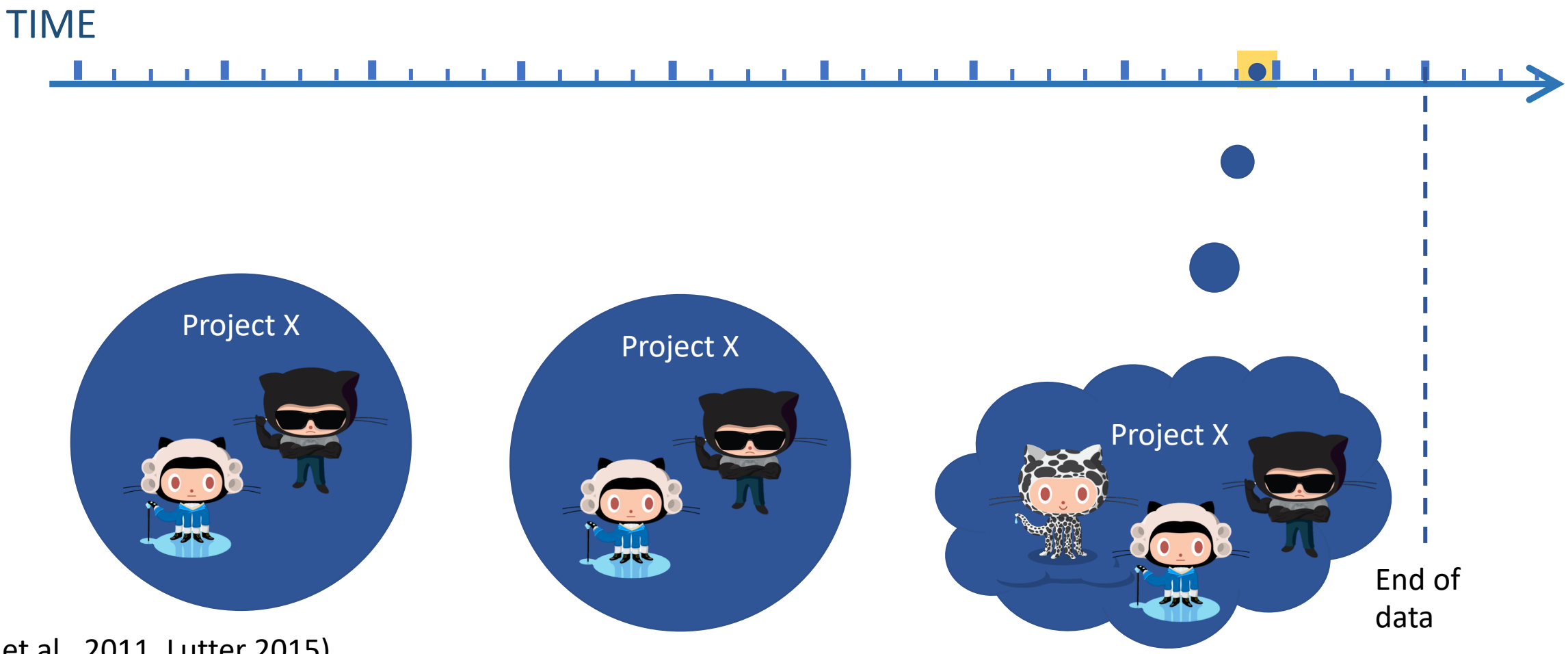
(de Vaan et al., 2011, Lutter 2015)

Bridging social capital – Language Diversity



(de Vaan et al., 2011, Lutter 2015)

