

How to Identify Programming Skills from Source Code?

Johnatan Oliveira, Eduardo Figueiredo

Department of Computer Science – UFMG, Brazil

Institute of Science, Technology and Innovation – UFLA/ICTIN, Brazil

johnatan.oliveira@ufla.br, figueiredo@dcc.ufmg.br

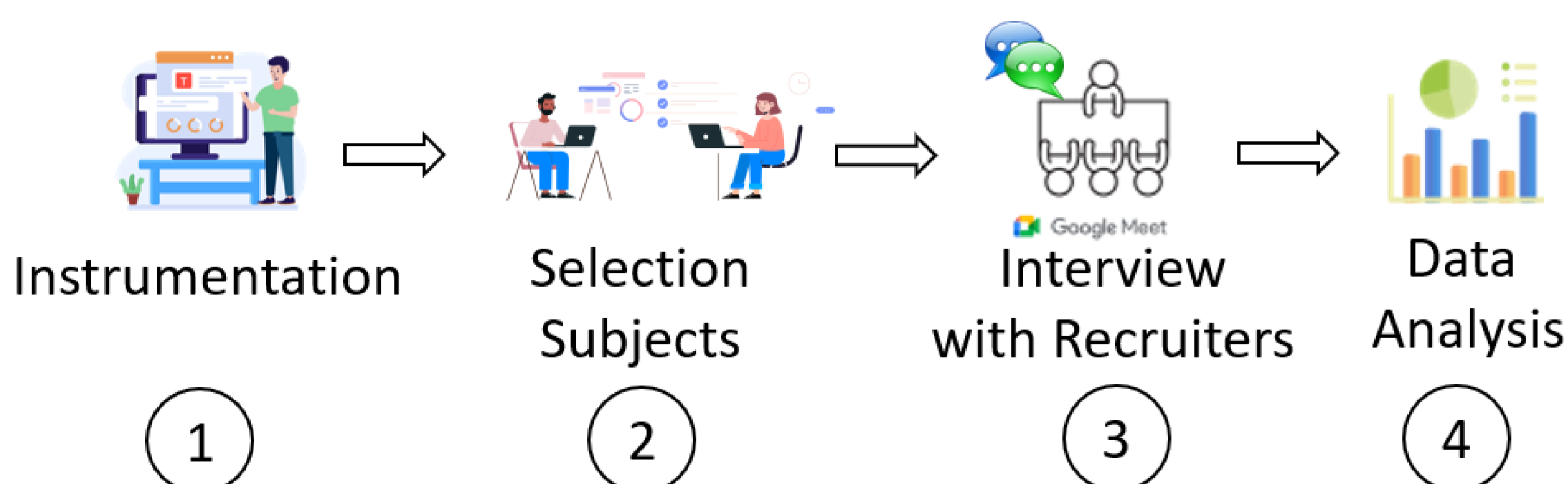
Motivation

The increasing complexity of software systems demands high quality and puts pressure on developers [1]. Identifying and hiring qualified developers is crucial, as poor hiring decisions can negatively impact project success [2]. Social coding platforms like GitHub, provide data to identify skills and expertise, optimizing team formation [3]. Despite their widespread adoption, limited research has explored the effective hard skill detection.

