**Leadership Behaviors in the Open Source Community**

Team Project Report

**Team 10**

Ashish Singh

Tanmay Bhoir

Anirudh Bezzam

# Summary

Open source model advocates open, collaborative, decentralized ownership and volunteer contributions. However, increasing growth in the adoption rates of open source software by organizations of all sizes makes an argument in favor of some sort of control of the direction of growth of an open source project. Having no say in a project’s growth can prove to be a competitive disadvantage A large organization cannot simply expect the use the power of its resources to influence an open source project. In addition to code contributions, it is important to communicate well with the global community of developers, be transparent in your actions, be supportive of other people’s work, and take other actions that advance the project in some way.

A paper called “**A Hierarchical Taxonomy of Leadership Behavior**” published by Gary Yukl, Angela Gordon, and Tom Taber identifies three meta-categories (Task Behavior, Relations Behavior, and Change Behavior). Within these meta-categories, the paper lists 12 behaviors that are important to be an effective leader.

We have tried to determine if there is a correlation between leadership actions listed in the research paper and the developers ratings, and hence their influence in an open source project. We will check if the scores for the leaders we have identified after analyzing their posts on the forum also correlate with their ratings. This will tell us if the taxonomy described in the research paper can be used to identify potential leaders or leadership behaviors. The dataset we have used is from a Bitcoin forum.

# Table of Contents

[Summary 2](#_Toc27051688)

[Table of Contents 3](#_Toc27051689)

[Introduction 4](#_Toc27051690)

[Motivation for the study 4](#_Toc27051691)

[Proposed Metrics 5](#_Toc27051692)

[Literature Review 6](#_Toc27051693)

[Data set details 8](#_Toc27051694)

[Analysis notes 8](#_Toc27051695)

[Results and Discussion 9](#_Toc27051696)

[Dictionary: Leadership behavior taxonomy 9](#_Toc27051697)

[Chart 1: Distribution of the 12 behaviors for the entire dataset 10](#_Toc27051698)

[Chart 2: Leadership taxonomy used across the three meta-categories 11](#_Toc27051699)

[Table 1: Developers with the highest leadership scores 12](#_Toc27051700)

[Chart 3: Distribution of behaviors for the top 5 developers 13](#_Toc27051701)

[Chart 4: Total words spoken for each of the 12 dimensions based on issue discussion intensity 14](#_Toc27051702)

[**Paired t-test** 15](#_Toc27051703)

[Chart 5: Leadership taxonomy used across the three meta-categories 16](#_Toc27051704)

[**Paired t-test** 17](#_Toc27051705)

[Chart 6: Developers with highest leadership actions – based on intensity of discussion 18](#_Toc27051706)

[**Paired t-test** 19](#_Toc27051707)

[Chart 7: Correlation of Leadership scores with developer ratings 20](#_Toc27051708)

[Conclusions 21](#_Toc27051709)

[Limitations 21](#_Toc27051710)

[Reflection 21](#_Toc27051711)

[References 22](#_Toc27051712)

[Appendices 23](#_Toc27051713)

[Steps to run the Python parser 23](#_Toc27051714)

# Introduction

## Motivation for the study

The rate of adoption of open source software has been growing exponentially. In a study conducted by Black Duck Software, in 2018, except for the computer hardware and semiconductors industry, in every other industry at least 90% of the audited codebases contained open source components.

Investing in open source software lowers Research & Development costs, development costs are shared with a global development community, product development is accelerated, greater innovation is achieved. Big companies are investing in open source projects. IBM has bought Red Hat for $34 billion, and Microsoft paid $7.5 billion for Github.

Why does it make sense to study leadership behaviors in the open source community?

* An organizations products may depend on open source technologies
* An open source project’s contributors may include competitors. Having no say in a project’s direction can prove to be a competitive disadvantage. Therefore, building and maintaining leadership in open source projects is key to corporate strategies and goals
* Unlike in a commercial project, creating influence in an open source project is not as simple as using your financial resources
* Open source means a global community where you need to be able to reach out to people who are from very different walks of life, doing very different things, in different companies, and be able to have that empathy and understating what people really need and want

The motivation for this study is to identify behaviors or actions that can help an organization or an individual achieve a leader status, and hence greater influence on an open source project/forum.

# Proposed Metrics

Realizing the importance of building leadership in an open source project, we want to identify if there are any emergent leadership behaviors or practices in the open source community?

