The who, what and how of the current research at the Brazilian Symposium on Software Engineering

Lidiany Cerqueira
UFBA - UFS
lidiany.cerqueira@ufba.br

Júlio César A. Silva
UEFS

Ícaro V. Alvim
UEFS

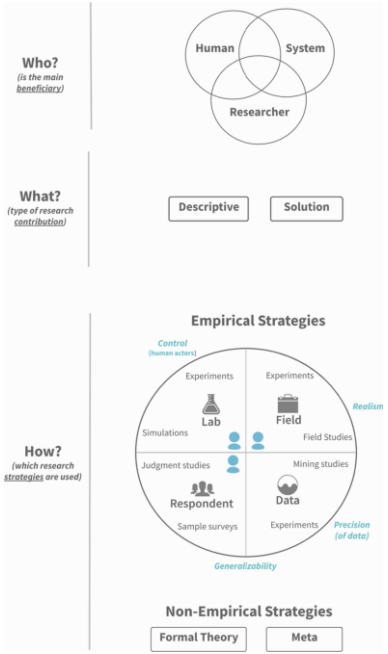
Manoel Mendonça
UFBA

José Amancio M. Santos
UEFS

Background

- How empirical software engineering research papers address human and social aspects?
  - a Socio-technical Framework: Who - What – How

- Software engineering is a socio-technical activity:
  - many contributions focus on technical aspects;
  - the impact of human factors on software development is still poorly understood.



Margaret-Anne Storey, Neil A Ernst, Courtney Williams, and Eirini Kalliamvakou. 2020. The who, what, how of software engineering research: a socio-technical framework. Empirical Software Engineering 25, 5 (2020), 4097–4129

Goal

- Analyze research published on the SBES'19-21 to understand how these works considers human aspects using the Who-What-How framework.

Research questions

RQ1: Who are the beneficiaries (technical systems, human stakeholders, researchers) of the research contributions?

RQ2: What is the predominant type of research contribution (descriptive or solution) provided?

RQ3: What strategies did the researchers use?

RQ4: How do the reported research strategies map to the beneficiary and types of contributions in these papers?

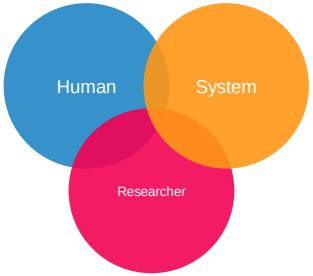
RQ5: How do the results compare to the original study?

RQ6: Was the framework helpful, and can it be used in different contexts?

Coding

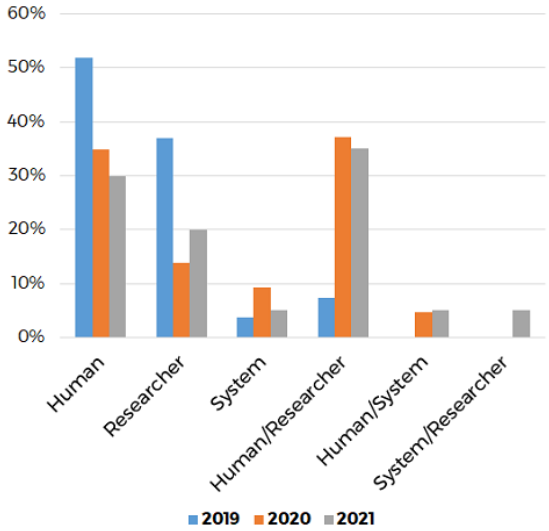
- We selected 90 papers from SBES Research track.
- Four researchers coded the papers individually, according to the framework.
- We then discussed the result together to reach a consensus.

Findings – RQ1: Who

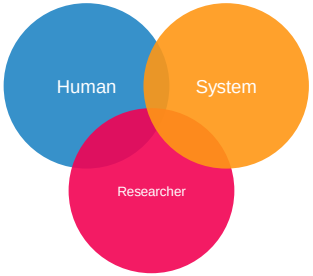


- Single beneficiaries

	Human	Researcher	System
2019	52%	37%	4%
2020	35%	14%	9%
2021	30%	20%	5%
	39%	22%	7%

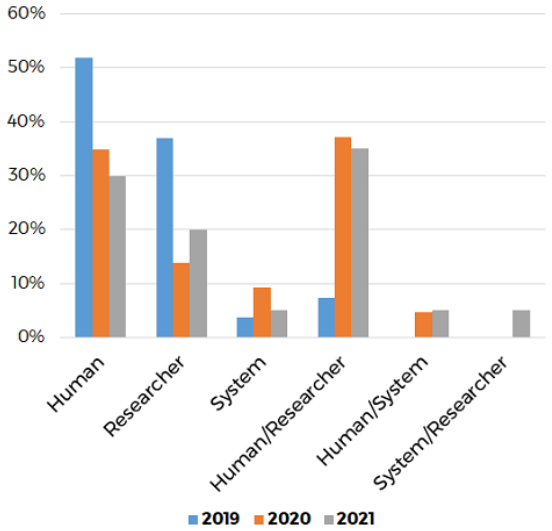


Findings – RQ1: Who



- Multiple beneficiaries

	Human-Researcher	Human-System	System-Researcher
2019	7%		
2020	37%	5%	
2021	35%	5%	5%

