

# A Theory of Influence in Software Ecosystems: A Developer's Perspective

Vinicius Serva Pereira<sup>1</sup>, Rodrigo Pereira dos Santos<sup>1</sup>

<sup>1</sup>Programa de Pós-Graduação em Informática  
Universidade Federal do Estado do Rio de Janeiro (UNIRIO)  
Av. Pasteur, 458 – Rio de Janeiro – RJ – CEP: 22290-240 – Brazil

{vinicius.pereira, rps}@uniriotech.br

***Abstract.** The social perspective is increasingly presented in software development. As a consequence, influence emerges as a relevant factor in the interaction between developers using the same technological platform, the so-called software ecosystems (SECO). Influence can be considered as the power that something or someone can have over something or a person and can be a determining factor for the success or failure of a SECO. This work aims to present a Master research proposal for a theory to explain how influence, as a power, manifests among developers in SECO. We intend to build the theory based on empirical evidence from our studies in GitHub SECO, in light of the theoretical framework underpinned by French and Raven's power taxonomy.*

## 1. Introduction

The collaborative perspective in software development does not have a linear evolution of technical and social characteristics. Among social factors, a critical factor is influence and it involves influencers - those who can start leading development and dictate how the project will progress. The term influence can be understood as a kind of power, where having power over someone is to make someone do something he/she would not otherwise do [Valença and Alves 2016].

With the study of influence, it is possible to understand the evolution of software ecosystems (SECO). A good influence can lead a project to a state of success, just as a bad influence could lead to a SECO failure [Carver et al. 2017]. Using GitHub projects as the setting of this research, a SECO refers to a collection of software projects that are developed and evolved together in the same environment [Lungu et al. 2010], as can be seen in the npm SECO and RubyGems SECO, which are formed by GitHub projects related to their libraries and packages [Constantinou and Mens 2017].

In this context, this Master research main objective is to understand the dynamics of influence among developers by the analysis of existing SECO in GitHub. To that extent, influence will be conceived as a kind of power in light of works developed by French et al. (1959) and Emerson (1962) about power taxonomy. As a result of our research, we intend to develop a substantive theory to explain how the power of influence affects the relationships of developers within a SECO.

This work is organized as follows: Section 2 describes a general view of the problem being studied by this research; Section 3 presents the research method and design; Section 4 addresses how this work will be evaluated; Section 5 reports some work we have done; finally, Section 6 presents the research proposal conclusion.

## 2. Problem Description

Distributed social coding is an increasing trend in software development. Developers who are considered to be important enough to dictate and lead development on how the project evolves are called influencers and influence is considered one of the critical factors for a SECO success or failure [Carver et al. 2017]. Therefore, a better understanding of the actors influencing others is a relevant factor for SECO project owners and managers, who are responsible for monitoring ecosystem health, the value of open source software and community trust on the platform to attract developers and users. The study of influence in the GitHub platform is a recent and relevant research subject, as can be seen by the works of Bana (2018), Hue et al. (2018), Liao et al. (2018). Despite the attention that the subject has been receiving, the main objective of those works are related to classifying the level of influence based on quantitative aspects of the developers and their network, leaving aside the comprehension of how and why the influence is affecting users.

Influence can be interpreted as a kind of power, where to have power over someone is to get someone to do something that he/she would not otherwise do [Valença and Alves 2016]. As a social power, influence can have different power forms, bases and sources within each relationship. According to the power taxonomy proposed by French et al. (1959), there are five different power forms: coercive, legitimate, expert, referent and reward, where power can behave differently within each form. This research intends to understand and theorize, in the context of SECO, what those sources of influence are, how they are connected to each power form and why they behave as they do.

## 3. Research Method

Before introducing our research methodology, it is important to state the main research question we have proposed: *How does influence take place in interactions among developers participating in a software ecosystem?*. The research methodology adopted in this work derives mainly from the work of Valença and Alves (2017), in which a theory of power between small-to-medium enterprises in SECO was proposed. To conceive the theory, we will use the building strategy defined by Eisenhardt (1989), which presents a roadmap to build theories based on empirical evidence from case studies. This process involves the use of one or more cases to create theoretical constructs, propositions and/or mid-range theory from empirical evidence. The proposed theory is intended to be classified as an explanation theory, which describes how and why a phenomenon happens [Gregor 2006].

We extend the following research design from Valença and Alves (2017), based on four sequential phases:

1. Definition and design: In this phase, we will build the conceptual framework that supports the proposed research, define the data collection protocol that will be used to extract the empirical evidence and select the SECO that will be analyzed;
2. Execution: The collection of the data will be performed by the means of surveys and semi-structured interviews with relevant developers from the selected ecosystem. The results will be analyzed and the power sources and forms will be identified using the coding procedures presented in the Grounded Theory method;

3. Evaluation: The main goal of this phase is to analyze complementary and confirmatory evidence obtained in the previous phase, allowing to refine and increase the accuracy of our study;
4. Conclusion: Finally, in this phase, the proposed theory will be formally built. Using the SECO cases selected in the first phase, we intend to create theoretical constructs and propositions from the collected evidence by using the roadmap proposed by Eisenhardt (1989).

