

## レビュースキルを飛躍的に向上させるトレーニングの提案

～ レビューアのレベル定義とレベルに合わせた効率的なトレーニング ～

## Proposal of the training method to improve reviewer's domain knowledge

- Define level of reviewers and propose efficient training way which matches the reviewer level -

2014 年度 SQiP 研究会第 3 分科会

The 3<sup>rd</sup> Section Meeting Team of SQiP Study Group in 2014

主査	中谷 一樹 <sup>1)</sup>		
副主査	原 佑貴子 <sup>2)</sup>	上田 裕之 <sup>3)</sup>	
アドバイザー	森崎 修司 <sup>4)</sup>		
研究員	○藤田 延介 <sup>5)</sup>	山田 将貴 <sup>6)</sup>	原山 拓実 <sup>7)</sup>
CHIEF INVESTIGATOR	: Kazuki Nakatani <sup>1)</sup>		
SUB-CHIEF INVESTIGATOR	: Yukiko Hara <sup>2)</sup>	Hiroyuki Ueda <sup>3)</sup>	
ADVISER	: Shuji Morisaki <sup>4)</sup>		
RESEARCHER	: ○Nobuyuki Fujita <sup>5)</sup>	Masaki Yamada <sup>6)</sup>	Takumi Harayama <sup>7)</sup>

**Abstract** For success of peer review in software development, an expert reviewer is necessary. We studied how reviewers to learn domain knowledge and to train more high quality reviews.

We embedded intentionally errors in documents in the last fiscal year SQiP Study Group, has propose a method to review it. But there is a problem that is an irregularity in effect by the person who receives the training.

We have considered the reason for the uneven effect. We studied training techniques tailored to the proficiency of the reviewers that applies a technique called scenario reviews. We propose a training method for reviewer by this method, confirmed effect.

**研究概要**

ソフトウェア開発において、有識者によるレビューは非常に重要である。我々はレビューアに早期にドメイン知識を習得し、より質の高いレビューができるようするトレーニング手法について研究してきた。

我々は昨年度 SQiP 研究会にて仕様書に意図的にエラーを埋め込み、それをレビューする手法を提案したが、トレーニングを受ける人により効果にばらつきがあった。

効果にばらつきがみられる原因について考察し、シナリオレビューという手法を応用したレビューアの習熟度に合わせたトレーニング手法について研究した。この手法によるレビューアのトレーニング手法を提案し、効果を確認した。

---

株式会社エスイーシー 情報通信事業本部 札幌システム事業部  
 SEC Corporation Information Communication Div. Sapporo System Dept.  
 北海道札幌市中央区南一条西 19 丁目 1-337 SEC 札幌ビル TEL: 011-640-7787 e-mail: nobu@secnet.co.jp  
 SEC Sapporo Bld., 1-337, Minami1-jo Nishi19-chome, Chuo-ku, Sapporo-shi, Hokkaido, 060-0061, Japan TEL: 011-640-7787

- |                    |   |
|--------------------|---|
| 1) T I S 株式会社      | TIS Inc.  |
| 2) 日本アイ・ビー・エム株式会社  | IBM Japan, Ltd                                    |
| 3) 株式会社 D T S      | DTS CORPORATION                                   |
| 4) 国立大学法人名古屋大学     | National University Corporation Nagoya University |
| 5) 株式会社エスイーシー      | SEC Corporation                                   |
| 6) キヤノンファインテック株式会社 | Canon Finetech Inc.                               |
| 7) アンリツ株式会社        | Anritsu Corporation                               |

## 1. はじめに

ソフトウェア開発において、レビューを成功に導くためには有識者、とりわけエキスパートレビューアの存在が欠かせない。作成担当者にはない知識、経験からの知見が必要なためである。

しかし、エキスパートレビューアは少数かつ多忙のため、全てのレビューに参加することができない。また、エキスパートレビューアは若手に知識や経験をどのように伝えるか頭を悩ませている。プロジェクトに余裕があれば、若手には試行錯誤、つまり実務の中で失敗させることで成長を促すことができる。しかし、近年の開発現場は短納期で失敗が許されない環境であり、若手に失敗経験を積ませる余裕が無い。そのためエキスパートレビューアに負荷が集中し若手に教育する十分な時間が無いという問題がある。

先行研究では、レビューの知識を体系化する議論<sup>[1]</sup>やエキスパートレビューアの欠陥検出メカニズムについて議論<sup>[2]</sup>されている。また、エキスパートレビューアの成功を若手レビューアが共に経験し、成功体験から知識や経験を伝える仕組み<sup>[3]</sup>が挙げられる。

