

最適なレビュー実施と現場が喜ぶメトリクスの研究

Research of metrics with which the best Review execution and site are pleased

主査 三浦 邦彦 矢崎総業（株）
副主査 藤巻 昇 （株）東芝
リーダー 中野 誠 T I S （株）
小澤 廣晃 （株）エスシーシー
吉田 直美 （株）日立製作所
伊藤 秀樹 アビームシステムズ（株）
有田 亮一 沖通信システム（株）

1. 研究概要

近年、ソフトウェア製品市場では顧客ニーズの多様化が進んでいる。その為、開発者側では今まで以上の生産性の向上と安定した高品質の確保が求められている。一般的な対応として、開発上流工程において成果物の不具合を摘出し対応を実施する事が考えられる。その一手段として活用されている活動にはレビューがある。

しかし、レビューをただ実施するだけではなく、これまで以上の品質確保の考慮および即時に対応を実施する事が必要となっている。そこで、本研究ではレビュー実施とその結果から即対応が出来るレビュー品質の評価を定量化する方式を検討しパイロットを実施した。

Abstract

Recently, the diversification of the customer needs is advanced in the software product market. Therefore, securing an improvement of productivity more than before and a steady high quality is requested on the developer side. It is possible to execute correspondence by the removal of the trouble of the product in the developmental style process as general correspondence. There is a Review in the activity that has been used as the one means.

However, the review is not enough by itself. More than ever, it is necessary to consider securing the level of quality and actions must be taken quickly based on the results of the review. We have created a method to quantify the evaluation of the quality of the review itself and allow for actions to be taken immediately; this method has been examined, and a pilot has been executed.

2. テーマ選定の理由と背景

近年、プログラムが高度化、複雑化、大容量化されるに従ってソフトウェア製品の品質見極めも複雑化しつつある。特にテスト工程でのバグは、バグ原因が上流工程（仕様／設計）段階の場合は、手戻り工数が発生し、結果として予算の超過や、納期遅延となる場合がある。そのために、いかに上流工程で品質を作り込むか、上流工程で品質を見極められるかが重要である。一方、現場では上流工程での品質確保を図るべく、設計レビュー時のレビューア、レビュー時間の確保をするものの、時間の効率化、指摘事項の質向上といった、常に生産性を意識した行動を求められる。本グループメンバーは、開発部門が半数以上と現場に近い部門で構成された。そこで、これらの問題を解決するために、レビューにおける成果物の品質評価を定量化することによって、現場自らが判断を行い適切な処置が行えること、つまり「現場が喜ぶメトリクス」を目的に以下の段階でのメトリクスについて研究を行った。

- ・ レビュー実施時の指摘事項の整理（指摘と原因分類）
- ・ レビュー実施後の評価（レビュー成果物の品質評価）

3. 活動目標

「現場が喜ぶメトリクス」を目的に、次のことを本グループの目標とした。

- 1) 現場が喜ぶレビュー指摘と原因分類の定義
- 2) 設計成果物の品質評価の精度向上
- 3) レビュー評価結果からのフォロー、安心感の提供

4. 活動内容

この一年の活動内容を「表 1:活動経過」に示す。

表 1：活動経過

日程	活動内容
2008/04/18	・ 研究会メンバー自己紹介と検討課題洗い出し。グループ分け
2008/05/22	臨時会開催 ・ 各社の現状課題の意識あわせ実施 ・ 工程の区分と測定指標（メトリクス）の雛形を検討し併せて各社の課題について意見交換
2008/06/08	・ 各社課題から絞り込み、スコープを検討（レビュー実施と終了判断）