**We will use the GQM methodology to structure our research:**

**Goal:** To explore leadership qualities in open source community

**Questions:**

Q1. Can the existing leadership taxonomy be applicable to open software community?

Q2. What are some types of leadership behaviors in open source community?

**Metrics**

M1: Number of posts per developer

M2: Number of posts indicating leadership actions per developer

M3: Number of core leaders identified

*For our research, we will define a core leader as someone who displays a high score for behaviors that have been identified as important to be an effective leader. A leader is also rated highly by the other developers, and may be able to influence the direction of the software product.*

M2: Distribution of the 12 items of the core leaders

M5: Correlation between a developers leadership rating and leadership behavior scores?

# Literature Review

We have used a research paper called “**A Hierarchical Taxonomy of Leadership Behavior**” published by Gary Yukl, Angela Gordon, and Tom Taber as a guide. The paper classifies leadership behaviors into three meta-categories: Task Behavior, Relations Behavior, and Change behavior. Within each of these meta-categories, the authors identify a total of 12 dimensions to more accurately identify leadership actions.

A large number of studies have tried to correlate leadership behavior and leadership effectiveness, that is, types of behaviors that enhance individual and collective performance.

A major problem in comparing the results of all the studies is the lack of agreement about which behavior categories are the most relevant. Additionally, a number of behavior categories and a number of taxonomies on leadership behaviors has emerged. This makes is difficult to compare the results from one set of studies to another.

The primary objectives of task behavior include high efficiency in the use of resources and personnel, and high reliability of operations, products, and services. The primary objectives of relations behavior include strong commitment to the unit and its mission, and a high level of mutual trust and cooperation among members. The primary objectives of change behavior include major innovative improvements

(in processes, products, or services), and adaptation to external changes.

A summary of the 12 behaviors and their relation to leadership effectiveness (from the studies done so far) is presented below.

*A high rating for the leader does not always mean a high follower performance.*

|  |  |
| --- | --- |
| Task Behavior | |
| Short-term planning | Several studies have found correlation between planning and managerial effectiveness |
| Clarifying responsibilities | Positive relation between clarifying responsibilities and managerial effectiveness, but not for all situations. Challenging goals results in higher performance as long as the goals are accepted |
| Monitoring operations and performance | Two observational studies found that leaders who did more monitoring were more effective.  In the survey studies, monitoring was related to leader effectiveness for some samples but not others |
| Relations Behavior | |
| Supporting - showing  consideration, acceptance, and concern for the  needs and feelings of other people | There is strong evidence that supporting is related to follower satisfaction with the leader.  However, only a weak, inconsistent relationship has been found between supporting and follower performance |
| Developing (coaching) | In the survey research, developing was correlated with follower performance in some studies but not others |
| Recognizing - involves giving praise and  showing appreciation to others for effective  performance, significant achievements, and  important contributions to the organization. | Most survey studies on the consequences of recognizing have found a positive correlation with subordinate satisfaction. However, results for effects on performance are less consistent in the survey studies |
| Consulting - involving followers in making important decisions | there is only a weak, inconsistent relationship with follower satisfaction and performance |
| Empowering - includes delegating and  providing more autonomy and discretion to  subordinates | The results from survey research on the  relationship between delegation and subordinate  performance have been inconsistent and difficult  to interpret, |
| Change Behaviors | |
| External Monitoring – monitor external environment and identify threats and opportunities for the organization | Studies have shown that leaders of high performing companies do more external monitoring than leaders of low performing companies |
| Envisioning change – articulating an inspiring vision of a better future | Several studies provide evidence for envisioning change as an important leadership behavior |
| Encouraging innovative thinking | There is evidence that this type of behavior is an effective leadership tool |
| Taking personal risks | Effects of risk taking on successful implementation of change have not been directly investigated |

# Data set details

The original data set had 18,657 comments on 3423 issues by 1926 commenters

To make the dataset easier to manage and the Python program to run faster, we will only consider issues that have at least 10 comments. This leaves us with **8137 comments from 882 unique commenters on 498 unique issues**. Additionally, we expect to find more meaningful analysis from issues that have been discussed in some detail.

## Analysis notes

In addition to analyzing the entire dataset, we want to see if the intensity of discussion affects the behaviors that are displayed. To do this we have also divided the dataset into issues that were discussed at least 30 times and issues that were discussed less than 30 times.

