

GitHub 利用履歴とアンケート調査に基づく チーム開発活動評価実験

伊藤恵^{*1} 松原克弥^{*1} 富永敦子^{*1}

^{*1} 公立はこだて未来大学

An Evaluation Experiment of Team Development Activities based on GitHub Usage History and Questionnaire Survey

Kei Ito^{*1} Katsuya Matsubara^{*1} Atsuko Tominaga^{*1}

^{*1} Future University Hakodate

In system development PBLs, many student teams use several standard tools, such as GitHub. By usage history data of such tools, we aim to evaluate objectively and quantitatively students' activities in PBL. However, evaluation using usage history data of such tools is greatly affected by proficiency level of students. So, it is also needed to evaluate of proficiency level of students. In previous research work, we create a checklist for proficiency level of GitHub and try to evaluate proficiency level by the checklist. This article reports the results of the continuous evaluation of proficiency level of GitHub, the subsequent questionnaire, and the analysis results with GitHub usage history of students.

キーワード：習熟度評価，システム開発 PBL，GitHub

1 はじめに

さまざまな教育分野において実践的な教育方法の一つとして Project Based Learning(以下 PBL)が行われており，情報系の学部や学科においても，実践的な IT 教育の一つとしてシステム開発 PBL が広く実施されるようになってきている⁽¹⁾⁽²⁾。このようなシステム開発 PBL における成績評価は発表会や最終成果物に対する評価が中心である。しかし，学習対象には発表や最終成果物だけでなく，そこに至るまでのプロセスが含まれていることも多い。ところが，システム開発 PBL において開発プロセスの評価方法は確立されているとは言えない⁽³⁾。

システム開発 PBL では GitHub¹などの定番ツールがよく使われる。GitHub はシステム開発等で使われるバージョン管理ツールの一つ Git を基にした Web 上のサービスであり，開発プロセスに紐づけられる利用履歴情報が相応に残される。我々はその情報に着目し，GitHub

を使用するシステム開発 PBL を対象に，利用履歴情報を用いた PBL プロセス評価手法の確立を目指している⁽⁴⁾。しかし，Git/GitHub はまあまあ使えるレベル，かなり使いこなしているレベルなど，使用者の習熟度合の幅が広く，システム開発 PBL においてもチーム間やチーム内で習熟度合の差が大きくなることも多く，それが利用履歴を用いたプロセス評価に大きく影響してしまう。そこで我々は，システム開発 PBL のプロセス評価に活用することを前提として GitHub の習熟度チェックリストを作成し，実際にこれを用いた習熟度の評価実験を繰り返してきた⁽⁵⁾。

本稿では継続的に実施している GitHub 習熟度評価実験と，それに併せて実施した評価結果に対するアンケート調査，および，GitHub 利用履歴分析の結果を報告する。

¹<https://github.com>

2 主観評価に基づく習熟度評価

まず我々は先行研究⁽⁵⁾において Git/GitHub の習熟度評価のためのアンケート項目とアンケート結果からの習熟度判定方法を策定した。以下に述べる。

2.1 アンケート項目

アンケート項目は、Git や GitHub に関する書籍やある程度整理された Web 上の情報⁽⁶⁾⁽⁷⁾を参考に項目候補を決め、複数の GitHub 熟練者で協議した上で決定した。表 1 がアンケート項目の一覧である。

表 1 に記載されている順に、それぞれの項目について「何のことか分からない」「コマンドは分かる/使ったことはある」「分かる/使っている」「使いこなしている/十分理解している」の 4 択で回答させる。

表中の「分類」は、後述する判定のためのものであり、次のように決めている。Git の基本的な操作に関する項目で、Git/GitHub の利用開始初期から使うと思われるものを「基本 1」、基本的な操作だが利用開始後しばらくしてから使うと思われるものを「基本 2」とした。また、GitHub Flow などのバージョン管理のワークフローにおいて多く使われるブランチ操作についても同様に、ブランチの利用開始初期から使うと思われるものを「ブランチ 1」、しばらくしてから使うと思われるものを「ブランチ 2」とした。