2008/07/10～11	<ul style="list-style-type: none"> 計画段階からのレビュー実施手順とレビュー実施後の終了判断、及び、 工程終了判断のメトリクスの検討 併せて現場に役立つセルフメトリクス評価方法の検討
2008/09/16	<ul style="list-style-type: none"> 試行に向けての準備（測定データ、データ区分の検討）
2008/11/14	<ul style="list-style-type: none"> 測定データの試行および評価 試行結果の検討、課題の整理と方向性を審議① 論文骨子の検討
2008/12/19	<ul style="list-style-type: none"> 試行結果の検討、課題の整理と方向性を審議② 論文骨子の検討
2009/01/09	<ul style="list-style-type: none"> 試行結果の検討、課題の整理と方向性を審議③ 論文骨子の検討
2009/01/09 ～ 02/01	<ul style="list-style-type: none"> 報告書作成 報告書レビュー
2009/02/02	<ul style="list-style-type: none"> 報告書提出

5. 研究成果および考察

本分科会ではV字型開発モデルにおけるソフトウェア開発での上流工程・外部設計成果物の品質を確保する方法について研究を行った。研究のスコープは成果物レビューで取得し得るメトリクスの活用とした。

5.1 成果物品質を評価する上でのメトリクスの見直し

5.1.1 レビュー指摘分類

レビュー結果の評価を行う上で、設計本質にかかわる指摘なのか、単に表記上の誤りで本質にかかわるものではないものを分別して成果物品質の評価の精度を向上させる目的として、レビュー指摘分類を下表の5パターンとした。

その観点は次の2点である。

- 設計成果物に対する指摘としてその設計内容については「漏れ」「相違」「あいまい（冗長）」で分類できると考えたこと
- 「ケアレスミス」や「記載に当たっての標準化ルールに沿っていない」といった設計書表記上による指摘を分別したかったこと

表 2：レビュー指摘分類

レビュー指摘分類名	定義
記述漏れ	設計成果物に記載されるべき事項が書かれていない
記述相違	設計成果物に記載されている事項が正しくない

記載あいまい	設計成果物に記載されている内容が冗長で複数の解釈ができる
パンチミス	誤字脱字や表記上の誤りで、設計内容に問題はないと判断されるもの
標準化違反	プロジェクトでの記述ルールに沿っていないもの

指摘分類を行うに当たっては、その内容が‘誤字脱字’や‘コピーペーストによる誤りのコピー増幅’についてはパンチミスに含まれるとしてカウントした。

5.1.2 レビュー指摘原因分類

指摘の原因区分をより細かく分類して不備を発生させた傾向を具体的に捉えやすくすることを目的として、見直し前の4パターンから8パターンとした。その観点は次の2点である。

- ・ 上流工程成果物が原因と考えられる場合、自工程での設計担当者には問題があるケース、上流工程成果物に問題があるケースを分別することによってレビュー対象成果物の品質評価の質が高まると考えたこと
- ・ 方式面や環境面での知識不足／理解不足が原因となる指摘は、業務要件に対する理解不足に比べて指摘内容の影響が大きいと考えたこと

表3：レビュー指摘原因分類

レビュー指摘原因分類名	定義
上流工程成果物の誤り	インプットとなる上流工程の成果物に誤りがあったことが原因
上流工程成果物の内容理解不足	上流工程の成果物に対する理解不足が原因
上流工程成果物の内容誤解	上流工程の成果物内容を正しく理解できなかったことが原因
環境理解不足	開発環境を理解できていないことが原因
実装方式スキル不足	開発方式に対するスキルが不足していることが原因
業務要件理解不足	業務要件を理解できていないことが原因
仕様漏れ	顧客からの業務要件仕様に漏れがあったことが原因
ケアレスミス	不注意による表記ミス

5.1.3 レビュー指摘密度（件／10 ページ）

指摘事項の発生割合のメトリクスである「レビュー指摘密度」は、指摘件数を10 ページあたりの発生件数に換算したものとした。

5.2 試行結果

前項記載にしたがって見直したメトリクスを活用して、実際の開発プロジェクトでの上流設計工程成果物レビューに適用し施行とした。

<参考：施行対象のプロジェクト内容>

- ・システム形態 クライアント／サーバー型
- ・開発規模 120人月
- ・業務 金融・営業支援系（同規模の6サブシステムから構成）
- ・開発形態 V字型開発

