

レビュースキル向上に向けた思考プロセス蓄積手法

- 優秀なレビューアの知見を AI に蓄積した学習サイクル「もぐつむ法」 -

研究員：西村 春香（株式会社日立製作所）
 中村 江里（ブラザー工業株式会社）
 木村 優花（株式会社日立ソリューションズ）
主査：中谷 一樹（TIS 株式会社）
副主査：上田 裕之（株式会社 DTS インサイト）
アドバイザー：安達 賢二（Software Quasol）

研究概要

本研究では、若手社員のレビュースキルを向上させるため、優秀なレビューアの思考プロセスを効率的に習得する手法として「もぐつむ法」を提案する。もぐつむ法は、優秀なレビューアの思考プロセスを AI に学習・蓄積させ、若手社員がその AI と共に繰り返しレビューを行い、結果を比較することで自身の弱点を把握し、改善を繰り返す学習プロセスである。検証として、この手法を若手社員 6 名に適用した結果、優秀なレビューアの思考プロセスの定着とレビュースキルの向上が確認できた。

1. はじめに

1.1 背景

ソフトウェア開発におけるレビューは、欠陥を早期に検出できる有効な手段として、品質向上に寄与し、多くの企業で広く採用されている。しかし、レビュー対象に対して、同じ時間をかけてレビューを実施しても、優秀なレビューアは重大な欠陥を検出できる一方で、十分なレビュー経験を持たないレビューアでは、それらを見逃してしまうことがある。

一般に、レビュー経験を積むことで欠陥検出の精度は向上すると考えられているが、その過程は個人の経験や属人的な学習に依存する部分が大きく、一定の時間を要する傾向がある。そのため、レビュー経験が浅い若手社員を優秀なレビューアへと成長させるには、実務を通じた長期的な経験の蓄積を前提とした育成になりやすいという課題がある。

そこで、若手社員の成長を早めることを目的に、レビュー経験の浅い段階から、優秀なレビューアが有する欠陥検出スキルを効率的に習得できる手法を提案したいと考えた。

1.2 研究の狙い

優秀なレビューアのスキルを効率的に習得するため、本研究では、業務経験年数 5 年以下の若手社員と優秀なレビューアの違いに着目した。優秀なレビューアは、ドメイン知識や欠陥知識を複合的に組み合わせ、それらに関連付けて活用し、欠陥検出につなげていると考えられる。一方、若手社員は知り得た知識をどのように使えば良いのかが分からない、あるいは活用したつもりでも実際には十分に活用できていないと考えられる。

そこで、本研究では、欠陥の見つけ方、すなわち、優秀なレビューアの思考プロセスがレビュースキル向上のカギであると考えた。また AI を活用することでそのスキル向上を加速できるのではないかと考えた。本研究では、「優秀なレビューアの思考プロセスを AI に学習させ蓄積し、若手社員の習得に活用することで、レビュースキルが向上する」という仮説を設定する。また本研究では機能仕様書をレビュー対象とする。

2. 提案する手法

2.1 提案する学習手法「もぐつむ法」

第 41 年度 ソフトウェアレビューコース（もぐつむグループ）

仮説に基づき、優秀なレビューアの思考プロセスを若手社員が効率的に学習する手法「もぐつむ法」を提案する。レビューにおける思考プロセスは暗黙的ノウハウであり、その継承は容易ではない。そこで本手法は、格言「やって見せ、説いてきかせて、させてみて誉めてやらねば、人は動かじ」^[1]に示される人材育成の本質を参考にしつつ、現代の動機づけ理論「モチベーション 3.0」^[2]踏まえた構成とした。“誉めてやらねば”に代表される他者からの報酬による動機付けではなく、自身の成長を確認しながら内発的動機付けを高める仕組みへと置き換えている。本手法における学習サイクルの工程を下記に示す。

