Artificial Intelligence (AI) Capabilities, Trust and Open Source Software Team Performance

Babu Veeresh Thummadi

[1] Lero, The Science Foundation Research Centre for Software,
Cork University Business School, University College Cork, Ireland.
vthummadi@ucc.ie

Abstract. In recent years, Artificial Intelligence (AI) has become a key element in digital platforms for improving performance. Despite vast body of knowledge it is yet unclear on how AI can be successfully integrated into platforms and what are the key mechanisms that drive the performance in digital platforms such as open source. To investigate this phenomena a survey has been conducted to understand how AI capabilities (i.e., capabilities associated with AI resources/usage) on Open Source Software (OSS) team performance. The analysis highlights the role of trust in driving OSS team performance and suggests that designers need to pay more attention to cognition when dealing with AI technologies and opportunities.

Keywords: open source, artificial intelligence (AI), AI capabilities, OSS team performance, trust.

1 Introduction

Artificial Intelligence (AI) can be defined as "a broad collection of computer-assisted systems for task performance, including but not limited to machine learning, automated reasoning, knowledge repositories, image recognition, and natural language processing" [1]. In recent years open source software development, a type of software development practice that uses voluntary workers for creating software with minimal restrictions on code usage, has come to the forefront in solving some of the grand challenges associated with development of AI technologies [2-4].

Despite the growing importance of AI in open source production, very little work has been done on important issues surrounding how AI can be used as a capability in enhancing Open Source Software (OSS) team performance [2, 5]. AI capability can be thought of as a unique feature of open source team that measures open source teams' inclination in seeking AI opportunities and resources. For example, AI can be used an infrastructure in the form of bots in OSS teams for streamlining open source process such as closing pull requests, troubleshooting, greeting new users etc. At the same time, OSS teams can also explore new business opportunities in AI to increase attractiveness of the project. As open source communities use AI in myriad of ways, it is unclear how AI capabilities can affect OSS team performance[4]. Hence I ask:

RQ1: How does AI capabilities affect open source software team performance? And what are the key mechanisms that drive open source team performance when using AI capabilities?

To investigate this question a theoretical model has been developed to understand the effects of AI capabilities on OSS team performance using existing literature. Then a survey has been carried out on Mechanical Turk (or simply "MTurk") to study the effects of AI capabilities on open source team performance [6]. A total of 223 responses have been recorded and the data has been cleaned extensively using various data cleaning strategies suggested in the literature. By employing various data clearing strategies, the final sample was reduced to 89 responses and the analysis was then carried out in smart PLS for understanding the relationships between the theorized variables. The analysis revealed that the effect of AI (specifically AI proactive stance i.e., projects' ability to acquire and exploit AI knowledge and innovations) on open source team performance was significant and was fully mediated by cognitive trust i.e., the trust that is generated by rational assessments. The analysis highlights the role of trust in driving OSS team performance and suggests that designers need to pay more attention to cognition when dealing with AI technologies and opportunities.

The paper is organized as follows. In the next section a background of AI and open source is discussed. Then key hypothesis are presented. This is followed by discussion of data collection and analysis. Then findings are reported. Finally the paper ends with discussion, conclusion and limitations.

2 Background

The field of AI was established around 1950s, though its adoption has been rather slow. AI as a field received recognition with the IBM "Deep Blue" intelligent computer program outpacing world-famous chess player Gary Kasparov in 1997. From then on AI gained steam and began to applied in organizations and societies for replacing or augmenting human intelligence [1]. For example, AI is currently used in wide variety of application for tasks such as autocompletion, crime detection, hiring, medical diagnosis, self-driving, recommendations etc. One key thing that distinguishes AI from simple automation programs such as auto reply is that AI has the capability in being unpredictable much like humans [7]. For example, when faced with a roadblock AI-based systems are expected to act more intelligently and take actions in harmony with the environment and not throw an automated response which can be harmful and dangerous. Hence von Krogh (2018) eloquently wrote, "AI has the qualities of being a new but poorly understood organizational phenomenon. By concentrating efforts on collecting quantitative and qualitative data on the aforementioned questions, we may discover unanticipated relationships and ways to resolve tensions and ambiguities in the research on AI within the domains organizational decision-making and problem-solving"[1].