5.2.1 指摘分類の見直しの実施

従来指摘の内容を「漏れ」「誤り」「あいまい」という点から分類していた結果に対して、一次精査として明らかな「誤字脱字」や表記のケアレスミスを分別することにより、指摘分類・原因をより設計内容に絞っての分析が可能となった。表4に、分別による件数の分布を示す。

表4：指摘分類見直し

<従来の指摘区分>			<見直した指摘区分>	
指摘分類	件数	→	指摘分類	件数
記載漏れ	77		記載漏れ	59
記載誤り	142		記載誤り	115
記載あいまい	52		記載あいまい	31
			パンチミス	5
			標準化違反	61

この結果、評価分析としては見直し後の「記載漏れ」「記載誤り」「記載あいまい」に分類されたものを対象とした。66件（24%）は評価分析対象外となった。

※この66件は「漏れ」「誤り」「あいまい」との指摘になっていたが、内容の再監にて誤字脱字等のパンチミス・標準化違反であることが明らかであるため設計内容の不備から分別した。その上で設計者にミスの傾向が見られれば改善指導を個々実施した。

5.2.2 レビュー指摘原因分類による見直し及び再監者による精査結果

原因区分の見直しとあわせて、一次評価者と再監者との見直しを行った。区分の見直しによる原因の細分化と再監による見直し結果は以下のとおりである。

表 5：指摘分類見直し結果

<従来の原因区分>		<見直した原因区分による一次評価者分類>		<見直した原因区分による再監結果>	
原因区分	件数	原因区分	件数	原因区分	件数
前工程設計書の誤り	73	上流工程成果物の誤り	73	上流工程成果物の誤り	14
前工程設計書の理解不足	88	上流工程成果物の内容理解不足	44	上流工程成果物の内容理解不足	81
仕様漏れ	3	前工程設計書内容誤解	23	前工程設計書内容誤解	60
ケアレスミス	41	環境理解不足	3	環境理解不足	5
		実装方式スキル不足	8	実装方式スキル不足	8
		業務要件理解不足	3	業務要件理解不足	10
		仕様漏れ	4	仕様漏れ	3
		ケアレスミス	41	ケアレスミス	26

分類と再監により、4項目の原因区分からは傾向が捉えにくかった状況から以下の傾向を掴むことが可能となった。

- 一次評価の際には上流設計成果物の不備に起因する件数が多かったが再監の結果、自工程設計作業に原因がある件数のほうが多く上流設計成果物の品質には大きな問題はないと判断
- システム実装方式や環境面に対する知識／スキル不足による件数は割合として大きなものではないのでこの面での懸念は小さいこと

また、上流工程成果物の理解不足や誤解については、再監者を固定化することにより全レビュー記録に対してサブシステム間を横断して同一観点で再監することが可能となり、評価判断が統一化されるとともに各担当者の傾向も見ることもできて個別指導を行う効果も確認された。

5.2.3 レビュー指摘密度のサブシステム間相対評価の実施

指摘件数・指摘率をサブシステム間で比較評価し、他システムと乖離している値となったものについては、差異が出たことの原因を精査して定性評価を行った。図1に表記されたデータ中にある、3と4は他サブシステム（1，2，5）と比べて指摘率が低く、レビュー・プロセスに問題があった可能性があるため、指摘の内容と開発機能の難易度についての考察を行った。その結果指摘率が低くなった原因としては以下の2点によるものであり、品質上の大きな問題はないと判断をした。

- 既存サブシステムの類似機能開発であるため、過去の開発資産が活用できた事
- 帳票出力が主機能で開発難易度が他サブシステムに比べて低かった事

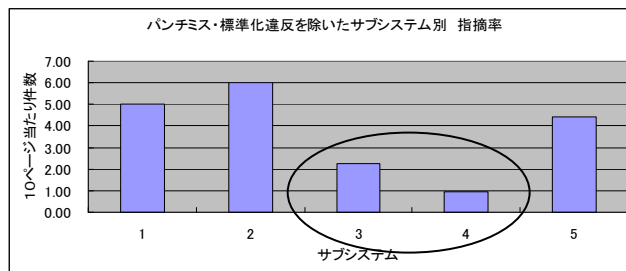


図 1：指摘率

5.3 考察

総括として、レビュー指摘分類、指摘原因区分を再定義することは、

- ① 指摘が設計内容の本質にかかわるものか（表記上の不備などの）本質にはかかわらないものかの分別の明確化
- ② 原因区分の細分化により設計担当者へのフィードバックをより具体的に実施することが可能