(1) 事前準備

優秀なレビューアと若手社員がレビューを実施し、それぞれレビュー結果とレビュー時の思考プロセスを記述する。

(2) 思考プロセスの伝授

優秀なレビューアの思考プロセスを若手社員に伝える。

(3) データ蓄積と振り返り

優秀なレビューアの思考プロセスを AI に蓄積し、AI にレビューを実施させる。

AI のレビュー結果を学ぶとともに、レビュー結果を若手社員の振り返りに使用することで、前工程で得た知見を自身の理解として定着させる。

(4) スキル向上検証

若手社員と優秀なレビューア役の AI で新たなレビューを実施し、自身のスキル成長度合いを体感する。

この工程を学習サイクルとして反復するほど、優秀なレビューアの思考プロセスが AI に蓄積するため、効果が高まると考えられる。本手法の命名は、この様子が若手社員と AI が共に思考プロセスを「もぐもぐ」と食べ、繰り返すことでナレッジが若手社員と AI の中に「つみ」あがるように見えることに由来する。

2. 2 思考プロセスについて

レビュー時の思考プロセスとは、レビューアが頭の中でほぼ無意識かつ高速に行っている思考のメカニズムのことである。具体的には、「日付というキーワードが出てきた」、「過去に日付に関する障害があった」、「日付の概念を間違っていないか」、「このまま実装すると後工程で問題になりそうだ」といった思考の連鎖、欠陥検出に至るまでの思考の道筋を表すものである。

レビューにおける思考プロセスに関する先行研究として HDR 法^[3]が挙げられる。HDR 法とは、優秀なレビューアが普段から行う暗黙知としての思考メカニズムを手法として形式知化、体系化したものであり、「兆候」「仮説」「検証」という関係性に基づいて思考を繰り返すメカニズムを提唱している。

本研究では HDR 法を参考に優秀なレビューアの思考プロセスを下記の関係性と考える。

- 兆候：欠陥の存在を示唆するデータや、キーワードなど、レビュー対象を俯瞰的に見て、何か特徴的なことはないかを探す。
- 仮説：兆候からどんな欠陥が混入しているか仮説を立てる。
- 検証：仮説に基づき、的を絞って欠陥検出や次の関連する兆候を発見する。

HDR 法では普段意識していない思考を馴染みのない兆候や仮説といった構造に落とし込み言語化する必要があるが、この作業を容易にできるかはレビューア個人の能力や経験に依存している。そのため、兆候や仮説の落とし込み作業に労力を割かれ指摘数が減少するケースや言語化が上手くいかないケースが存在した^[3]。

しかし本研究では、優秀なレビューアの思考プロセスを AI にナレッジとして蓄積するため、AI が認識可能な形で具象化する必要がある。そこで、HDR 法で用いられる「兆候」や「仮説」といった馴染みの薄い用語を用いず、欠陥検出の気づきを得たキーワードや特徴、そこから連想した障害事例や仮説、欠陥検出までの思考の連鎖をまとめて

第 41 年度 ソフトウェアレビューコース（もぐつむグループ）

「指摘理由」として記述させる方式とする。また指摘理由は、具体的な形式は定めず、詳細にテキスト化することのみを求める。これは、作業負荷を軽減し、馴染みのない単語を除外し不慣れた作業を排除することを目的としている。

2. 3 AI に自社の障害事例をインプットする経緯

本研究では、AI に自社の過去の障害事例をインプットする。過去の障害事例をインプットする理由としては、次の 2 つの理由に基づく。1 つめは、何もインプットしない AI は思考プロセスをナレッジとして与えても一般的な知識や障害事例しか情報を持たないためである。この場合、レビュー対象に対して本来検出したい重大な欠陥を発見できない可能性がある。2 つめは、もぐつむ法では、AI がレビューを実施する際に AI は優秀なレビューア役を行うため、優秀なレビューアが持つ製品や業界固有のドメイン知識を持つ方が望ましいためである。

2. 4 AI のプロンプトについて

本研究では使用する AI やバージョンを固定しない方針を採用する。これは企業によって利用可能な AI に差があり、AI 自体の性能に違いがあるためである。そこで AI へのプロンプトは下記に示す指示の流れとする。詳細は付録の別紙 A に示す。

- (1) 役割と目的を伝える
- (2) HDR 法を教える（論文や要約をインプット）
- (3) 実際の思考プロセスを教える（もぐつむ法で実施したレビュー結果をインプット）
- (4) 参考にする過去の障害事例を教える
- (5) アウトプット内容とフォーマットを定義する
- (6) レビュー対象を渡してレビューを実施させる
- (7) 必要に応じてアウトプット粒度や内容の修正を依頼する