我々は、昨年度の研究にて、現場で実際に使用している仕様書にわざと欠陥を埋め込み、これを教材として使用するトレーニングする手法、EIDeR-Training 法（Error Injected Document Review - Training 法）を提案した<sup>[4]</sup>。しかし、EIDeR-Training 法にはトレーニングの受け手によって効果にばらつきがあり、実用性を示すに至らなかった。特に教材によっては、若手レビューアへの効果を確認できないケースもあった。これは、トレーニングの受け手のレベルとトレーニング内容がマッチしていなかったことに起因する。

我々は、前述の問題を解決しレビューアのレビュースキルを飛躍的に向上させるトレーニング方法を検討するため、以下の課題（RQ1-3）を設定し研究した。

- RQ1. 若手レビューアを（短時間で、かつ実害のある失敗はさせずに）レビューを任せられる人材に成長させるにはどのようにしたらよいか
- RQ2. レビューアのレベルを定義し測定することは可能か
- RQ3. レベル別に観点を絞った教材を用意することで、トレーニングの効果をあげられないか

## 2. 問題点と提案手法

はじめに、EIDeR-Training 法を考案するにあたって挙げた問題点を整理する。EIDeR-Training 法を考案するにあたって、はじめにエキスパートレビューアの知識を整理した。エキスパートレビューアに必要とされる知識は、表 1 に示す 3 種類に分類できると考える。レビューの知識、一般的な技術知識は、独学や前プロジェクトの経験から身に付けることができる。しかし、ドメイン知識は体系化されていることが少ないため、業務を通じて身に付けるのが一般的である。

表 1：有効なレビューをするための 3つの知識

必要な知識	概要	学習方法
レビューの知識	有効なレビューをするための知識。	レビューオリエンテーションキット <sup>[1]</sup> 間違いだらけの設計レビュー <sup>[5]</sup> など
一般的な技術知識	広く知られている技術的な知識。 すでに体系化、書籍化されている。	対象の参考書 OJT（On the Job Training）
ドメイン知識	専門分野の知識。 専門知識が体系化されない。	OJT

我々は、体系化されにくいドメイン知識に着目した。着目した理由は、レビューにおいては、ドメイン知識に由来する指摘が重大な欠陥に繋がることが多いと考えたためである。また、ドメイン知識は定義のあいまいな言葉である。表 2 に本研究で扱うドメイン知識の定義を示す。人や組織に関わるドメイン知識を廃し、技術的な面にフォーカスをあてた。若手レビューアは、OJT を通して、ドメイン知識を習得している（表 1）。ドメイン知識は、

体系化されていることが少ないため習得に時間をかけている。また、OJTによる教育には以下2点の問題がある。

(1) 実務においては失敗が許されない問題

実務においては、納期が厳しいこと、失敗すると信頼を失うことから、組織は若手レビューアに試行錯誤をする余裕が与えられない。そのため、レビューではエキスパートレビューアが多く発言するようになり、若手レビューアの成長を阻害してしまう。

(2) エキスパートレビューアが教育に時間をかけられない問題

エキスパートレビューアに負荷が集中しているため、エキスパートレビューアは、若手の教育ができない。

表 2：本研究におけるドメイン知識の定義

定義したドメイン知識	具体例
レビュー対象の知識	ソフトウェアが対象とする業務への理解。 病院内向けの電子カルテシステム開発の場合、院内の業務プロセス自体の知識。
ステークホルダの関心	ユーザやシステムに関係する人の関心への理解。 発注した病院によって重視するものが違い、他院に通用した仕様が受け入れられないなどの知識。 例えば端末をオフラインにしても使えることを前提とするか、常にオンラインであることを前提とするかなど設計思想が異なる。
常に変化し続けている知識体系	時間とともに体系そのものが変わる特徴を持つ。 当初紙ベースでのコミュニケーションが電子化、多端末での同時参照、患者情報の入力や検索方法など要求や機能が増え、変化し続ける。

これらの問題を解決するために EIDeR-Training 法を提案し、実用性を検証するにあたってレビュー実験を行った。昨年度の実験は、若年層からエキスパート層までのトレーニーに対し、EIDeR-Training 法の教材を使って2回レビュー実験を行った。その結果、1回目よりも2回目のレビュー実験の欠陥指摘率が上がったケースと下がったケースが出た。効果にばらつきが出た原因はレビューアのレベルが混在しており、教材の難易度がトレーニーのレベルと合っていないと考えられる<sup>[4]</sup>。