ただし、これらはあくまでも基本的な操作の範疇であり、ある程度 GitHub を使い慣れた者はほとんどの項目に「使いこなしている/十分理解している」と回答する可能性がある。より熟練度が高い場合の評価項目について今後検討予定である。

2.2 アンケート結果からの判定方法

前節で述べたアンケートの回答結果から以下の手順で習熟度を判定する。

1. 点数化

まずアンケート回答一つ一つの点数化を行う。「何のことか分からない」を 0 点、「コマンドは分かる/使ったことはある」を 1 点、「分かる/使っている」を 2 点、「使いこなしている/十分理解している」を 3 点とする。

2. 分類ごと集計

次に表 1 の分類ごとに集計を行う。具体的には

「基本 1」の 5 項目の平均点、「基本 2」の 4 項目の平均点、「ブランチ 1」の 5 項目の平均点、「ブランチ 2」の 6 項目の平均点を計算する。それぞれ最低点は 0.0 点、最高点は 3.0 点となる。

3. 習熟度判定

4 つの平均点から以下の基準で習熟度を判定する。

- 4 つの平均点がすべて 1.0 未満なら「入門者」
- 上記以外で 1.0 未満が 2 分類以上なら「初級者」
- 上記以外で 2.0 未満が 2 分類以上なら「中級者」
- それ以外なら「上級者」

3 評価実験

著者ら所属大学学部 3 年生の必修通年 PBL 科目「システム情報科学実習」²において、GitHub を利用しつつシステム開発を行うプロジェクトを対象に、前節で述べた主観評価に基づく習熟度評価を開始時・中間・終了時の 3 回行い、また、その判定結果に対する個々の学習者本人へのアンケート調査を行ったほか、そのうちの 1 プロジェクトについて同じ期間中の GitHub 利用履歴との比較分析を行った。

3.1 実験対象

著者ら所属大学において学部 3 年生を対象として開講される必修通年科目「システム情報科学実習」では、毎年度 20 数個程度のプロジェクトが実施される。その中にはシステム開発を行うプロジェクトもそうでないプロジェクトもあるが、本実験ではそのうちの GitHub を利用してシステム開発を行うプロジェクトを対象とする。具体的には上記の条件を満たす 3 つのプロジェクト A,B,C の協力を得て、それらのプロジェクトの所属学生それぞれ 15 名、15 名、11 名の計 41 名を対象とした。プロジェクトへの配属は学生本人の希望に基づいて行われ、対象学生の多くはシステム開発や開発スキルの向上などに関心がある。前期は 5 月初旬ごろにプロジェクトが実質開始し、7 月上旬には中間発表会と中間報告書提出がある。また、後期は 9 月下旬にプロ

²https://www.fun.ac.jp/edu_career/project_learning/

表 1: アンケート項目

分類	項目
基本 1	コミットができる
基本 2	コミット間の diff がみれる
基本 1	リモートのリポジトリにプッシュができる
基本 1	リモートのリポジトリから pull ができる
基本 1	リモートのリポジトリのクローンができる
基本 2	git commit -amend ときいて使い方がわかる
基本 2	.gitignore が何のためにあるファイルか知っている
基本 2	コミットをどの程度の粒度にすべきか分かる
基本 1	コミットメッセージにはどんな内容を書くべきか分かる
ブランチ 1	ブランチの作成ができる
ブランチ 1	ブランチの削除ができる
ブランチ 2	あるブランチとあるブランチ 2 の差分がすぐに出せることを知っている
ブランチ 2	あるブランチ A とあるブランチ B を比較して、A のみに加えられた差分をすぐに出せることを知っている
ブランチ 1	merge ができる
ブランチ 2	merge に失敗した後、merge 前に戻るやりかたを知っている
ブランチ 1	rebase ができる
ブランチ 2	rebase に失敗した後、merge 前に戻るやりかたを知っている
ブランチ 2	merge 中のコンフリクトを解消できる
ブランチ 1	どんな名前がブランチ名として適しているか分かる
ブランチ 2	どのくらいの頻度 (単位) でマージするのがよいか分かる

