Track 1 Lightning Talk: Forking as a Tool for Software Sustainability—An Empirical Study.

SARAH ALHOZAIMY
ROBERT HAINES AND CAROLINE JAY
SCHOOL OF COMPUTER SCIENCE
UNIVERSITY OF MANCHESTER
SEPTEMBER 2017
Ensuring sustainability has been identified as one of the key challenges in the development of software and how to achieve it remains an open question.

SSI defines 'sustainable software' as that which 'you use today will be available—and continue to be improved and supported—in the future'.
PURPOSE

Examines the link between the use of *forking*— the process of cloning a repository, such that development can progress separately to the original project—and the active life of the project.
SOFTWARE SUSTAINMENT DEFINITIONS

Sustainment of the original project, $S$, is defined as the time period from the first commit of a repository through to its last commit, measured in days.

$$S = t_{\text{last-commit}} - t_{\text{initial-commit}}$$
SOFTWARE SUSTAINMENT DEFINITIONS

- Software sustainment with forking, SF, is defined as the number of days from the first commit of the original repository through to the last commit when considering all forks.

\[ SF = t_{\text{last-commit-on-all-forks}} - t_{\text{initial-commit}} \]
9,118 projects retrieved from GitHub with 71 different languages.

GitHub projects selection criteria:
- The project was created between 1st January and 31st December 2009.
- Five continuous days from each month were sampled.
- The project had at least one commit.
- The first commit occurred on or after 1st of January 2009.
FORK TREE

- **Depth 0:** $F_0$
- **Depth 1:** $F_{10}$, $F_{11}$, $F_{12}$
- **Depth 2:** $F_{20}$, $F_{21}$, $F_{22}$
Is forking related to sustainment of the original project?

Mean sustainment $S$:
- $\text{Forks} \geq 1 \rightarrow 788 \text{ days}$
- $\text{Forks} = 0 \rightarrow 303 \text{ days}$
Does forking sustain software beyond the original project?

Mean sustainment SF:
- Forks ≥ 1 → 1,111 days
- Forks = 0 → 303 days
Conclusion

- Forking is positively related to software sustainment.

- Forking helps to extend the active life of the original project and beyond the original project.

- Open source repositories such as GitHub that support forking may play a key role in supporting the sustainability of software.
Thank You!