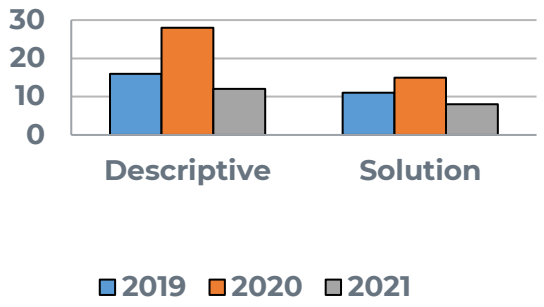


Descriptive | Solution

Findings – RQ2: What

- Most papers published in all three years present descriptive works.
- SBES has a track that focuses on tools, while we covered only the research track.
- This may be related to most works being coded as descriptive as well as to discovering fewer works benefiting a system.

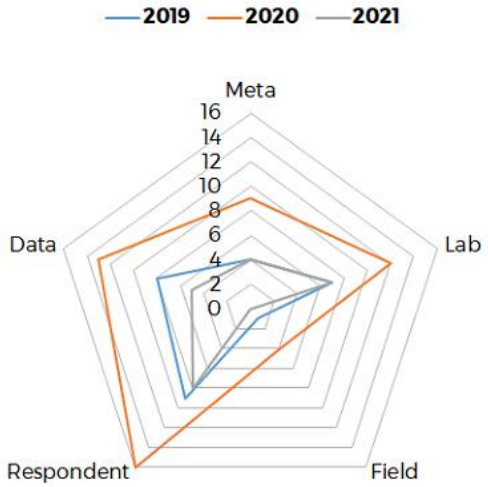
	Descriptive	Solutions
2019	59%	41%
2020	65%	35%
2021	62%	38%
	62%	38%



Findings – RQ3: How

- The most adopted research strategies are:
  - Respondent, Data, Lab, Meta and Field.

	Respondent	Data	Lab	Meta	Field
2019	9	8	7	4	1
2020	16	13	12	9	4
2021	8	5	7	4	0



- We did not find any papers using formal theory strategy.
- We found higher use of respondent strategies.

Discussion

- There are benefits and drawbacks when choosing these strategies:
  - data strategies - easier to replicate
  - respondent - easily scaled to larger populations
- Most works adopting respondent strategies used surveys and interviews:
  - higher generalizability;
  - ease of creating and distributing online questionnaires.

Discussion

- Data mining studies used Github or StackOverflow as a data source.
- None of the works used **Formal theory** as a research strategy:
  we can suppose that researchers are still **trying to understand and describe** the problems and **not yet proposing new theories**.

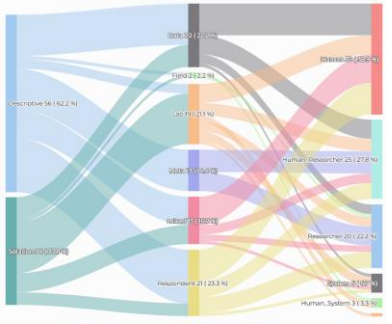
- We did not find any field studies in 2021:
  - Field studies can be more difficult to set up and run;
  - Possible change in the research context caused by the COVID-19 pandemic;
  - Field studies may have become even more difficult to execute;
  - Relationship between software development industry and academic research groups.

Findings - Triangulation

- How these studies used different strategies or methods.
- 15 papers (16,7%) used mixed strategies.

1st	2nd	3rd	Count	Ref
Data	Lab	Respondent	1	[48]
Data	Lab		2	[10,46]
Data	Respondent		2	[28,30]
Field	Respondent		2	[15,33]
Lab	Respondent		3	[22,34,36]
Meta	Lab		1	[17]
Meta	Respondent		3	[7,21,29]
Respondent	Field		1	[42]

Findings – RQ4 - Mapping



- Few studies describing solutions (what) have researcher (who) as a single beneficiary (3.3%).
- Considering works benefiting both humans and researchers, this number increases to 10%.
- Most studies using lab strategy (how) aims to present a solution (what).
- 54% of the papers who use the strategy meta (how) also focus on human (who) and aim to describe (what) the nature of problems.
- Over 46% of studies claim to benefit humans in some way, but use research strategies such as data or lab.

Findings – RQ5 - Comparison

- The authors of the framework analyzed 151 papers from the ICSE and the EMSE:
  - We found more respondent studies (they discovered a higher use of data strategies).
  - We found more descriptive studies (62%).
  - They found more works adopting formal theory than meta and field strategies.
  - The original study also detected few studies triangulating their research strategies.

Findings – RQ6 – Framework evaluation

- We considered that the framework improves understanding of research contributions.

- The framework can be adopted in different contexts.

Conclusions

- We present a replication study, using the Who-What-How framework to analyze papers from SBES'19-21 to assess how these papers consider or study social aspects.
- We focused our analysis on the Research track.
- While coding, we found some works that did not explicitly identify the research beneficiaries.
- Positioning the studies according to the framework can benefit future research.
- Our findings can provide some answers and induce discussions about research method choice and human involvement on software engineering research.
- Our work also offers insights into how the Brazilian research community is addressing social and human aspects in their research.

Thanks!

Any questions?

lidiany.cerqueira@ufba.br

References

- https://dl.acm.org/doi/10.1145/3555228.3555241