Bridging social capital – Share of Newcomers

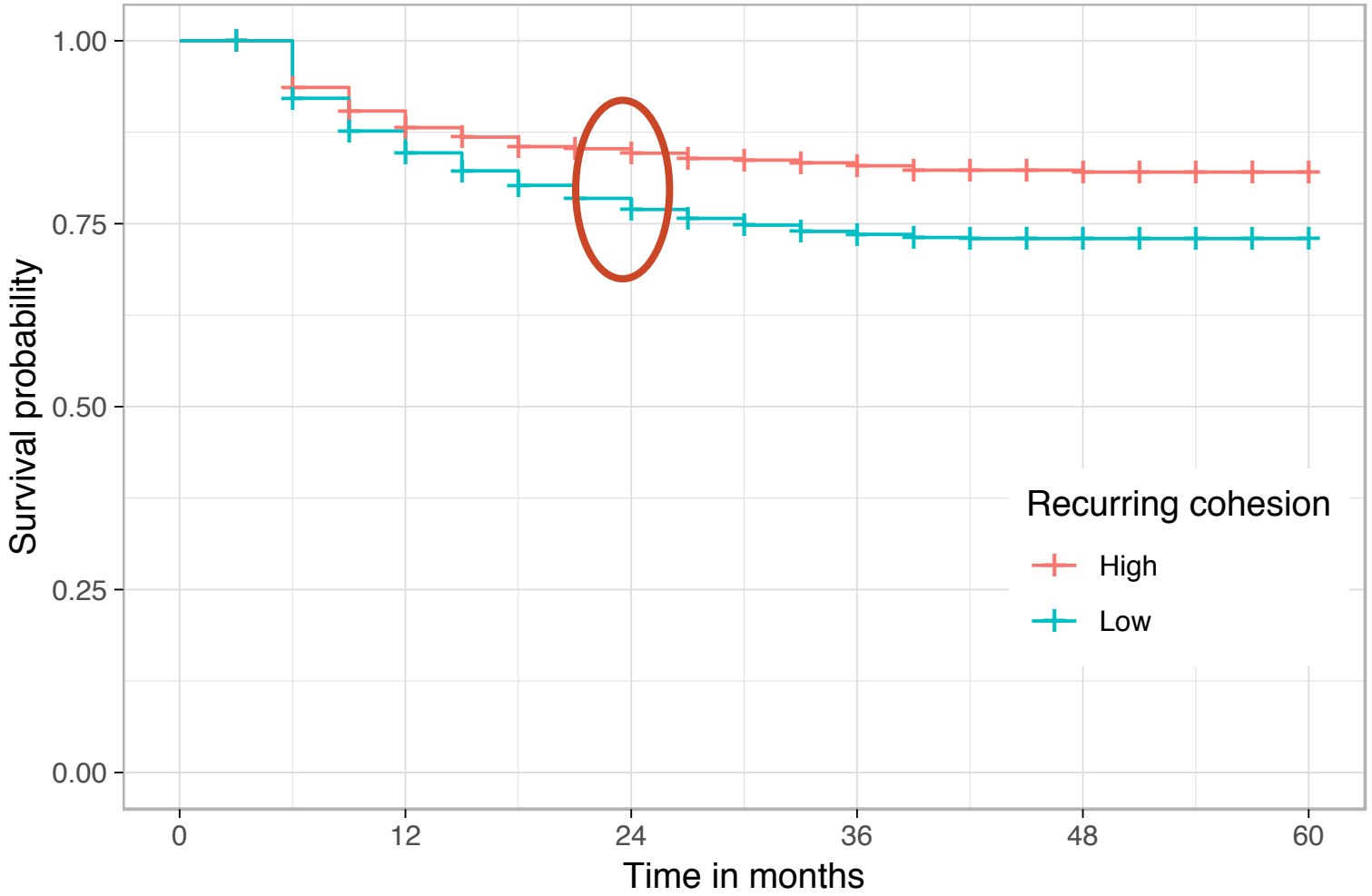


(de Vaan et al., 2011, Lutter 2015)