3. 検証・評価

3. 1 検証内容

提案手法「もぐつむ法」の有効性を検証するため、表 3-1 の検証参加者で 2. 1 節の全工程を 1 回実施し、結果を評価した。

表 3-1 検証参加者

参加者, AI	条件
対象者（若手社員）	経験年数 5 年以下のレビューア 6 名（3 社から各 2 名）
優秀なレビューア	経験年数 6 年以上のレビューア。レビュー対象につき 1 名以上
AI	各社で社内情報の入力利用の許可が得られている AI を選択 ※各社の事情により AI の種類・バージョンは、本論文には非掲載

本検証での AI への事前のインプット情報は、各社でのその製品及び類似関連製品の過去の障害事例とした。粒度やフォーマットは自由とした。

また、レビューは表 3-2 の条件で実施した。実際のレビューでは、優秀なレビューアはレビュー対象以外の情報（プロジェクトの進捗やレビュー対象の作成者のペルソナなど）を兆候とすることがある。そこで本検証では、レビュー対象の作成者に関するペルソナを事前情報として提供し、指摘理由に利用してもよいことを伝えている。

表 3-2 レビュー条件

項目	内容
レビュー方法	紙上レビュー

第 41 年度 ソフトウェアレビューコース（もぐつむグループ）

レビュー時間	30 分
レビュー対象	機能仕様書（A 社：IT 管理系システム，B 社：デバイス管理系システム，C 社：交通系システム） ※仕様書は実際の製品の機能仕様書をベースに，一般的な欠陥と過去の障害事例由来の欠陥を意図的に複数混入させ，レビュー時間内でレビューが可能な文字数に改変している。 ※仕様書には機密情報が含まれるため，本論文には非掲載
事前展開情報	レビュー対象の作成者に関するペルソナ。 （例）5 年以上続くプロジェクトの既存改修，作成者は新規参入者，作成者の経歴は 5 年目
レビューのアウトプット	指定したフォーマットの指摘票に，指摘箇所，指摘内容，指摘理由の 3 点を記載する。詳細は付録の別紙 B に示す。

3. 2 検証手順

本検証は，表 3-3 に示す手順で実施した。レビュー対象物の内容や難易度による影響を抑えるため，レビュー対象物 A および B を用い，参加者ごとに実施順を入れ替えて検証を行った。表 3-3 では 1 回目にレビュー対象物 A，2 回目にレビュー対象物 B のレビューを実施した場合の手順で示す。

表 3-3 検証手順

手順	実施内容	実施者	補足
(1) 事前準備			
(1-1)	レビュー 1 回目	若手社員 優秀なレビューア	若手社員はレビュー対象物 A のみ，優秀なレビューアは A 及び B に対し，30 分の紙上レビューを実施し，指摘箇所，指摘内容，指摘理由を指摘票に記入
(2) 思考プロセスの伝授			
(2-1)	対面での思考プロセス共有	若手社員 優秀なレビューア	優秀なレビューアから指摘内容および導出までの思考プロセスを共有
(3) データ蓄積と振り返り			
(3-1)	AI による学習およびレビュー	AI	AI に障害事例および指摘票 A をインプットし，レビュー対象物 A に対し，AI がレビューを実施
(3-2)	振り返り 1 回目	若手社員	若手社員が優秀なレビューアと AI の指摘票 A を参照し，追加指摘を記載
(4) スキル向上検証			
(4-1)	レビュー 2 回目	若手社員	レビュー対象物 B に対し，30 分間の紙上レビューを実施し，指摘箇所，指摘内容，指摘理由を指摘票 B に記入
(4-2)	AI による学習およびレビュー	AI	AI に障害事例および指摘票 B をインプットし，レビュー対象物 B に対し，AI がレビューを実施
(4-3)	振り返り 2 回目	若手社員	若手社員が優秀なレビューアと AI の指摘票 B を参照し，追加指摘を記載
(4-4)	採点	研究チーム	指摘票 A および B を所定の基準に従って点数化

3. 3 評価方法

本研究では、指摘内容を重大度に応じて点数化して定量的に評価した。点数化で使ったフローチャートを図 3-1 で示す。本研究では、軽微な欠陥よりも重大欠陥の検出を重視し、過去に発生した社外事故につながる可能性のある欠陥を「重大欠陥」と定義して該当する指摘には高い点数を設定した。この評価方法を用いて 3 社それぞれ評価した。評価は、下記の 2 つの観点から行った。