When organizations or platforms use AI they use them for multiple purposes either in problem-solving or decision-making activities [1]. Within the first few decades of the introduction of AI, problem solving has been the prime focus and has been extensively used to solve problems and more recently the shift has been towards decision-making and how we can make AI more autonomous (i.e., the organizations and platforms are moving towards strong AI from weak AI that can intelligently decide and simultaneously solve problems without human intervention or minimal human intervention). Within the realm of digital platforms, open source software development platforms are using AI as a capability for creating cutting edge AI software that address global problems and societal challenges and for streamlining open source processes such as closing pull request, or greeting a new member etc [4, 5]. For example, open source projects use AI as infrastructure and use bots for 1) license creation 2) reviewing and 3) chatting. Despite wide scale usage, bots are still are considered problematic, as they pose challenges in social interactions in the open source platforms. Many open source workers agree that bots are poor in social interactions [5].

In conclusion, use of AI in open source has been challenging and more work is needed in terms of understanding the relationship between AI capabilities and OSS team performance. In the next section key hypothesis are discussed.

3 Hypothesis Development

In this section the key hypothesis are developed (see Figure 1). In recent IS literature, AI has been viewed as a capability as it enables platforms and organizations to sense, comprehend, act, and learn [8]. This concept of AI capabilities holds parallels to existing IT capabilities construct that is routinely used in IS literature [8] and hence AI capabilities can be defined as a "platform/firm's ability to acquire, deploy, combine, and reconfigure AI resources in support and enhancement of business strategies and work processes"[9]. Following prior research on IT capabilities, AI capability can be conceptualized as a latent construct reflected in three dimensions: *AI infrastructure capability* (the technological foundation), *AI business spanning capability* (business-AI strategic thinking and partnership), and *AI proactive stance* (opportunity orientation) [9]. All these three dimensions can play a key role in improving the open source development process and team performance.

For example, having a strong AI infrastructure capability and using bots can increase the confidence of the workers in making the processes more efficient by reducing unnecessary bells and whistles and hence AI infrastructure capability can have a positive effect on open source team performance by inducing novelty and streamlining existing process [5]. In a similar vein, AI business opportunities and orientation can also make open source workers tuned to the projects as open source workers like nut-cracking problems [10, 11]. Hence the orientation and business opportunities of AI can have a positive impact on open source team performance. Hence, I hypothesize:

H1/H2/H3: *AI [Infrastructure/ Business Spanning/ Proactive stance] can have a positive effect on OSS Team Performance.*