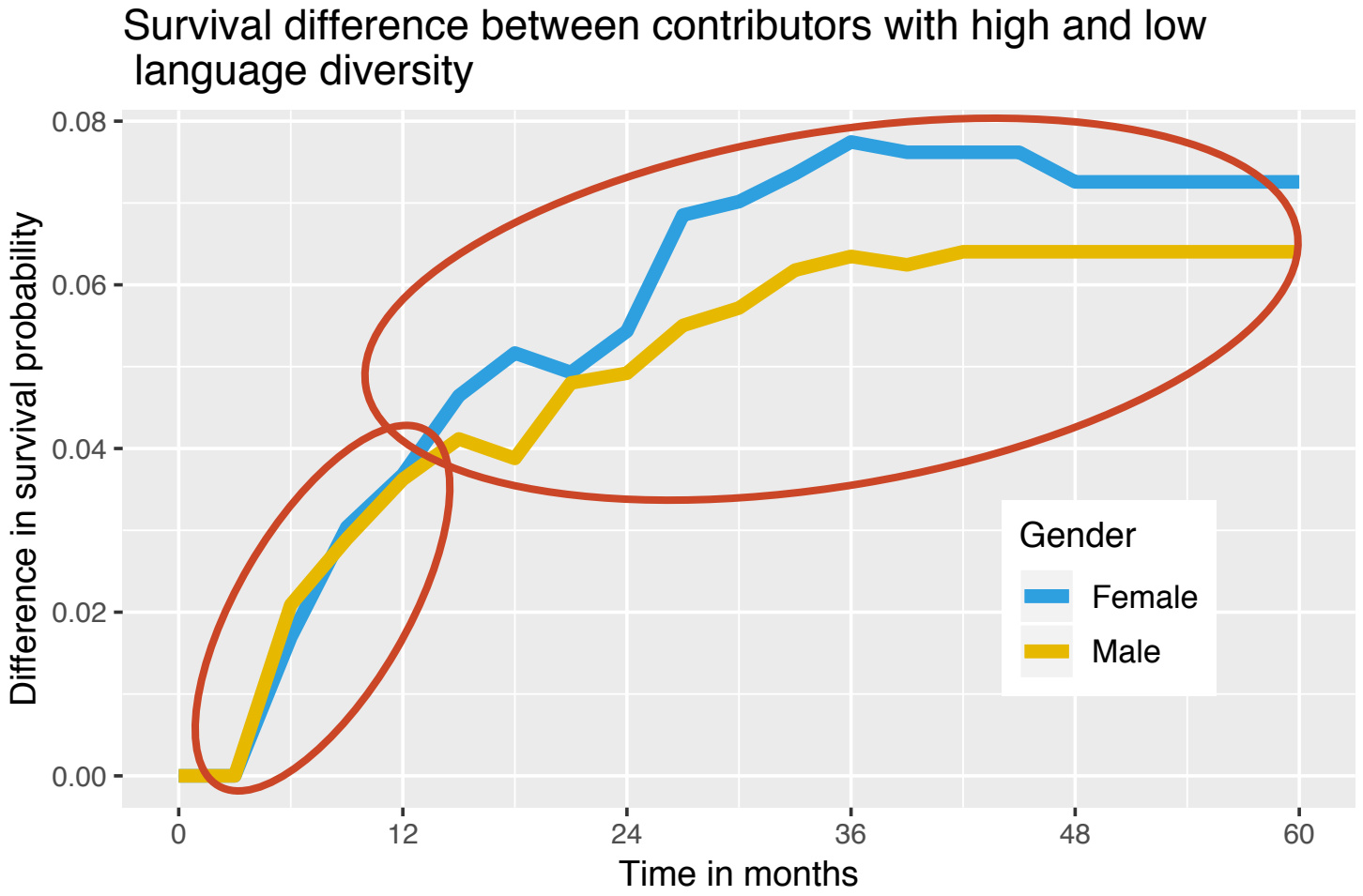
COX regression model

Contributor	Time	Active	Social capital	Control variables
	2008 Jan – Mar	True	Team familiarity Recurring cohesion Language diversity Share of newcomers	Project size Project owner ……
	2008 Jan – Mar	True	Team familiarity Recurring cohesion Language diversity Share of newcomers	Project size Project owner ……
	2009 Apr – Jun	False	Team familiarity Recurring cohesion Language diversity Share of newcomers	Project Size Not project owner ……

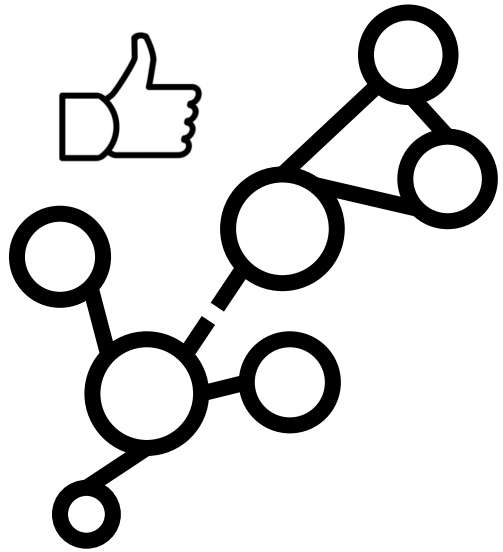
H1: more social capital ~ more prolonged engagement