## 2.1 解決したい問題・課題

エキスパートレビューアに負荷が集中する問題を解決するには、若手レビューアを育成する必要がある。有効なレビューを行うために必要な知識は、3種類あることを示し、レビューの知識については先行研究<sup>[1]</sup>で検討されており、一般的な技術の知識についても、独学で習得できることを示した(図1)。若手レビューアが、現状より重大欠陥が指摘出来るようになること、エキスパートレビューアが保有する一部のドメイン知識を身に付けることを狙い(図2)、EIDeR-Training 法を提案した。しかし、昨年度の実験結果から効果が現れない場合があることがわかった。そこで EIDeR-Training 法を用い効率のよいドメイン知識の習得を目指すにあたり、課題(上記 RQ1-3)を解決し、重大な欠陥を指摘できる人材にトレーニングする方法について研究した。

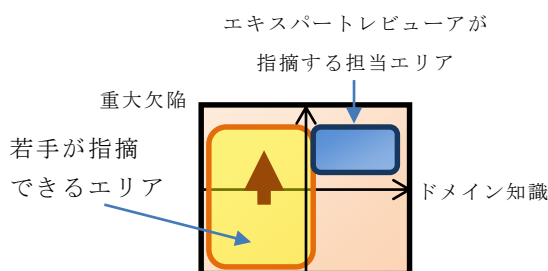


図 1：レビューの知識の習得

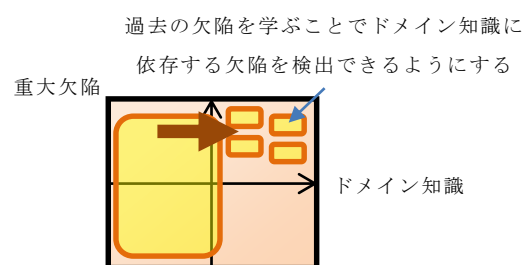


図 2：本研究のめざすところ

## 2.2 提案

我々は、若手がレビューを任せられる人材に成長するために、以下2つの手法を組み合わせることを提案する。

1つ目は、シナリオレビュー<sup>[5]</sup>により具体的な役割を与えてレビューに参加させる。

2つ目は、EIDeR-Training法を用いる。エキスパートレビューアは、若手レビューアに任せたい新規プロジェクトに類似した過去のプロジェクトの仕様書から、解答(図3)と教材を作成する。教材は、欠陥があった状態の仕様書を用意すると手軽である。解答と教材を作成する際に使用する欠陥は、トレーニング対象者のレベルに合わせて選択する。



図3: EIDeR-Training法 解答イメージ

昨年度の研究では、レベルの異なるレビューアに対して、同一のトレーニング教材を与えて実験を実施した。このため、想定通りの結果が得られたレビューアと効果の出なかったレビューアがいた。そこで、期待するレビューアの到達点を表3のように定める。なお、ITSSではドメイン知識に関する定義はないためITSSとのレベル対比は本研究独自のものである。

EIDeR-Training法の教材を作成する際には、それぞれレビューアのレベルに合った欠陥情報を選択し、教材を作成する。昨年度の研究の実績からEIDeR-Training法の教材を作成する時間は約20分であり、欠陥情報さえ揃っていれば教材作成の負担は軽い。

レベル0のレビューアは、まだレビューの経験が浅く、レビューに関する知識も少ない。したがって、レベル0のレビューアに必要なものは、レビュー文化の浸透であると考えられる。そこで、この層についてはレビューオリエンテーションキット<sup>[1]</sup>による教育や、レビューの演習(インスペクションのロール・プレイング)を実施する。

レベル1のレビューアは、レビューに関する知識はあるがレビューの経験が浅く、欠陥の発見に慣れていない。そこで、EIDeR-Training法を使用したトレーニングによって、レビューの疑似体験を積む。また、このレベルのレビューアはドメイン知識よりも一般的な知識をまず十分に蓄積する必要がある。したがって、教材に埋め込む欠陥は、ドメイン知識に基づいた欠陥ではなく、一般的な知識に基づいた欠陥に限定する。

レベル2のレビューアは、ドメイン知識に基づいた指摘を期待している。レベル2に対してもEIDeR-Training法を使用したトレーニングを実施する。教材に埋め込む欠陥は、過去の欠陥情報やドメイン知識に基づいた欠陥とする。ただし、欠陥の範囲は担当機能部単体で起こるものなど、影響範囲の狭いものに限る。