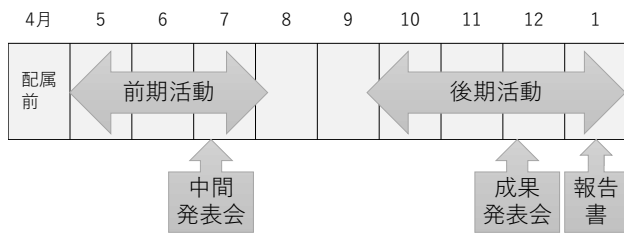


図 1: 実験対象科目の全体スケジュール

プロジェクトが再開され、12月上旬に成果発表会、1月に最終報告書提出がある(図1)。

3.2 習熟度評価

5月のプロジェクト開始直後、7月の前期活動終了後、12月の成果発表会終了後に第2節で述べた方法でアンケート調査とその回答に基づくGit/GitHub習熟度判定を行った。対象学生41名中3回ともアンケート回答が得られた36名分の判定結果が表2である。表では、3回のアンケート実施をそれぞれ開始時、中間、終了時と示し、学生の所属プロジェクト別に終了時の判定の降順に記載している。A01からA14がAプロジェクトの学生、B01からB15がBプロジェクトの学生、C01からC07がCプロジェクトの学生である。

著者ら所属大学ではGitHubそのものを教える授業はないが、プロジェクト開始時点の結果から一部の学生は既にGitHubをある程度使えていることが分かる。前期活動終了時点までに判定結果がやや向上し、成果発表終了後までに大きく向上している。この理由として、前期には各プロジェクト内で行われたGitHub勉強会等により、少し習得が進んだが、前期中はフィールド調査や要求獲得などの活動が多く、GitHubを実際に活用する機会は多くなかったこと、後期には開発が本格化し、プロジェクト内で実際にGitHubの利用が増えたことが考えられる。ただし、どのプロジェクトもごく一部の学生は終了時でもあまり判定結果が高くない。これは、実際に開発が本格化した後期に、チーム内での役割分担が明確化したことで、GitHubをよく利用する学生とそうでない学生とに分かれたことが原因と考えられる。

表の基本1、基本2、ブランチ1、ブランチ2のそれぞれについて、3時点の比較をするために、被験者内

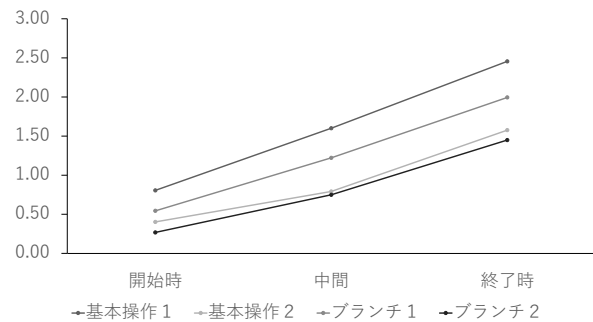


図 2: 各カテゴリの平均値推移

1 要因分散分析を行った結果、基本1、基本2、ブランチ1、ブランチ2ともに、1%水準で有意差が認められた($F(2, 70) = 91.86$; $F(2, 70) = 68.31$; $F(2, 70) = 113.88$; $F(2, 70) = 51.75$)。多重比較(Holm法)の結果、基本1、基本2、ブランチ1、ブランチ2ともに、5%水準で、開始時<中間<終了時という結果であった。この結果から、プロジェクトの活動が進むにつれて、基本1、基本2、ブランチ1、ブランチ2のスキルが有意に上がっていったということが言える(図2)。

3.3 判定結果に対するアンケート調査

7月に実施した2回目の習熟度判定の後、3プロジェクトの所属学生41名全員に、本人および同じプロジェクトの他学生の1,2回目の判定結果を開示した。その上で、自分自身の判定結果と同じプロジェクトの他学生の判定結果について、どう感じたかアンケート調査を行ったところ、22名から回答が得られた。その結果は図3のグラフの通りであった。

自分自身の判定結果に対しては、回答者22名中17名が大変妥当、2名がやや妥当と答えており、3名はやや妥当でないと答えていた。やや妥当でないと回答した3名のうちの1名は自由記述欄に「自分のレベルが上級者になっていたが、上級者と言われるほど実感できていない。」と記述していた。このことから2.1節で述べたようにより習熟度の高い場合の評価項目の必要性が感じられる。また、同じチームの他のメンバーの判定結果に対しては、回答者22名中12名が大変妥当、9名がやや妥当と答えており、1名がやや妥当でないと答えていた。やや妥当でないと答えた1名は自由記述欄に「全体的に初心者、入門者が想像よりも多いなと思った。」と記述していた。なお、どちらの問いに対しても「全く妥当でない」と回答したものはなかった。