# Results and Discussion

## Dictionary: Leadership behavior taxonomy

**Step 1:** Identify words that might be used by leaders in each of the 12 behaviors

**Step 2:** Break down the words into their roots so we can find any variations spoken by the developers.

The preliminary set is:

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Clarifying roles** | assign | explain | objective | expect | clarif |  |  |  |
| **Monitoring operations** | progress | quality | evaluat | performance | monitor |  |  |  |
| **Short-term planning** | plan | resource | schedule | coordinate |  |  |  |  |
| **Consulting** | consult | check | encourag | decision making | suggestion | idea |  |  |
| **Supporting** | support | considerat | sympathy | upset | anxious | encourag | stress | difficult |
| **Recognizing** | recogni | praise | highlight | contribution | reward |  |  |  |
| **Developing** | coach | advi | opportuni | develop | improve |  |  |  |
| **Empowering** | responsibilit | trust | approv |  |  |  |  |  |
| **Envisioning change** | change | envision |  |  |  |  |  |  |
| **Taking risks for a change** | risk | sacrifi |  |  |  |  |  |  |
| **Encouraging Innovative Thinking** | innovat | creat |  |  |  |  |  |  |
| **External Monitoring** | trend | event | environment | threat | clarif |  |  |  |

## Chart 1: Distribution of the 12 behaviors for the entire dataset

**Supporting,** which is defined as “showing concern for the needs of others” has the highest score.

*Note: Supporting has shown to be related to high satisfaction with the leader, but only a weak relationship with follower performance*

**Taking personal risks (Change behavior)** to implement change has the lowest score

## Chart 2: Leadership taxonomy used across the three meta-categories

**Change behavior,** which includes major innovative improvements

(in processes, products, or services), and adaptation to external changes, scores higher than task behavior Task, and Relations Behavior

## Table 1: Developers with the highest leadership scores

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Developer** | **Number of comments** | **Total leadership score** | **Clarifying roles** | **Monitoring operations** | **Short-term planning** | **Consulting** | **Supporting** | **Recognizing** | **Developing** | **Empowering** | **Envisioning change** | **Taking risks for a change** | **Encouraging Innovative Thinking** | **External Monitoring** |
| **laanwj** | 1182 | **462** | 47 | 38 | 36 | 64 | 84 | 4 | 46 | 13 | 0 | 10 | 41 | 79 |
| **dooglus** | 152 | **246** | 24 | 63 | 7 | 7 | 0 | 1 | 13 | 6 | 0 | 0 | 51 | 74 |
| **sipa** | 498 | **198** | 27 | 14 | 15 | 26 | 26 | 0 | 25 | 16 | 1 | 7 | 22 | 19 |
| **lovesh** | 16 | **140** | 4 | 0 | 1 | 0 | 62 | 5 | 0 | 0 | 0 | 0 | 36 | 32 |
| **dthorpe** | 5 | **136** | 0 | 133 | 2 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 0 | 1 |
| **jonasschnelli** | 302 | **130** | 6 | 3 | 2 | 13 | 23 | 2 | 14 | 0 | 0 | 2 | 19 | 46 |
| **Diapolo** | 412 | **118** | 6 | 9 | 6 | 22 | 12 | 1 | 8 | 0 | 0 | 0 | 42 | 12 |
| **gmaxwell** | 179 | **104** | 17 | 11 | 2 | 3 | 23 | 4 | 13 | 4 | 0 | 4 | 19 | 4 |
| **achow101** | 89 | **99** | 5 | 8 | 1 | 4 | 27 | 2 | 1 | 1 | 0 | 1 | 17 | 32 |
| **rebroad** | 148 | **90** | 9 | 4 | 1 | 15 | 11 | 0 | 14 | 19 | 0 | 2 | 13 | 2 |

The table shows the top 10 commenters based on total leadership scores

**Envisioning Change**, and **Recognizing** behaviors seem to have low scores for all the top developers

## Chart 3: Distribution of behaviors for the top 5 developers

For the Top 3 developers, the total score is spread across the behaviors

*For lovesh and dthorpe, nearly their entire score is made of up one behavior. We might have to look at the words in the dictionary to understand this.*

**We will now explore if the intensity of discussion affects the scores or the distribution of leadership behaviors**

## Chart 4: Total words spoken for each of the 12 dimensions based on issue discussion intensity

### **Paired t-test**

**We will now check if there is a significant difference between the two groups. We will conduct a paired t-test on the two sets**