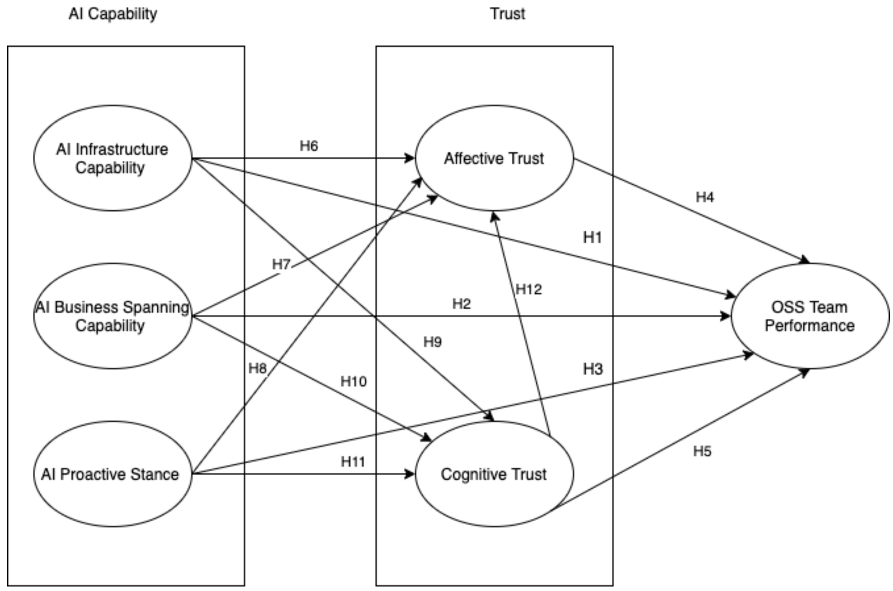

Figure 1: Theorized model on the effects of AI on OSS Team performance

Prior works in open source development suggest that trust is an important factor that leads to effectiveness of the open source project. Stewart and Gosain (2006) distinguish two different forms of trust: cognitive and affective trust [12]. Cognitive trust arise out assessing the members and bots rationally and affective trust reflects the emotional connections between the actors (both humans and bots). By trusting and giving more power and decision making to AI technologies open source workers can focus on other more important project decisions and activities such as coming up with new requirements that can lead to better outcomes [5]. Both these types of trust can be critical in AI context as the workers' positive trust can have a positive effect on open source team performance. Hence I hypothesize:

H4/H5: *[Cognitive/ Affective] Trust can have a positive effect on OSS Team Performance.*

Trust has been the key theme in AI research in the recent few years as humans find it hard to trust and give power to bots and algorithms taking over the life [7]. Much similar to how human-human trust develops through frequent interactions, the relationship between actors and bots can follow a similar trajectory. When open source workers rationally/emotionally believe and develop cognitive/affective trust on the bots and algorithms, the effects of AI capabilities [business spanning/ infrastructure/ proactive stance] on open source team performance can become mediated through trust[12]. Hence, I hypothesize:

H6/H7/H8/H9/H10/H11: *[Cognitive/Affective] Trust mediates the effect of AI [Infrastructure/ Business Spanning/Proactive stance] on OSS Team Performance.*

Prior studies on open source development also hypothesize that cognitive trust can have a positive effect on affective trust, hence I include this in our model as well[12]. Hence, I hypothesize:

H12: *Cognitive Trust has a positive effect on affective trust.*

In the next section, research design, data collection, survey design and data analysis are discussed.

4 Research Methods

4.1. Data collection

I first developed a survey instrument based on the existing scales (see Appendix for the scale and items) [9, 12, 13]. These items were asked to rate on a scale of 1-7. The survey was conducted on Amazon Mechanical Turk much similar to prior studies that were published in social sciences [6] and survey received 223 responses. Many researchers might question the legitimacy of the data and hence I chose to restrict the responders based on the ratings of the responders (for example, only responders with 95% accuracy scores were asked for responses, this is a feature in MTurk which was utilized for increasing accuracy of data), and also the survey explicitly asks in the first page that the survey is limited to open source workers with at least 1 years of open source and AI experience.

4.2. Data analysis

Data was checked for reliability and then descriptive statistics were used to get a look at the data. The very first step was to clean the raw data. The responders who had answered the reverse coded question incorrectly were eliminated. Responses to control questions not adhering to survey guidelines were eliminated too. The survey on finalization was given to few candidates for pilot test. It was noticed that on an average they needed 180 seconds to complete the survey. Hence the cut off was set to eliminate responders who would respond to the survey within 180 seconds. Finally, missing values were dealt by elimination as well. The biggest loss of responses were the wrong answers to the reverse coded control question. Using the above strategies I was able to improve the quality of the data which is considered an issue in the studies conducted on MTurk. After cleaning the data there were 89 observations and this data was analysed using smart PLS using an exploratory factor analysis[14].

5 Findings

To analyse the survey data, I used the formative model and performed the exploratory factor analysis to understand how the items in survey getting loaded. I performed the single harman test to rule out any common method bias (and the variance less than 50%). The EFA revealed that the items were getting loaded very well and the loadings were above .4 as suggested by literature [12, 15]. One item was dropped in affective trust and cognitive trust constructs. The results from the EFA is listed in Table 1. For testing the reliability and validity, I used the threshold values set by Fornell and Larker (1981) and ensured that the item loadings were above 0.7, construct's composite reliability (CR) scores were above 0.8 and average variance extracted (AVE) above 0.5. From Table 1 we can see that the item loadings were all above .7[16].