表 2: 判定結果

学 生	開始時					中間					終了時				
	基 1	基 2	ブ 1	ブ 2	判定	基 1	基 2	ブ 1	ブ 2	判定	基 1	基 2	ブ 1	ブ 2	判定
A01	1.40	1.25	1.00	1.00	中級	2.80	2.75	3.00	2.83	上級	3.00	3.00	3.00	3.00	上級
A02	3.00	2.50	2.60	1.83	上級	3.00	1.75	2.60	2.00	上級	3.00	2.75	2.80	2.00	上級
A03	0.00	0.00	0.00	0.00	入門	1.20	0.00	0.60	0.17	初級	2.80	2.50	2.20	1.33	上級
A04	0.80	0.25	0.40	0.00	入門	2.00	1.25	1.60	0.67	中級	3.00	1.75	2.20	1.50	中級
A05	1.60	0.50	1.00	0.00	初級	1.60	0.25	0.20	0.00	初級	2.80	1.25	2.20	1.67	中級
A06	1.80	1.25	1.20	0.67	中級	2.00	1.50	1.60	1.33	中級	1.80	1.75	1.60	1.50	中級
A07	1.80	1.00	0.80	0.50	初級	1.80	1.50	1.80	1.33	中級	2.00	1.50	1.80	1.17	中級
A08	1.80	0.50	1.40	0.50	初級	2.00	1.25	1.80	1.17	中級	2.00	1.50	1.60	1.17	中級
A09	0.20	0.00	0.00	0.00	入門	1.20	0.00	0.20	0.00	初級	2.00	1.50	1.40	1.17	中級
A10	0.20	0.00	0.20	0.00	入門	0.80	0.25	0.80	0.33	入門	2.00	0.75	1.80	1.00	中級
A11	0.20	0.00	0.40	0.00	入門	1.80	0.50	1.20	0.50	初級	2.00	1.25	1.40	0.50	中級
A12	0.00	0.00	0.00	0.00	入門	0.60	0.00	0.80	0.33	入門	2.40	0.50	1.60	0.33	初級
A13	0.00	0.00	0.00	0.00	入門	2.00	0.25	0.80	0.33	初級	1.80	0.25	2.00	0.50	初級
A14	0.00	0.00	0.00	0.00	入門	0.20	0.00	0.00	0.00	入門	1.80	0.50	0.60	0.00	初級
	基 1	基 2	ブ 1	ブ 2	判定	基 1	基 2	ブ 1	ブ 2	判定	基 1	基 2	ブ 1	ブ 2	判定
B01	3.00	3.00	2.40	2.67	上級	3.00	3.00	3.00	3.00	上級	3.00	3.00	3.00	3.00	上級
B02	2.00	0.25	1.60	0.33	初級	3.00	2.25	2.60	2.17	上級	3.00	3.00	3.00	3.00	上級
B03	2.00	1.50	1.60	0.50	中級	3.00	2.50	2.60	2.33	上級	3.00	3.00	2.80	2.50	上級
B04	2.00	1.00	1.60	1.00	中級	2.40	1.25	1.60	1.17	中級	3.00	2.50	3.00	2.67	上級
B05	0.00	0.00	0.00	0.00	入門	1.20	0.25	0.60	0.00	初級	2.80	2.00	2.80	3.00	上級
B06	0.80	0.50	0.60	0.00	入門	2.00	1.25	1.80	1.33	中級	3.00	2.25	2.60	2.33	上級
B07	0.00	0.00	0.00	0.00	入門	2.00	0.00	1.20	0.00	初級	3.00	1.75	2.20	2.00	上級
B08	0.60	0.25	0.20	0.00	入門	2.00	0.50	0.80	0.33	初級	3.00	1.75	2.00	1.67	中級
B09	0.00	0.00	0.00	0.00	入門	0.80	0.50	0.20	0.00	入門	3.00	1.75	2.00	1.33	中級
B10	0.00	0.00	0.00	0.00	入門	1.00	0.75	1.00	0.83	初級	2.00	1.75	1.80	1.83	中級
B11	0.00	0.00	0.00	0.00	入門	1.60	0.75	1.20	0.83	初級	2.80	1.50	1.60	1.00	中級
B12	0.00	0.00	0.00	0.00	入門	0.60	0.00	0.80	0.00	入門	2.00	0.75	1.40	1.17	中級
B13	0.00	0.00	0.00	0.00	入門	1.00	0.00	0.40	0.00	初級	2.00	0.75	1.60	0.67	初級
B14	0.00	0.00	0.00	0.00	入門	1.20	0.75	0.80	0.50	初級	1.40	0.50	1.00	0.17	初級
B15	0.00	0.00	0.00	0.00	入門	0.00	0.00	0.20	0.00	入門	1.20	0.00	1.00	0.00	初級
	基 1	基 2	ブ 1	ブ 2	判定	基 1	基 2	ブ 1	ブ 2	判定	基 1	基 2	ブ 1	ブ 2	判定
C01	2.00	0.25	0.80	0.33	初級	1.00	0.50	0.60	0.33	初級	2.60	2.50	2.20	2.67	上級
C02	1.00	0.25	0.60	0.17	初級	1.80	1.25	2.00	1.17	中級	2.80	1.50	2.00	1.33	中級
C03	0.40	0.00	0.00	0.00	入門	1.20	0.25	0.80	0.17	初級	3.00	1.75	2.00	0.67	上級
C04	0.80	0.00	0.20	0.00	入門	1.80	0.50	1.60	0.83	初級	2.20	1.50	2.00	1.67	中級
C05	0.80	0.00	0.60	0.17	入門	1.00	0.25	1.00	0.17	初級	2.80	1.00	2.20	1.00	中級
C06	0.80	0.25	0.40	0.00	入門	2.00	0.50	1.40	0.67	初級	2.60	0.75	2.00	1.33	上級
C07	0.00	0.00	0.00	0.00	入門	1.00	0.25	0.80	0.17	初級	1.80	0.75	1.40	0.33	初級