H0: μ1 - μ2 = 0  
H1: μ1 - μ2 ≠ 0

The null hypothesis is that there is no difference

|  |  |  |
| --- | --- | --- |
| t-Test: Paired Two Sample for Means | |  |
| 12 Behaviors distribution | Under 30 | Over 30 |
| Mean | 305.25 | 99.91666667 |
| Variance | 41358.02273 | 6512.083333 |
| Observations | 12 | 12 |
| Pearson Correlation | 0.924144012 |  |
| Hypothesized Mean Difference | 0 |  |
| df | 11 |  |
| t Stat | 5.371134785 |  |
| P(T<=t) one-tail | 0.000113196 |  |
| t Critical one-tail | 1.795884819 |  |
| P(T<=t) two-tail | 0.000226392 |  |
| t Critical two-tail | 2.20098516 |  |

The null hypothesis states that there is no difference in the means of the two sets.

The p value for the paired t-tests is significantly smaller than the Alpha value 0.05 meaning that we can reject the null hypothesis.

We can say that the two groups are significantly different. This means that at the individual behavior level a different set of behaviors is displayed for issues discussed at different intensities

## Chart 5: Leadership taxonomy used across the three meta-categories

### **Paired t-test**

**We will now check if there is a significant difference between the distribution of scores for the three meta-categories. We will conduct a paired t-test on the two sets**

H0: μ1 - μ2 = 0  
H1: μ1 - μ2 ≠ 0

The null hypothesis is that there is no difference

|  |  |  |
| --- | --- | --- |
| t-Test: Paired Two Sample for Means | |  |
| **Three meta-categories** | **Under 30** | **Over 30** |
| Mean | 2070 | 399.6667 |
| Variance | 1035811 | 2508.333 |
| Observations | 3 | 3 |
| Pearson Correlation | 0.632404 |  |
| Hypothesized Mean Difference | 0 |  |
| df | 2 |  |
| t Stat | 2.931689 |  |
| P(T<=t) one-tail | 0.049659 |  |
| t Critical one-tail | 2.919986 |  |
| P(T<=t) two-tail | 0.099318 |  |
| t Critical two-tail | 4.302653 |  |

The p value for the paired t-tests is greater than the Alpha value 0.05

We cannot reject the null hypothesis

At the meta-category level there doesn’t seem to be a difference between behaviors that are displayed for issues that are discussed at high and low intensity

## Chart 6: Developers with highest leadership actions – based on intensity of discussion

### **Paired t-test**

To conduct a paired t-test for the behaviors discussed by the top developers for issues that are discussed at different intensities, we will first add the scores of the each behavior for the top 5 developers

|  |  |  |
| --- | --- | --- |
|  | **Under 30** | **Over 30** |
| **Clarifying roles** | 75 | 30 |
| **Monitoring operations** | 191 | 54 |
| **Short-term planning** | 55 | 12 |
| **Consulting** | 107 | 11 |
| **Supporting** | 120 | 109 |
| **Recognizing** | 6 | 10 |
| **Developing** | 82 | 17 |
| **Empowering** | 18 | 13 |
| **Envisioning change** | 0 | 1 |
| **Taking risks for a change** | 17 | 2 |
| **Encouraging Innovative Thinking** | 96 | 90 |
| **External Monitoring** | 138 | 168 |

|  |  |  |
| --- | --- | --- |
| t-Test: Paired Two Sample for Means | |  |
|  | *Under 30* | *Over 30* |
| Mean | 75.41666667 | 43.08333 |
| Variance | 3487.356061 | 2779.538 |
| Observations | 12 | 12 |
| Pearson Correlation | 0.643453446 |  |
| Hypothesized Mean Difference | 0 |  |
| df | 11 |  |
| t Stat | 2.355936558 |  |
| P(T<=t) one-tail | 0.019043062 |  |
| t Critical one-tail | 1.795884819 |  |
| **P(T<=t) two-tail** | **0.038086124** |  |
| t Critical two-tail | 2.20098516 |  |

The null hypothesis states that there is no difference in the means of the two sets.

The p value for the paired t-tests is significantly smaller than the Alpha value 0.05 meaning that we can reject the null hypothesis.