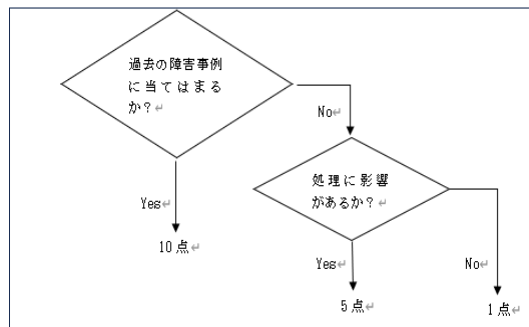


図 3-1 点数化で使ったフローチャート

(a) もぐつむ法による学習効果の評価

もぐつむ法による学習効果は、紙上レビュー1回目と2回目におけるレビューの点数の比較によって評価した。表 3-3 でいうと、手順 (1-1) と (4-1) の比較である。具体的には、2 回目のレビューの点数が 1 回目の点数を上回った場合、もぐつむ法による学習効果が現れたものと判断した。

(b) もぐつむ法による思考プロセス定着状況の評価

もぐつむ法による思考プロセス定着状況は、若手社員の全指摘件数に対する、優秀なレビューアおよび AI との同一指摘件数の割合について 1 回目と 2 回目を比較することで評価した。2 回目の割合が 1 回目の割合を上回った場合、1 回目の学習過程で参照した優秀なレビューアおよび AI の指摘内容が、若手社員の思考プロセスとして定着し、実際に 2 回目のレビューに反映できていると判断できる。

3. 4 実験結果

3. 3 節の (a) を確認するため、若手社員 6 名が実施したレビュー結果および振り返りを図 3-1 のフローチャートに従って点数化したものを図 3-2 に示す。図 3-2 に示す手順番号は、表 3-3 の番号と対応しており、各手順におけるレビュー結果を示している。なお、振り返りの点数は、追加指摘分のみを対象とし、自身のレビュー結果は含まない。

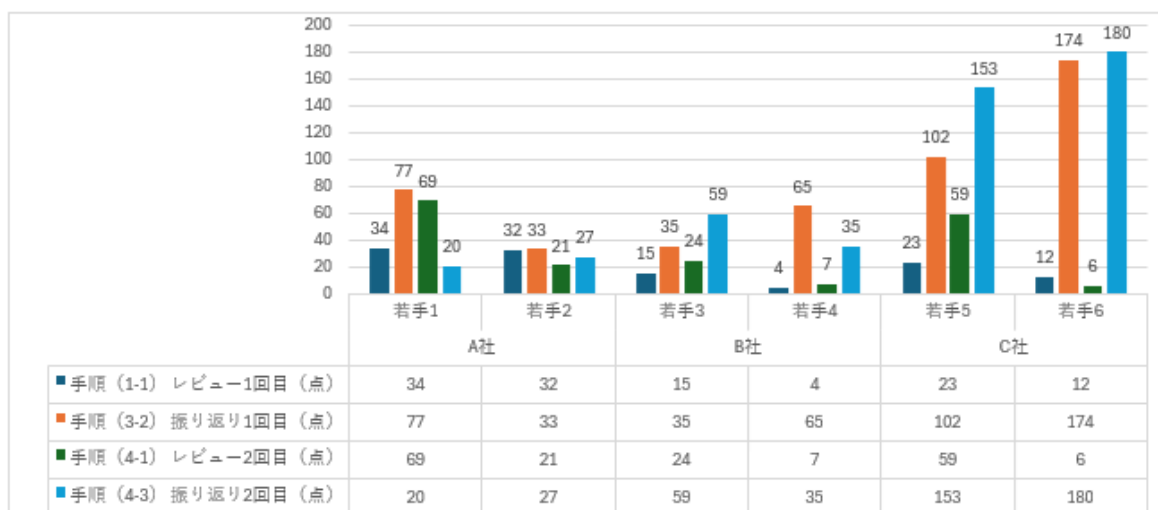


図 3-2 若手社員別 レビュー結果および振り返りの点数 (点)

また、図 3-2 に示す手順 (1-1) および手順 (4-1) の総得点のうち、重大欠陥に該当する指摘 (つまり、最大得点の 10 点) の割合を算出した。その結果を図 3-3 に示す。