H2: Language diversity interacts with gender

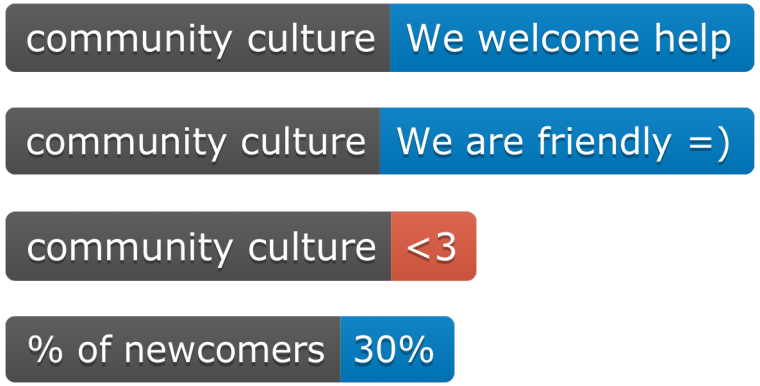


What next?



Recommend projects that can help build social capital

Offer mentorship

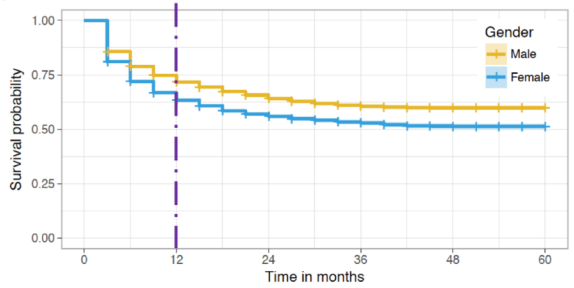


Use badges to show a community's culture

Social capital explains prolonged engagement

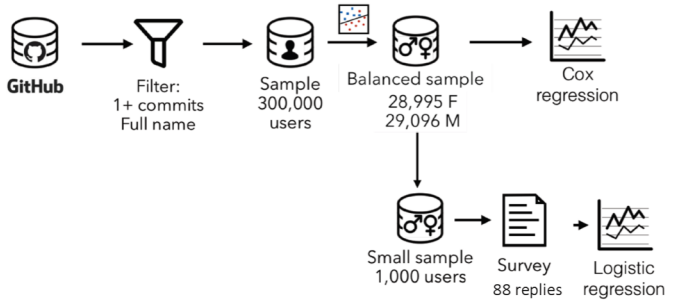
On GitHub, women disengage earlier than men

- After one year ca. 70% of men are still contributing to GitHub projects but only ca. 60% of women



Carnegie Mellon University School of Computer Science STRIDEL 17

Large-scale mixed-methods study

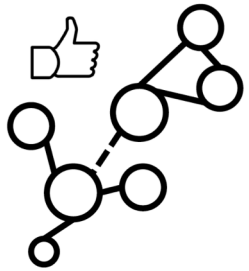


Carnegie Mellon University School of Computer Science STRIDEL 19

Social capital explains sustained participation

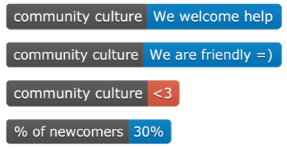
Carnegie Mellon University School of Computer Science STRIDEL 46

What can we do?



Recommend projects that can build social capital

Offer mentorship



Use badges to show a community's culture

Carnegie Mellon University School of Computer Science STRIDEL 47

Acknowledgements:

Contact: Sophie Qiu

@sophiehsqq

hsqq@cmu.edu

@alexander_nolte

alexander.nolte@udo.edu

@aserebrenik

a.serebrenik@tue.nl

@b_vasilescu

vasilescu@cmu.edu

Code and data:

DOI 10.5281/zenodo.2566684