We can say that the two groups are significantly different. We can say that there is significant difference in the behaviors displayed by the top developers for issues that are discussed at different intensity levels

## Chart 7: Correlation of Leadership scores with developer ratings

|  |  |  |
| --- | --- | --- |
| **Developer** | **Merit Score/Ratings** | **Leadership score** |
| **laanwj** | 667 | 462 |
| **sipa** | 228 | 198 |
| **jonasschnelli** | 330 | 130 |
| **Diapolo** | 500 | 118 |
| **gmaxwell** | 260 | 104 |
| **MarcoFalke** | 259 | 86 |
| **gavinandresen** | 200 | 76 |
| **luke-jr** | 158 | 76 |
| **TheBlueMatt** | 200 | 66 |
| **fanquake** | 444 | 61 |
|  | **Correlation coefficient** | **0.71** |

**A high positive value of correlation coefficient (0.71) between the top developers leadership scores and their ratings on the forums suggests that we might be able to assess leadership qualities of individuals in the open source community**

# Conclusions

We observed a strong positive correlation between leadership scores (calculated based on the use of leadership taxonomy) and developer ratings. This suggests that leadership taxonomy could be used to identify potential leaders in the open source community. Note: *High leadership rating does not always mean high follower performance*

We observe that there is significant difference in the behaviors displayed for issues that are discussed at more or less intensities

At the meta-category level, however, the difference wasn’t as significant

The above two results lead us to think that leadership actions should be looked at individual behavior level and not at the meta-category level. At the meta-category level the two sets (less and high intensity) did not show much difference but showed significant difference at the behavior level. A different set of behaviors will be more effective depending on the situation, so more emphasis should be given to the results of specific behaviors.

# Limitations

The results can vary depending on the words selected for the dictionary

* There are some words that could be used in a completely unrelated context than to suggest leadership qualities. This can give skewed results for some developers. For example, “change” was being used to assign a task and not to reflect “envisioning change”. We removed change from the dictionary but there could be other words as well

We have only considered the 12 behaviors listed in the research paper. Even the research paper admits that to limit of the questionnaire they used in their research, they did not include all the leadership behaviors identified in prior research. We may want to do additional research by other possible leadership actions.

# Reflection

Our research indicates that existing leadership taxonomy can be applied to identify leadership behaviors in the open source community. We noted a strong positive correlation between a developers ratings and their interactions on the open source forums.

We should perform additional research to identify which types of behaviors are more effective in what types of projects or issues.

We can explore applying the same technique for posts on other social media platforms to identify potential leaders or leadership qualities.

# References

A Hierarchical Taxonomy of Leadership Behavior: Integrating a Half Century of Behavior Research – Gary Yukl, Angela Gordon, Tom Taber

Leadership and Inclusion in the Open Source Community – Kiran Oliver (<https://thenewstack.io/leadership-and-inclusion-in-the-open-source-community/>)

Building Leadership in an Open Source Community – The Linux Foundation (<https://www.linuxfoundation.org/resources/open-source-guides/building-leadership-in-an-open-source-community/>)

The Future of Open Source Software – Forbes.com – Taylor Armerding (<https://www.linuxfoundation.org/resources/open-source-guides/building-leadership-in-an-open-source-community/>)

# Appendices

  
*Python parser to calculate leadership scores*

*  
Dictionary of words for each behavior. The dictionary can be modified to add more behaviors and words.*

*  
Sample output file. Includes results of running the parser on the Bitcoin dataset.*

## Steps to run the Python parser

1. Create a dictionary

- Create a spreadsheet called “dictionary.xls”

- Create a worksheet called “dictionary”

- Add the 12 Behaviors starting from cell B2 to B13

- Add words you want to look for against each of the behavior. Write the words in lowercase.

2. Prepare data set for parser

- Create a new spreadsheet. Name it anything

- Create a worksheet “data”

- Put issue IDs in column A, starting from cell A2; Names of commenters in column B, starting from cell B2; Comments in column C, starting from column C2

- Make the comments column all lowercase

3. Run the parser

- Open the parser in any editor

- Add the paths for the spreadsheet created in step 2, and the “dictionary” spreadsheet created in step 1 in variables “loc” and “loc\_dictionary”

- Run the Python program

4. Analyze the results

- The program will calculate the number of comments, the score for each behavior, and the total score for each of the commenters in the “Analysis” sheet

- The program will calculate the number of comments for each issue in the “Issues Analysis” sheet.

- The file might take a couple of minutes to open. Once the file is open value paste the “Analysis” and the “Issues Analysis” sheets so the file opens faster next time.