第 41 年度 ソフトウェアレビューコース（もぐつむグループ）

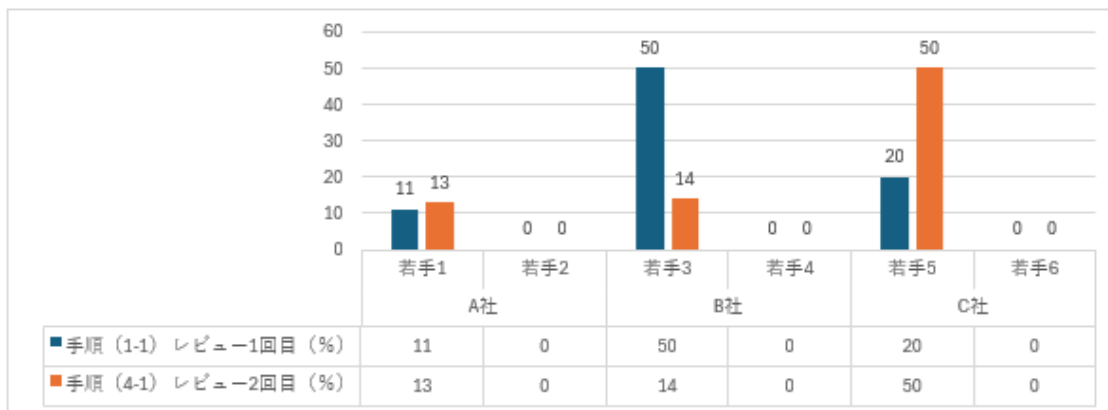


図 3-3 若手社員別 重大欠陥に該当する指摘割合 (%)

3. 3 節の (b) を確認するため、レビューの指摘内容を分析した。具体的には、1 回目と 2 回目のレビューの総指摘件数のうち、優秀なレビューアおよび AI との同一指摘がどの程度を占めるのかを割合で表した。その結果を図 3-4 に示す。

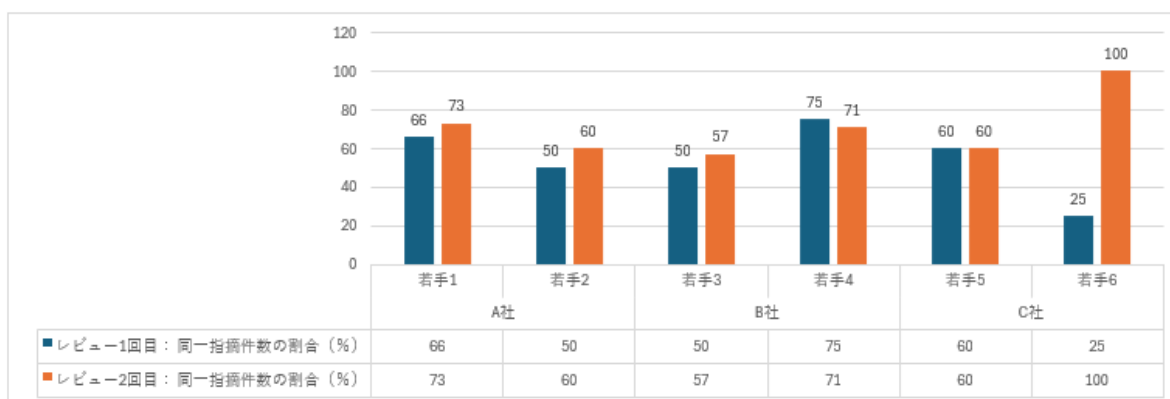


図 3-4 若手社員別 優秀なレビューアおよび AI との同一指摘割合 (%)

3. 5 評価

3. 5. 1 もぐつむ法による学習効果の評価

図 3-2 の手順 (1-1) と手順 (4-1) のレビュー点数の変化に着目すると、6 名中 4 名において、2 回目のレビュー点数が 1 回目を上回る結果となった。特に、若手 1 および若手 3 は大きく向上している。一方、若手 2 および若手 4 では下回る結果となった。