#### **4. Evaluation**

According to Sjöberg et al.(2008), there are, there are six criteria suitable for evaluating empirically based theories, which is the case of the theory proposed by this work. We intend to rate as low, moderate and high the degree to which our theory meets each criterion after its proposition. The criteria are the following:

1. Testability: Refers to the degree to which a theory is formulated so that empirical testing is possible;
2. Empirical support: The degree to which the theory is supported by empirical studies that confirm its validity;
3. Explanatory power: The degree to which a theory explains all known observations within its scope;
4. Parsimony: Refers to the degree to which a theory is economically constructed with minimum concepts and propositions;
5. Generality: Evaluates the independence of the theory within a specific setting;
6. Utility: Refers to the degree to which the theory supports the relevant areas of its context.

#### **5. Performed Activities**

This research is currently at its initial stage (Definition and design phase), as proposed by the research design stated in Section 3. Currently, we have performed a literature review, which identified that the current studies were focused on classifying the level of influence based solely on quantitative aspects regarding data on developers and their network, leaving aside the comprehension of how and why the influence is affecting users, as stated in Section 2.

In addition, we have already identified some socio-technical characteristics and perceptions of developers that characterize influence in a SECO [Farias et al. 2019], from which we will perform coding procedures to elaborate the semi-structured interviews. The results will also be used as reference for the sources of power that influence can possess. Finally, we have defined the cases that will be analyzed, which will be npm SECO and RubyGems SECO, as both were classified as ecosystems by previous research [Constantinou and Mens 2017] and allow to collect socio-technical characteristics from data available at GitHub.

#### **6. Final Remarks**

Influence is a relevant factor in the understanding of the evolution of SECO. A good influence can lead a project to a state of success, just as a bad influence could lead to a SECO failure. However, current studies focus mainly on classifying the level of influence

based solely on quantitative aspects regarding data on developers and their network, leaving aside the comprehension of how and why the influence occurs. Thus, as proposed by this research, we believe that a new specific theory that better explains the dynamics of influence among developers in SECO may have its place in literature.

## Referências

- Bana, R. (2018). Influence Indexing of Developers , Repositories , Technologies and Programming languages on Social Coding Community GitHub. *2018 Eleventh International Conference on Contemporary Computing (IC3)*, pages 1–6.
- Carver, J. C., Muccini, H., and Yamashita, A. (2017). Distributed Teams, Developer Participation, and More. *IEEE Software*, 34(3):114–116.
- Constantinou, E. and Mens, T. (2017). An empirical comparison of developer retention in the RubyGems and npm software ecosystems. *Innovations in Systems and Software Engineering*, 13(2-3):101–115.
- Eisenhardt, K. M. (1989). Building Theories from Case Study. *Academy of Management Review*, 14(4):532–550.
- Emerson, R. M. (1962). Power-Dependence Relations. *American Sociological Review*, 27(1):31.
- Farias, V., Wiese, I., and Santos, R. (2019). What Characterizes an Influencer in Software Ecosystems? *IEEE Software*, 36(1):42–47.
- French, J. R. P., Raven, B., and Cartwright, D. (1959). The bases of Social Power. In Cartwright, D., editor, *Studies in Social Power*, pages 311–320.
- Gregor (2006). The Nature of Theory in Information Systems. *MIS Quarterly*, 30(3):611.
- Hu, Y., Wang, S., Ren, Y., and Choo, K. K. R. (2018). User influence analysis for Github developer social networks. *Expert Systems with Applications*, 108:108–118.
- Liao, Z., Jin, H., Li, Y., Zhao, B., Wu, J., and Shengzong, L. (2018). DevRank: Mining influential developers in Github. *2017 IEEE Global Communications Conference, GLOBECOM 2017 - Proceedings*, 2018-Janua(1):1–6.
- Lungu, M., Lanza, M., Gîrba, T., and Robbes, R. (2010). The Small Project Observatory: Visualizing software ecosystems. *Science of Computer Programming*, 75(4):264–275.
- Sjoberg, D. I., Dybå, T., Anda, B. C., and Hannay, J. E. (2008). Building theories in software engineering. *Guide to Advanced Empirical Software Engineering*, pages 312–336.
- Valença, G. and Alves, C. (2016). Understanding how power influences business and requirements decisions in software ecosystems. In *Proceedings of the 31st Annual ACM Symposium on Applied Computing - SAC '16*, pages 1258–1263, New York, New York, USA. ACM Press.
- Valença, G. and Alves, C. (2017). A theory of power in emerging software ecosystems formed by small-to-medium enterprises. *Journal of Systems and Software*, 134:76–104.