

# Influence in Software Ecosystem and Software Communities

Jingchang Chen

A dark blue diagonal gradient bar that starts from the bottom left corner and extends towards the top right corner, covering the lower half of the slide.

# What is a software ecosystem (SECO)?

An evolution from software product line. (Bosch, 2009)

# What is a software ecosystem?

*“A set of businesses functioning as a unit and interacting with a shared market for software and services, together with the relationships among them. These relationships are frequently under-pinned by a common technological platform or market and operate through the exchange of information, resources and artifacts” (Jansen et al., 2009)*

*“A software ecosystem consists of the set of software solutions that enable, support and automate the activities and transactions by the actors in the associated social or business ecosystem and the organizations that provide these solutions” (Bosch, 2009)*

A set of connected services based on a common technology foundation in a shared market.

# Real-world examples

Microsoft Windows

Apple iOS, MacOS

Linux

Google Chrome

...

VS Code

React

USB Standard

# Why SECO is popular?

## Benefits (Barbosa & Alves, 2011):

- Reduce the (individual) development cost
- Welcome new players
- Accelerate the evolution
- Create new markets...

Especially good for small to medium enterprises, which is the majority in the industry (Fischer et al., 2000)

## Challenge:

- Requirement engineering process - Influence matters! (Schaarschmidt et al., 2015; Linåker et al., 2019; Damian et al., 2021)

# What brings influence?

There is a tie between contribution and influence, especially in open-source software communities.