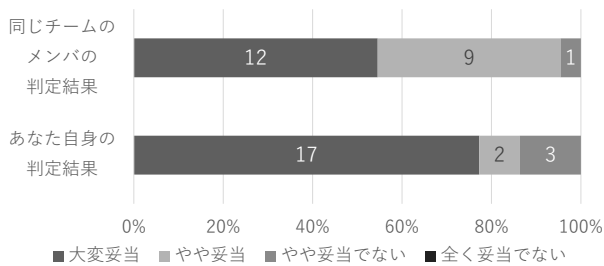


図 3: 判定結果に対するアンケート結果

3.4 GitHub 利用履歴との比較分析

主観評価に基づく習熟度評価と並行して、実験対象プロジェクトのうちの1つ(Bプロジェクト)を対象に、同じ時期の学生のGitHub利用履歴を取得し、分析した。Git/GitHubではコミットという単位でバージョン管理対象となるプログラムやドキュメントの変更を記録していく。コミットの粒度のばらつきには依存するものの、平均的にはコミット数の多い学生ほど、よくGitHubを使っていると言える。

表3はBプロジェクトの所属学生15名のプロジェクト期間中のGitHubコミット数を月ごとに示したものである。前期活動期間の5月から7月、夏休み期間の8月から9月は全体的にコミット数が少なく、開発が本格化した後期の10,11月にコミット数が急増している。この表では10,11月の合計コミット数の降順にBプロジェクトの学生を並べている。表の左側には3.2節で述べた3回分のスキル判定結果を再掲している。

コミット数が多い学生ほどスキルが向上することが想定されたが、コミット数が非常に多い4名のうち1名はコミット数が多いにも関わらずスキル判定結果は「中級者」に留まっている。コミット数が多くても使い方は上達していなかった可能性もあり、Git/GitHubの使い方のより詳細な調査が必要と言える。一方、コミット数が少ない学生はあまりスキルが向上しないことが予想されたが、10,11月のコミット数が30未満の下位学生の中にも開始時「入門者」から終了時「中級者」に上がった学生が3名、開始時「入門者」から終了時「上級者」に上がった学生が1名いた。コミットデータは「システム情報科学実習」でBプロジェクトが利用している複数のGitHubリポジトリから取得したが、学生に