図 3-3 の重大欠陥に該当する指摘の変化に着目すると、6 名中 2 名が 2 回目の重大欠陥数が 1 回目を上回る結果となった。一方、若手 2, 4, 6 については変わらず 0% のまま、若手 3 では下回る結果となった。ただし、若手 3 の 1 回目のレビューは指摘件数 2 件中 1 件が重大欠陥に該当する指摘であったのに対し、2 回目は指摘件数 7 件中 1 件が重大欠陥に該当する指摘であり、両者では指摘件数の母数が大きく異なる。この結果は、2 回目のレビューで重大欠陥の指摘件数が減少したということではなく、レビュー全体の指摘件数が増加したことにより、相対的に割合が低下したものと解釈できる。

3. 5. 2 もぐつむ法による思考プロセス定着状況の評価

図 3-4 の若手社員のレビューにおいて、優秀なレビューアおよび AI との同一指摘の変化に着目すると、6 名中 5 名が 2 回目の同一指摘数が 1 回目を上回る結果となった。一方、

第41年度 ソフトウェアレビューコース（もぐつむグループ）

若手4はやや減少しているが、1回目のレビューは指摘件数4件中3件が同一指摘、2回目は指摘件数7件中5件が同一指摘に該当する指摘であり、両者では指摘件数の母数が異なる。この結果は、レビュー全体の指摘件数が増加したことにより、相対的に割合が低下したものと解釈できる。

3.6 考察

3.6.1 若手社員の点数の伸びが悪い要因

3.5.1節で述べた通り、若手社員6名中2名の点数については、2回目の点数が1回目を下回る結果となった。この要因を明らかにするため、若手社員と優秀なレビューアの指摘内容を比較・分析した。

分析の結果、優秀なレビューアは仕様書の記載内容そのものだけでなく、仕様書の外に存在する要因（以下、外的要素）に着目してレビューを行っていることが明らかとなった。例えば、仕様書作成者のペルソナ情報から考えられる行動を予測した指摘や、自社で過去に発生した障害事例を踏まえ、同様の問題が再発しないかという観点で指摘が確認された。具体的に挙げられた外的要素の内訳は下記の通りである。

- ・キーワード・図・文脈からの指摘
- ・仕様書作成者のペルソナ情報を考慮した指摘
- ・自社における過去の障害事例を踏まえた指摘
- ・後工程の作業者の立場（例：設計者、テスト設計者、テスト実施者）を考慮した指摘
- ・エンドユーザーの立場を考慮した指摘

優秀なレビューアの総指摘件数のうち、外的要素に着目した指摘がどの程度あるかを割合で示した。その分析結果を図3-5に示す。

一方、若手社員のレビュー指摘内容は外的要素への着目が不足している傾向が見られた。特に、仕様書の一部の記述のみを根拠として指摘を行うケースが多く、結果として指摘内容の視野が限定的になっていることが示唆される。これらのことから、若手社員のレビュー点数が十分に伸びなかった一因として、外的要素を考慮した思考プロセスを十分に持っていない点が挙げられる。なお、外的要素の内訳ごとの割合については付録の別紙Cに示す。

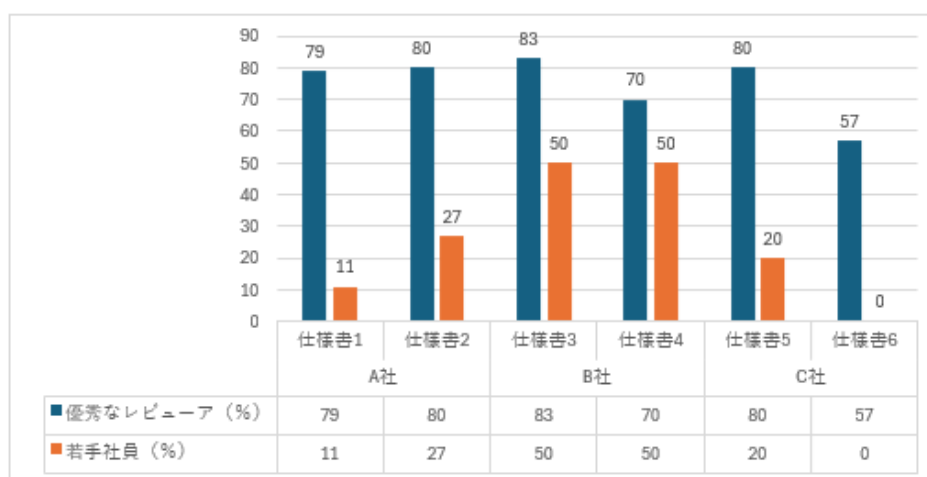


図3-5 仕様書別 外的要素に着目した指摘割合 (%)