Table 1: Factor loadings

	1	2	3	4	5	6
AIB11	**0.798**	0.51	0.57	0.625	0.596	0.539
AIB21	**0.809**	0.527	0.638	0.632	0.648	0.692
AIB31	**0.775**	0.447	0.644	0.614	0.565	0.574
AIB41	**0.764**	0.525	0.617	0.655	0.568	0.511
AIIC11	0.545	**0.877**	0.478	0.466	0.462	0.43
AIIC21	0.598	**0.917**	0.582	0.557	0.513	0.565
AIPS11	0.672	0.475	**0.791**	0.67	0.637	0.644
AIPS21	0.628	0.559	**0.791**	0.559	0.575	0.643
AIPS31	0.5	0.4	**0.733**	0.484	0.529	0.531
AIPS41	0.638	0.425	**0.809**	0.643	0.633	0.599
AT11	0.607	0.387	0.566	**0.822**	0.616	0.628
AT21	0.589	0.496	0.595	**0.774**	0.564	0.536
AT31	0.714	0.536	0.685	**0.827**	0.722	0.638
AT51	0.676	0.432	0.604	**0.815**	0.672	0.611
CT11	0.671	0.428	0.629	0.724	**0.806**	0.682
CT21	0.61	0.384	0.643	0.615	**0.785**	0.633
CT31	0.506	0.405	0.527	0.613	**0.766**	0.574

CT51	0.575	0.433	0.616	0.625	**0.784**	0.672
CT61	0.575	0.48	0.543	0.523	**0.754**	0.53
TP11	0.605	0.436	0.615	0.607	0.679	**0.792**
TP21	0.558	0.434	0.662	0.556	0.588	**0.791**
TP31	0.656	0.48	0.622	0.614	0.682	**0.803**
TP41	0.57	0.374	0.511	0.601	0.566	**0.714**
TP51	0.447	0.435	0.578	0.495	0.546	**0.756**

I also checked for the convergent validity and the AVE and the cronbach's alpha were in the expected thresholds as prescribed by Fornell and Larker (1981)[16]. See Table 2 for the results pertaining to the convergent validity.

Table 2: Construct validity

	Cronbach's Alpha	Composite Reliability	Average Variance Extracted (AVE)
AI Business Spanning Capability	0.795	0.866	0.619
AI Infrastructure Capability	0.759	0.892	0.805
AI Proactive stance	0.788	0.863	0.611
Affective trust	0.825	0.884	0.655
Cognitive trust	0.838	0.885	0.607
OSS Team performance	0.830	0.880	0.596

For testing discriminant validity I used the criteria that is prescribed by Chin (1998) and Fornell and Locker (1981)[16, 17]. See Table 3 and 4 for the results pertaining to the divergent validity. First, I investigated the cross-loadings for the six factor identified by the model. Through visual inspection by row in Table 1 we can see that all constructs' indicators load highest into the respective constructs than other constructs. Second, I used the criteria prescribed by Fronell-Locker criterion and evaluated to see if the square root of AVE value are higher than the correlations with other constructs (square root of AVE is shown in bold in Table 3). From Table 3, we can see most of the constructs pass this criterion expect for AI Business spanning capability, Cognitive trust and OSS team performance. The square root of AVE for AI Business spanning capability is .789 as shown in the diagonal and the correlations to affective trust is higher by about .02. Based on this we can consider the discriminant validity results to be satisfactory.

Table 3: Discriminant validity

	1	2	3	4	5	6
1.AI Business Spanning Capability	**0.787**					
2.AI Infrastructure Capability	0.639	**0.897**				
3.AI Proactive stance	0.785	0.595	**0.782**			
4.Affective trust	0.802	0.574	0.759	**0.81**		
5.Cognitive trust	0.757	0.545	0.762	0.799	**0.779**	
6.OSS Team performance	0.74	0.561	0.776	0.747	0.797	**0.772**

Then I carried out path analysis. The analysis shows that the effects of AI proactive stance were significant on OSS team performance, though I did not find any significant effects of AI Infrastructure capability and AI business spanning capability on OSS Team performance (see Figure 2). Furthermore, the cognitive trust was found to be mediating the effects of AI proactive stance. This provides support for hypothesis H3, H5, H10, H12.