- Lines of code
- Number of pull requests

Overall, changes to the source code

# Methodology, Data Source, Tools



# Methodology

De

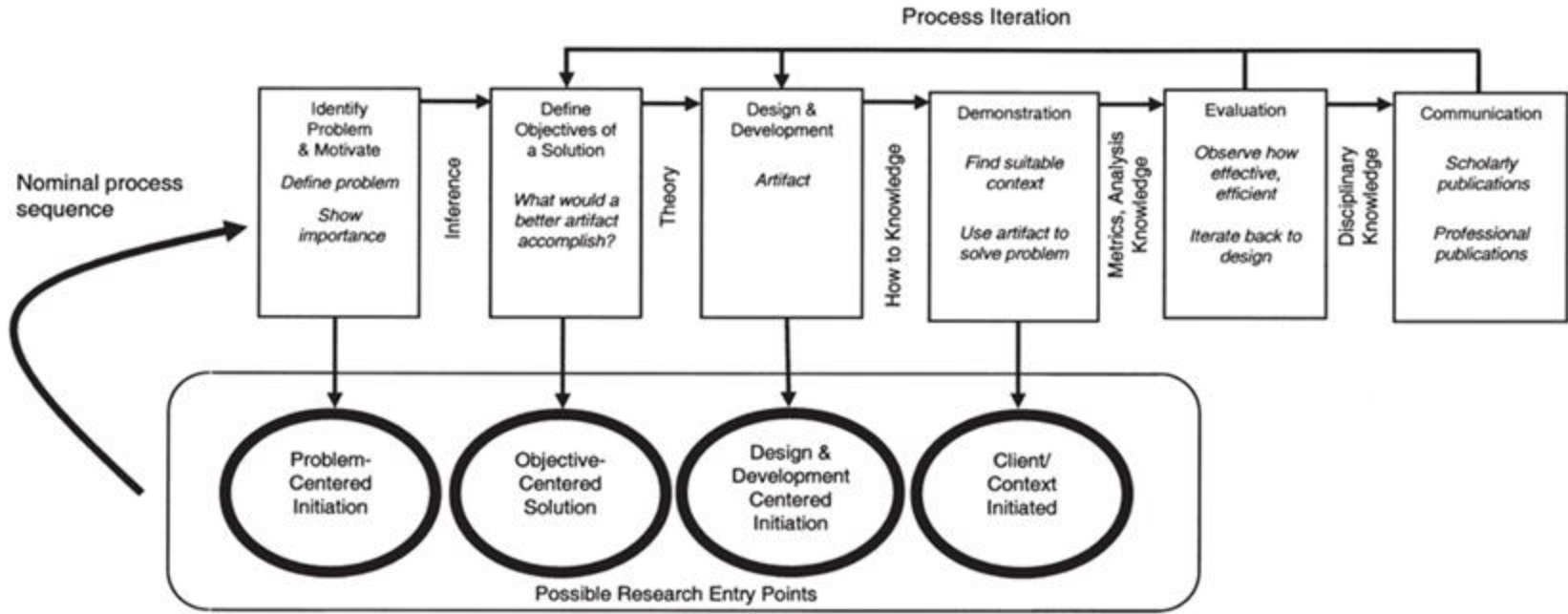


Figure 1. DSRM Process Model

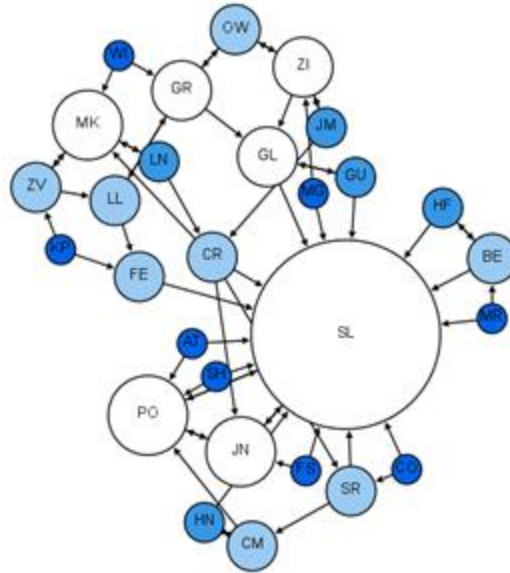


# Data Source

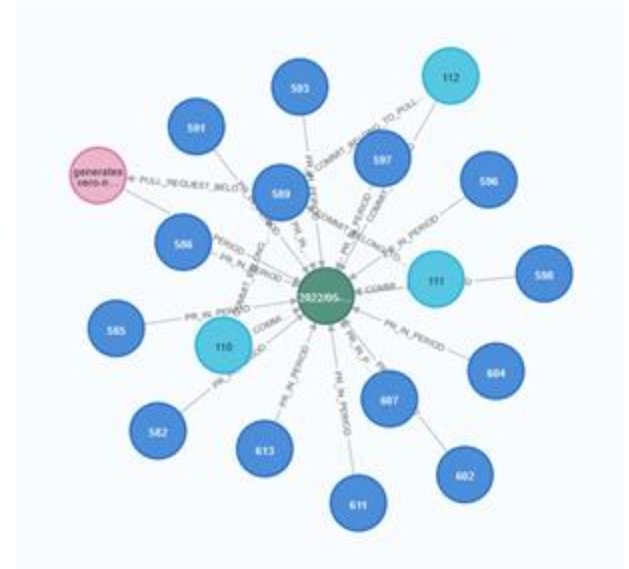
- Private
- Open-source
  - GitHub
  - GitTorrent (Gousios & Spinellis, 2012)
  - GHArchive

# Tools

- Social Network Analysis
  - A tool to study relationships
- Neo4j
  - A graphic relational database



\*\* A social network. Cited from Wikipedia



# My study



# Xero

## 1. Breaking down its structure



## 2. Identifying available public data sources

- Official website.
- Developer forum (deprecated).
- User forum.
- Xero API customer feedback platform.
- Application store.
- Official documents.
- GitHub repositories.

# Xero

## 3. Mining GitHub

- GitHub RESTful API

## 4. Building networks

- Node: Individual/Organisation
- Edge: Both commit in one patch (pull request)
- Aggregated by time period (versions/tags)



# Xero

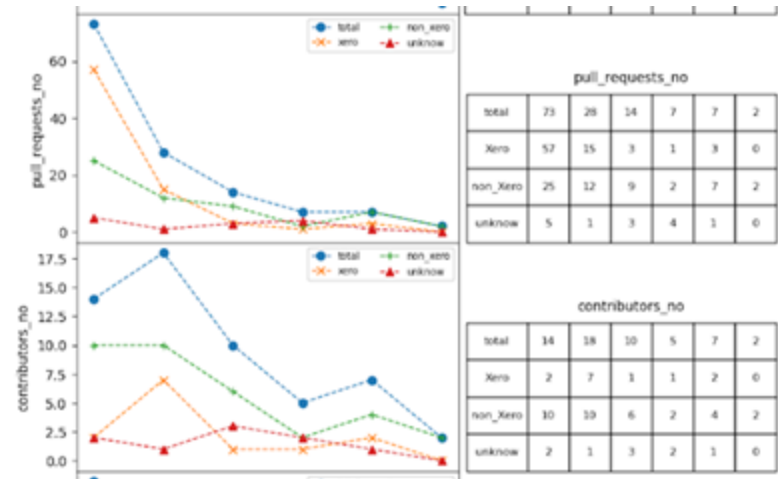
Results:

1. A large number of outsourcers are hired

Period	Orchestrator	Niche Player	End User	External Developer	Outsource	Unclear
2017/05-2017/10	1	1	2	0	6	1
2017/11-2018/04	1	1	0	2	4	3
2018/05-2018/10	1	0	1	0	1	3
2018/11-2019/04	1	0	0	0	2	0
2019/05-2019/10	1	0	1	1	1	1
2019/11-2020/04	0	0	1	0	1	0
Total	1	2	4	3	12	8

1. Most contributors are short-term contributors, focusing on the issue they encountered

3. In an SDK project, the amount of contribution and contributors decreases when the product becomes more mature



# React

## 1. Mining React GitHub repo

- Use GHArchive and GitTorrent to recover the missing pieces

## 2. Building social network

- Node: Individual
- Edge: Comment in the same pull request
- Aggregated by time period (versions/tags)
- Strong tie VS weak tie