等の効果があり、成果物品質に対する評価の精度を上げるための方法として有効であったと考える。以下詳細を記述する。

5.3.1 誤記・標準化ルール違反を除外することの効果

- ・ 評価の対象を設計内容に合わせることで可能となり客観的な評価が可能

5.3.2 レビュー結果再監によるレビュー結果精度向上

- ・ 複数人からなる一次評価に対して、同一視点での評価をプロジェクト・リーダーが実施することにより評価のばらつきを抑えて平準化することを実現
- ・ 一次評価との差異が出た場合には評価者との評価方法の差異是正により精度の向上とプロジェクト内での評価方針の統一化を実現

5.3.3 原因区分の細分化・プロジェクト特性に則した内容への変更によるレビュー結果精度向上

- ・ 標準的な区分を細分化し、また選択肢としてプロジェクト特性へのテーラリングを行うことにより指摘の原因がわかりやすくなり、またお仕着せでない具体的な原因の特定化を行うことが容易
- ・ 「その他」を原因区分として原則使用禁止としたことにより具体的な原因追求を実施

5.3.4 相対的評価による特異な結果からの問題サブシステムの検知

- ・ サブシステム間で指摘率について突出傾向が見られる場合は、基準値との乖離とは別に突出傾向の原因分析を行うことにより問題傾向の把握が可能

5.3.5 指摘結果のフィードバックによる設計者との情報共有化拡大

- ・ レビュー結果は設計者及びプロジェクトチーム内への評価報告による設計書品質の情報共有化を実現

最後に補足するならば、これまでの品質評価活動を通じてレビュー結果の評価は定量評価の数値が一人歩きする傾向を感じていたが、今回その結果に加えて定性評価を行い、その結果をチーム全体に連携することにより、より設計成果物の品質に対する理解・問題点の理解が全体として深まり、次工程での留意事項として活用されることに

つながった、と感じる。指摘の原因を細分化してその背景を定性評価し、その内容の共有化はチームの弱点やプロジェクトの留意点を明らかにすることにつながり、弱点対策の立案に対しても効果があったと感じる。

5.4 今後に向けて

今回の活動を通じて、今後も考察を続けていく事項を付記する。

- ① 基準値の収集は収集基準を定めて行わないと基準にならないのではないかと？
(基準値は収集対象を長く広くすれば多くのプロジェクトから収集が可能であるが、収集基準をあわせておかないと精度にムラのあるデータ群から求められた基準値となり、結局「使えない」ものになってしまう可能性があると思料する)
- ② 現場にとってどのようにフィードバックをすすめることが効果があるのか？
- ③ 今回の上流設計の品質評価が最終成果物の品質向上にどのように寄与したかはトレースが必要であり、稼働後の品質状況を踏まえての追加評価を行うことがより効果的であろう。
- ④ レビュー結果の再監負荷・定性評価に関するプロジェクト・リーダーの作業負荷が大きくなるため、一定規模以上になった場合の評価品質をどのように担保すればよいか？

6. 終わりに

今回の研究会には、色々な業務環境で働いている方が集まりメトリクスについて研究を進めた。その中で、我々はレビューに関わるメトリクスに焦点を絞った研究活動を実施した。今回選定したテーマであるレビュー成果物の品質評価を定量化する事については、なかなか議論が進まなかった。しかし、その議論をする機会があったからこそ、メンバー自身の業務への気づきを与えてくれた。今回の考察を元に各メンバーは今後も継続して検討し、自身の業務への展開を実施する。

7. 参考文献

- [1] 日科技連出版社, 21世紀へのソフトウェア品質保証技術, 菅野文友 吉澤 正監修
- [2] 翔泳社, PSP ガイドブック, Watts S. Humphrey