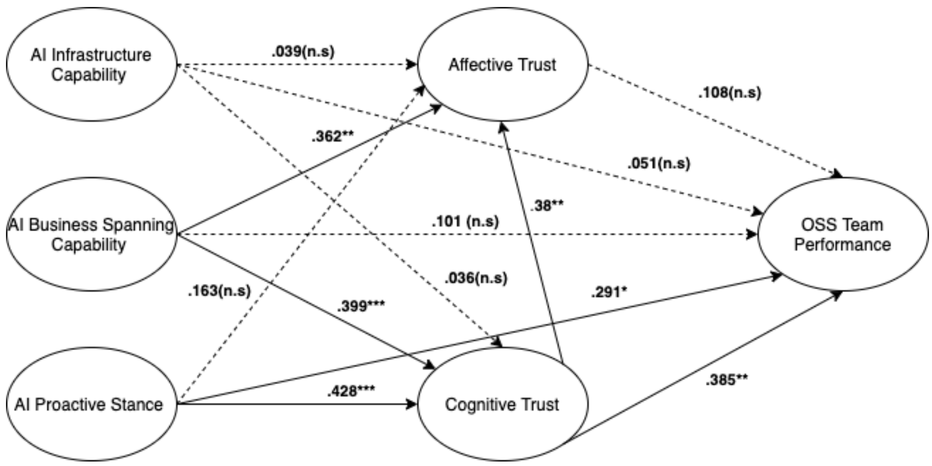

Figure 2: Path model of the effects of AI Capabilities on OSS Team performance (* p <.05, **p <.01, ***p <.001, non-significant paths are indicated with dotted lines)

The results from the study should be useful for open source teams for organizing the projects. By successfully incorporating AI capabilities, open source teams can become more productive by increasing the cognitive trust between the members. In Table 4, the model fit statistics are reported and the SRMR is less than <.08 and hence is considered a good model fit [18].

Table 4: Model Fit

	Saturated Model	Estimated Model
SRMR	0.074	0.074

Chi-Square	481.118	481.118
NFI	0.676	0.676

6 Discussion

In this section I address the original research question: *How does AI innovations get adopted in open source platforms? And what are the key mechanisms that drive open source team performance when using AI innovations?*

The analysis shows that cognitive trust is important in the AI and Open source team performance relationship. Specifically, the results show that having a AI proactive stance can have a positive effect on OSS team performance. However, what was surprising was that AI infrastructure and AI business spanning capabilities did not have an effect on OSS team performance. This suggests that AI is still considered problematic and usage of bots and algorithms in OSS communities does not necessarily increase OSS team performance but rather it is the inclination and proactiveness of the OSS teams in seeking AI opportunities that drives performance [5]. Further, cognitive trust mediated the relationship between AI proactive stance and OSS team performance suggesting that OSS teams should seek and enhance cognitive trust more than affective trust for better OSS team performance. Open source workers are more likely to be technology savvy and hence the relationships might differ from a normal user, for example, in case of a different setting, affective trust might be more important. Hence caution has to be observed in terms of generalizing these results into other environments such as robots in healthcare, self-driving cars etc. Also this study should be considered exploratory as with MTurk we do not know the authenticity of data and hence care should be taken in terms of generalizations, a more detailed study on specific open source community could be carried out to validate or invalidate the current study and suggest possible guidelines for reconciling the MTurk design with an email survey.