3.6.2 若手社員の点数向上に向けた学習手法の考察

前節までの結果および分析より、若手社員のレビューの点数の伸びを改善するには、仕様書の記載内容に依存した限定的な視点から脱却し、外的要素を考慮した思考プロセスを

第 41 年度 ソフトウェアレビューコース（もぐつむグループ）

習得することが重要であることが示唆された。

この課題に対し、まずは、優秀なレビューアの思考プロセスをより多く AI に蓄積することが重要である。特に若手社員が弱点とする外的要素を考慮して指摘に至ったものを AI に蓄積し、「どのような観点で」「何を根拠として指摘に至ったのか」という思考プロセスと、その具体事例を多く積み上げていく。その上で、それらすべてを一度に習得するのは難しいため、徐々に習得していくことが望ましい。外的要素の種類や指摘に至るまでの思考パターンは多様にあるため、若手社員は、自身の不足している思考プロセスを学んでは実践で使い、習得できたら次の不足している思考プロセスの学習と実践へとコマを進めていく。このような学習サイクルを回すことで若手社員の確実な成長につながると考えている。

4. まとめ

本研究では、「もぐつむ法」を提案した。検証の結果、各社、対象システムや利用 AI ツールに違いがあるものの、レビュー 2 回目の点数が向上した参加者が多く、すべての会社でもぐつむ法の効果が確認できた。また、もぐつむ法によって自身に不足している思考プロセスを認識し、それを学習したうえで再度レビューを行うことで、思考プロセスが定着していく可能性が示唆された。

これらのことから、もぐつむ法は、若手社員が自身の伸びしろを把握し、思考プロセスを段階的に定着させていくための有効な学習サイクルとなり得る。

一方で、重大欠陥に該当する指摘については、もぐつむ法を適用した場合においても十分に検出できていないことが確認された。この結果は、もぐつむ方がレビュースキル全般の底上げには有効である一方で、重大欠陥の検出に必要な高度な判断や経験に基づく直感については、短時間の学習では十分に定着しにくい可能性を示している。

5. 展望

本研究では長期的に運用した場合の効果検証までは至っていない。今後は、提案したもぐつむ法を継続的に運用し、若手社員のレビュー点数の推移や指摘内容の質的变化を定量的に評価する必要がある。

また、本研究は若手社員の育成に加え、優秀なレビューアに対しても有効であると考えられる。もぐつむ法では思考プロセスの暗黙知を言語化することで、優秀なレビューアが自身の思考プロセスを客観的に把握できるようになり、無意識的に行っていたレビューを意識的に実施することが可能と考えられる。その結果、優秀なレビューアにとって自身の思考プロセスを自己分析することが可能となる。さらに、これらの思考プロセスを AI に蓄積していくことで、特定のレビューアに依存していた思考プロセスを組織全体で共有可能な形にでき、レビュー知識の属人化を防ぐ効果も期待できる。

近年の傾向として、今後ソフトウェアレビューを AI に全面的にゆだねる運用形態が進展する可能性がある。このような運用がすすんだ場合、個々のレビュー機会が減少し、レビューに必要な思考プロセスが十分に身につかないことが懸念される。その結果、AI が提示するレビュー結果の妥当性を人間が適切に判断できなくなるという、新たな課題が生じる可能性がある。この課題に対し、もぐつむ法は有効な打開策となり得ると考えられる。もぐつむ法は、若手社員が AI を通じて優秀なレビューアの思考プロセスを学習し、それらを自身の思考として内在化できると考えられる。この学習サイクルを繰り返すことで、AI が提示するレビュー結果についても、人間がその妥当性を判断できると考えられる。

参考文献

- [1] 新人物往来社（編）、山本五十六のすべて、新人物往来社、1985
- [2] ダニエル・ピンク（著）、大前 研一（訳）、モチベーション 3.0 - 持続する「やる気！」をいかに引き出すか、講談社、2015
- [3] 上田 裕之、高橋 功、高橋 実雄、中谷 一樹、HDR 法：仮説駆動型レビュー手法の提案 - HDR 法の実践による生産性と品質の同時向上 -、ソフトウェア品質管理研究会、2012