Strong Tie = Reciprocated PR Comment (7ppl)  
Weak Tie = Non-Reciprocated PR Comment (7ppl)



## 3. Finding cliques

- Bron-Kerbosch algorithm

## 4. Labeling comments

- Manually label (code) a part of them
- Train machine learning model to label the rest

Raw Quote Example	Theme	Relevant Codes
"You guys told me to do the sketchy hack instead of always reflecting the correct value. I don't like the solution and sebastian advocated for, but I needed to choose my battles in the interest of moving this diff forward."	Conflict	Disagreement, Disappointment, Doubt, Highlighting Issues, Correction, Contrasting Needs or Perspective
"I've opened #2271 to track the status of 'setInnerHTML' vs 'createNodesFromMarkup'"	Project Workflow	Process, Making a Plan, Labelling, Taking Responsibility, Assignment, Ending Discussions, Redirection
"@jsh you made this change. What did you mean?"	Interdependence	Relying on Social Connections, Accountability, Accepting Work, Reference, Demonstration
"I know you are just following what's already here, but I think we should consolidate these constants."	Improvement	Giving Advice, Giving Clarification, Agreement, Refinement, Making Suggestions, Gratitude
"Should you mind explaining what this is good for?"	Inquiry	Seeking Advice, Seeking Clarification, Asking for Help
"rusted_hands: wooooo!"	Miscellaneous	-

# React

Results:

1. Conflicts exist in discussions in higher diverse teams

(Diversity Blau Index Score)	Ethnic	Nationality	Gender
<i>Conflict Present-(23/39)</i>	0.096	0.515	0.152
<i>Conflict Absent-(16/39)</i>	0.041	0.486	0.145

1. These conflicts are usually beneficial (productive conflicts) (Schulz-Hardt et al., 2002), leading to more merged pull requests

Collaborative Groups	Conflict	No Conflict
<i>Average Merged Pull Requests / Group</i>	172.00	53.00
<i>Average Merged Pull Requests / Group / Version</i>	20.31	10.75

3. The teams with conflicts also have more interdependence and inquiry comments in discussions, which shows a higher level of psychological safety

	Interdependence	Inquiry	p value
<i>Groups (Conflict Present)</i>	16	971	0.000
<i>Groups (Conflict Absent)</i>	0	194	0.000

3. These teams also last longer, compared with groups with lower diversity and low/no conflicts



Thanks for watching!

# References

- J. Bosch, "From software product lines to software ecosystems," in *Proceedings of the 13th International Software Product Line Conference*, 2009, vol. 9, pp. 111-119.
- S. Jansen, A. Finkelstein, and S. Brinkkemper, "A sense of community: A research agenda for software ecosystems," in *2009 31st International Conference on Software Engineering - Companion Volume*, 16-24 May 2009 2009, pp. 187-190, doi: 10.1109/ICSE-COMPANION.2009.5070978.
- O. Barbosa and C. Alves, "A systematic mapping study on software ecosystems," in *Proceedings of the Third International Workshop on Software Ecosystems*, 2011, pp. 15-26.
- E. Fischer, E. Fisher, and R. Reuber, Industrial clusters and SME promotion in developing countries (no. 3). *Commonwealth Secretariat*, 2000.
- M. Schaarschmidt, G. Walsh, and H. F. O. von Kortzfleisch, "How do firms influence open source software communities? A framework and empirical analysis of different governance modes," *Information and Organization*, vol. 25, no. 2, pp. 99-114, 2015/04/01/ 2015, doi: <https://doi.org/10.1016/j.infoandorg.2015.03.001>.
- J. Linåker, B. Regnell, and D. Damian, "A Community Strategy Framework—How to obtain influence on requirements in meritocratic open source software communities?," *Information and Software Technology*, vol. 112, pp. 102-114, 2019.
- D. Damian, J. Linåker, D. Johnson, T. Clear, and K. Blincoe, "Challenges and Strategies for Managing Requirements Selection in Software Ecosystems," *IEEE Software*, vol. 38, no. 6, pp. 76-87, 2021.
- K. Peffers, T. Tuunanen, M. A. Rothenberger, and S. Chatterjee, "A design science research methodology for information system s research," *Journal of management information systems*, vol. 24, no. 3, pp. 45-77, 2007.
- A. Marin and B. Wellman, "Social network analysis: An introduction," *The SAGE handbook of social network analysis*, pp. 11-25, 2011.
- S. Schulz-Hardt, M. Jochims, and D. Frey, "Productive conflict in group decision making: Genuine and contrived dissent as strategies to counteract biased information seeking," *Organizational Behavior and Human Decision Processes*, vol. 88, no. 2, pp. 563–586, 2002.