I also report here some limitations and opportunities future research. When conducting the research I was limited by the method of collecting data. The initial idea was to conduct survey of teams, which would have narrowed down the technology stack as well as would have given a better idea of OSS as a team. However, I had difficulty securing such teams and hence had to conduct a survey for individual developers. In the future research a small programming question can be set up in the survey to determine whether they have the requisite open source experience which would increase the authenticity of the data collected. MTurk users could also be asked for providing their GitHub profile that could display their experience and reduce the fraud [6]. In conclusion, this study explores the nascent dimension of AI and unpacks the relationships of trust and performance in open source development.

Acknowledgements

This work was supported with the financial support of the Science Foundation Ireland grant 13/RC/2094 and co-funded under the European Regional Development

Fund through the Southern & Eastern Regional Operational Programme to Lero - the Science Foundation Ireland Research Centre for Software (www.lero.ie). Author would like to thank Adwait Bhalero for research assistance and the participants of AIS DIGIT community for the helpful comments on the earlier version of the paper. The errors are my own.

References

1. Von Krogh, G., *Artificial intelligence in organizations: New opportunities for phenomenon-based theorizing.* Academy of Management Discoveries, 2018.
2. Hukal, P., et al., *Bots coordinating work in open source software projects.* Computer, 2019. **52**(9): p. 52-60.
3. Von Hippel, E. and G. Von Krogh, *Open source software and the" private-collective" innovation model: Issues for organization science.* Organization Science, 2003: p. 209-223.
4. Ng, A., *What artificial intelligence can and can't do right now.* Harvard Business Review, 2016. **9**(11).
5. Wessel, M., et al., *The power of bots: Characterizing and understanding bots in oss projects.* Proceedings of the ACM on Human-Computer Interaction, 2018. **2**(CSCW): p. 1-19.
6. Hauser, D.J. and N. Schwarz, *Attentive Turkers: MTurk participants perform better on online attention checks than do subject pool participants.* Behavior research methods, 2016. **48**(1): p. 400-407.
7. Glikson, E. and A.W. Woolley, *Human trust in artificial intelligence: Review of empirical research.* Academy of Management Annals, 2020. **14**(2): p. 627-660.
8. Bawack, R., S.F. Wamba, and K. Carillo, *Where Information Systems Research Meets Artificial Intelligence Practice: Towards the Development of an AI Capability Framework.* Technology, 2019. **12**: p. 15-2019.
9. Lu, Y. and K. Ramamurthy, *Understanding the link between information technology capability and organizational agility: An empirical examination.* MIS quarterly, 2011: p. 931-954.
10. Raymond, E., *The cathedral and the bazaar: musings on Linux and open source by an accidental revolutionary*. 2001: O'Reilly & Associates, Inc. Sebastopol, CA, USA.
11. Von Krogh, G., S. Spaeth, and K. Lakhani, *Community, joining, and specialization in open source software innovation: a case study.* Research policy, 2003. **32**(7): p. 1217-1241.
12. Stewart, K.J. and S. Gosain, *The impact of ideology on effectiveness in open source software development teams.* Mis Quarterly, 2006: p. 291-314.
13. Kostopoulos, K.C., Y.E. Spanos, and G.P. Prastacos, *Structure and function of team learning emergence: A multilevel empirical validation.* Journal of Management, 2013. **39**(6): p. 1430-1461.

14. Hair, J.F., C.M. Ringle, and M. Sarstedt, *PLS-SEM: Indeed a silver bullet.* Journal of Marketing theory and Practice, 2011. **19**(2): p. 139-152.
15. Stevens, J.P., *Applied multivariate statistics for the social sciences*. 2012: Routledge.
16. Fornell, C. and D.F. Larcker, *Evaluating structural equation models with unobservable variables and measurement error.* Journal of marketing research, 1981. **18**(1): p. 39-50.
17. Chin, W.W., *Commentary: Issues and opinion on structural equation modeling*. 1998, JSTOR.
18. Hu, L.t. and P.M. Bentler, *Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives.* Structural equation modeling: a multidisciplinary journal, 1999. **6**(1): p. 1-55.

Appendix

AI Infrastructure capability

Relative to other open source projects, please evaluate your open source projects' AI (artificial intelligence) infrastructure capabilities (for example bots, recommendations etc.,) in the following areas 1-7 scale (1 = strongly disagree, 7= strongly agree).