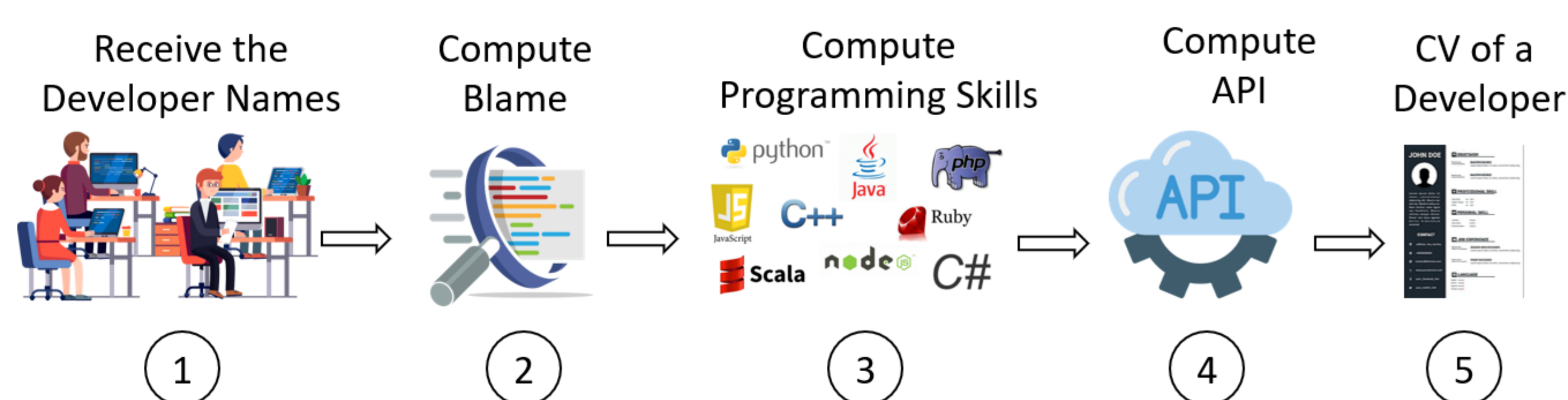
Research Questions

- RQ1-How do companies recruit software developers?
- RQ2-What are the possible application of the method to mine a developer profile?
- RQ3-How do the method results complement GitHub information?

Understanding the Hiring Process



Mining Programming Skills from GitHub



Contributions

- Comprehensive Understanding of GitHub Profiles [3].
- Reduction of Recruitment Effort [4].
- Enhanced Evaluation Metrics [2].
- Identification of Developer Profiles [1,3].
- Detection of Profile Pollution [3,4].

Results

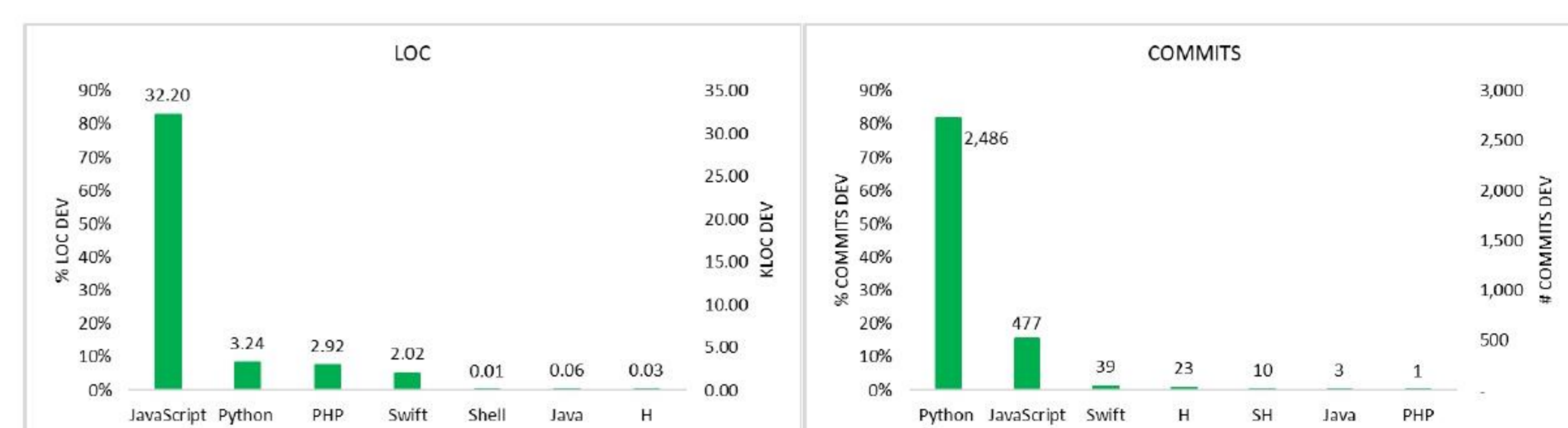
John
Software developer

SHORT BIO

Full Stack Python. John currently works in Washington, D.C. Developer Network as the Director of Developer Content.

GITHUB REPOSITORIES: 106 **GITHUB SINCE: 2010**

PROGRAMMING SKILLS



Most used API by developer *

JavaScript
Angular (337), React (323), Vue.js (124), Ember.js (105), Node.js (95), Backbone.js (71)

Python
AIOHTTP (9), Bottle (4), CubicWeb (3)

Conclusion

Selecting the right software developer is crucial for ensuring quality and timely delivery in projects. A good developer possesses technical skills, collaborates effectively, and identifies issues early. This study highlights the importance of combining hard and soft skills in recruitment. It suggests organizations invest in advanced tools and methods for analyzing GitHub profiles and standardize evaluation protocols.

Future Work

- Apply machine learning methods to identify software engineering skills on GitHub
- Investigate the most important technical skills for software developers in different domains.

References

- [1] Oliveira et al. (2020) [SBES].
- [2] Oliveira et al. (2021) [JSERD].
- [3] Oliveira et al. (2022) [CIbSE].
- [4] Oliveira et al. (2022) [QUATIC].