よっては他の活動あるいは個人的に別のGitHubリポジトリを使用していることも考えられ、その結果スキルが向上した学生の可能性もある。

4 おわりに

先行研究において策定したアンケート調査に基づくGitHub習熟度評価手法を用い、学部3年生を対象とした通年必修のPBL科目において、開始時・中間・終了時の3回評価実験を行った。分析結果からPBL科目の進行に応じて各スキルの評価値が有意に上がっていることが分かった。また、中間段階で行った習熟度判定結果に対するアンケート調査でも、結果を妥当と考える回答が多かった。GitHub利用履歴データ(コミット数)との対比では、全体的な傾向としてはコミット数の多い学生がスキル向上していたが、一部に例外的な結果となった学生もいた。

アンケート調査に基づく評価のため、客観的な評価にはなっていないものの、評価実験を行った範囲では、対象学生のスキルを大よそ判定できていると考えられる。

今後は、より習熟度が高い場合の判定を可能とするアンケート項目の検討や、GitHubのコミット数以外のデータを用いた分析を行うほか、長期に渡るPBLではなく、他の活動が影響しにくい短期のハッカソンでの評価実験の実施を検討していく。

参考文献

- (1) enPiT事務局: “高度IT人材を育成する産学協働の実践教育ネットワーク”(2017)
- (2) 福田晃, 鶴林尚靖, 荒木啓二郎, 峯恒憲, 日下部茂, 金子邦彦, 亀井靖高, 廣重法道: “情報工学系大学教員のためのPBL実践ガイド”, 九州大学大学院システム情報科学府情報知能工学専攻社会情報システム工学コース(2012)
- (3) 日戸直紘, 伊藤恵, 大場みち子: “能力成熟度モデル統合に基づいたPBLにおける定量的学習評価手法の提案”, 日本ソフトウェア科学会第34回大会論文集, 日本ソフトウェア科学会(2017)
- (4) 松原克弥, 伊藤恵, 木塚あゆみ: “コード更新履歴の可視化による分散PBLでのチーム活動評価の試み”, 第4回実践的IT教育シンポジウム(rePiT2018)論文集, 日本ソフトウェア科学会, pp. 4955(2018)
- (5) 伊藤恵, 松原克弥, 富永敦子: “GitHub習熟度評価手法の提案とシステム開発PBLにおける習熟度評価の試み”,

表 3: コミット数との対比

学生	開始時	中間	終了時	5月	6月	7月	8月	9月	10月	11月	10,11計
B04	中級	中級	上級	0	16	1	0	16	80	48	128
B02	初級	上級	上級	0	0	12	0	6	51	62	113
B10	入門	初級	中級	0	0	0	0	1	64	37	101
B01	上級	上級	上級	0	0	0	0	1	61	30	91
B07	入門	初級	上級	0	0	0	0	0	16	28	44
B05	入門	初級	上級	0	9	0	0	0	17	25	42
B08	入門	初級	中級	0	0	0	0	0	28	7	35
B03	中級	上級	上級	0	0	0	0	0	23	10	33
B13	入門	初級	初級	0	0	0	0	0	4	23	27
B11	入門	初級	中級	0	0	0	0	0	5	21	26
B12	入門	入門	中級	0	0	2	0	2	17	8	25
B09	入門	入門	中級	0	0	0	0	0	14	6	20
B14	入門	初級	初級	0	0	0	0	0	15	4	19
B06	入門	中級	上級	0	0	0	0	0	2	10	12
B15	入門	入門	初級	0	0	0	0	0	0	0	0

日本ソフトウェア科学会第 36 回大会講演論文集, 日本ソフトウェア科学会 (2019)

(6) @kyanro@github: git 理解度チェックリスト,
<https://qiita.com/kyanro@github/items/3f0d4303882e668e1e85> (2019 年 12 月 1 日確認).

(7) 大塚弘記: GitHub 実践入門, 技術評論社, (2014)