AIIC1: Bot services (for example chatbots, trouble shooting, automation services etc.,) are helpful in navigating my open source work

AIIC2: I like the recommendations and automatic notifications features

(Source: Lu, Y., & K.(Ram) Ramamurthy. (2011). Understanding the link between information technology capability and organizational agility: An empirical examination. *MIS quarterly*, 931-954.)

AI business spanning capability

Relative to other open source projects, please evaluate your open source projects' AI management capability in responding to the following on a 1 to 7 scale (1 = poorer than most, 7 = superior to most).

AIB1: Developing a clear vision on how AI contributes to business value

AIB2: Integrating open source project planning and AI planning effectively

AIB3: Enabling functional areas and general management's ability to understand the value in AI investments

AIB4: Establishing an effective and flexible AI planning process and developing a robust AI plan

(Source: Lu, Y., & K.(Ram) Ramamurthy. (2011). Understanding the link between information technology capability and organizational agility: An empirical examination. *MIS quarterly*, 931-954.)

AI Proactive stance

Relative to other open source projects, please evaluate your open source projects' capability in acquiring, assimilating, transforming, and exploiting AI knowledge in the following areas on a 1 to 7 scale (1 = strongly disagree, 7= strongly agree).

AIPS1:We constantly keep current with new AI innovations
AIPS2:We are capable of and continue to experiment with new AI as necessary
AIPS3:We have a climate that is supportive of trying out new ways of using AI
AIPS4:We constantly seek new ways to enhance the effectiveness of AI use

(Source: Lu, Y., & K.(Ram) Ramamurthy. (2011). Understanding the link between information technology capability and organizational agility: An empirical examination. *MIS quarterly*, 931-954.)

Affective Trust

Each of the statements below refers to how the participants in your open source project(s) feel about each other. Please indicate the extent to which you agree or disagree with each statement about the group using the following scale (1 = strongly disagree, 7= strongly agree).

AT1: Members of the team have made considerable emotional investments in our working relationships.

AT2: Members of the team have a sharing relationship with each other. We can freely share our ideas, feelings, and hopes.

AT3: On this team we can talk freely with each other about difficulties we are having and know that others will want to listen.

AT4: Members of the team would feel a sense of loss we could no longer work together.

AT5: If a member for this group shared problems with other members, they would respond constructively and caringly.

Source: Stewart, K. J., & Gosain, S. (2006). The impact of ideology on effectiveness in open source software development teams. *Mis Quarterly*, 291-314.

Cognitive Trust

Each of the statements below refers to how the participants in your open source project(s) feel about each other. Please indicate the extent to which you agree or disagree with each statement about the group using the following scale (1 = strongly disagree, 7= strongly agree).

CT1. Members of the team know that everyone on the team approaches their work with professionalism and dedication.

CT2. Given the track records of the team members, we see no reason to doubt each other's competence and preparation for a job.

CT3. Members of the team believe they will be able to rely on other members of the team not to make a job more difficult by careless work.

CT4. Members of the team are concerned with monitoring each other's work*.

CT5. Members of the team believe that other members should be trusted and respected as coworkers.

CT6. Members of the team consider each other to be trustworthy.

Source: Stewart, K. J., & Gosain, S. (2006). The impact of ideology on effectiveness in open source software development teams. *Mis Quarterly*, 291-314.

OSS Team performance

Each of the statements below refers to how well your open source project(s) are positioned in the following activities. Please indicate the extent to which you agree or disagree with each statement about the group using the following scale (1 = strongly disagree, 7= strongly agree).

TP1: Our open source team effectively used its resources.
TP2: Our open source team was within the proposed budget.
TP3: Our open source team was within the proposed time-schedule.
TP4: Our open source team was able to meet its goals.
TP5: Our open source team was able to respond quickly to problems.

Source: Kostopoulos et al. (2012): Structure and Function of Team Learning Emergence: A Multilevel Empirical Validation. Journal of Management, Vol. 39, No. 6, pp. 1430–1461