レベル3のレビューアには、製品全体に対する指摘を期待している。レベル3に対してもレベル2と同様にEIDeR-Training法を使用したトレーニングを実施する。ただし、教材に埋め込む欠陥の内容は、複数の機能部が関連するなど、影響範囲の広い欠陥とする。

レベル4のレビューアに関しては、レベル3のレビューアが経験を積むことによって到達できるレベルであり、トレーニングの範囲外と考える。

本研究では、EIDeR-Training法のトレーニング教材を作成する際にシナリオレビューにおけるシナリオを流用する。シナリオレビューでは、シナリオを複数のレビューアで分割して担当する方法を一般的に用いる。本研究では、シナリオの各観点をレベル分けし、

それぞれのレベル毎にトレーニングすることで、効果的にレビューアのトレーニングができると考えた。また、トレーニングの教材を表3のレベルに合わせて作成することで、各レビューアの実際のレベルを測定可能と考えた。

表 3：レビューア・レベルの定義

レベル	レビューにおける役割	定義	教育方法	ドメイン知識の範囲
レベル 0	なし	未経験者 レビュー技術、技術知識、ドメイン知識が十分ではない。	レビューオリエンテーションキット <sup>11)</sup> ロール・プレイング	なし
レベル 1	レビューの参加者	開発メンバ・テスト実施者 ITSS レベル 1 相当 レビュー技術についての知識を習得している。 開発するソフトウェアに使用する設計手法、開発言語、動作環境、テスト方法、テスト環境について知識を所有している。 知識に基づき一般的な欠陥を指摘できる。	EIDeR-Training 法	なし
レベル 2	レビューの参加者	開発メンバ・テスト実施者 ITSS レベル 2 相当 レビューに複数回参加し、欠陥を指摘した経験がある。 開発するソフトウェアの担当部分の機能、構造、特徴、特性を理解している。 担当部分の機能に基づき、欠陥を指摘できる。	EIDeR-Training 法	担当部分のみ
レベル 3	レビューの責任者	開発メンバ・テスト設計者 ITSS レベル 3 相当 開発するソフトウェアの担当部分だけではなく、担当部分を含んだ全体の機能、構造、特徴、特性を理解している。 製品全体の機能に基づき、欠陥を指摘できる。要求されている品質（機能性、信頼性等）の目標に基づき問題を指摘できる。	EIDeR-Training 法	担当部分を含んだ製品全体
レベル 4	レビューの責任者	開発リーダー・テストチームリーダー ITSS レベル 4 以上相当 開発するソフトウェアの欠陥だけではなく、あるべき姿から見た場合の製品の不足点、問題点を指摘できる。	OJT	担当製品だけではなく、製品を利用する業界全体の知識

### 3. 提案手法の検証

#### 3.1 実験の目的

我々は、レビューアのレベルに合わせて作成した教材を使った EIDeR-Training 法の効果・有用性を検証するため、2 つのレビュー実験を実施した。効果を確認した主な項目は、以下 2 点である。実験 1 では、(1) の効果を確認するために行う。効果・有用性は、実験前後のシナリオに基づく欠陥指摘件数から測定し、RQ1 を確認する。実験 2 では、(2) の効果を確認するために行う。効果・有用性は、各レベルの教材を用意し、各レベルの協力者に全てのレベルの教材を解いてもらい、各教材の不具合指摘率によってレビューアのレベルを測定できないか実験し、RQ2 を確認する。RQ3 は、実験 1, 2 から確認する。

- (1) 若手レビューアのシナリオに基づく指摘件数は増えたか？
- (2) レビューアのレベル定義は妥当か？

#### 3.2 実験方法

本研究では、表 3 のレベル 1～3 の該当者に対して、表 4 と表 5 に示す手順で実験を行った。実験手順の要点は、以下の通りである。

準備-1 では、過去の欠陥情報をシナリオレビューのシナリオの形にまとめる。準備-2

では、表 6 の観点に基づき、シナリオを分類した。

実験 1 は、トレーニングによってレビューアが成長したかどうかを確認するため、2 種類の教材を使い、1 回目と 2 回目の結果を比較する。

実験 2 は、各レベルの教材を作成し協力者全員にレビューしてもらった。各教材の不具合指摘率によってレビューアのレベルを測定できないか実験する。

表 4：実験 1 手順表

	実施内容
準備-1	欠陥情報まとめ (シナリオ)
準備-2	シナリオの観点分け (表 6)
準備-3	レベル毎の教材作成 (2 種類)
準備-4	実験協力者のレベル分け (表 3)
実験-1	トレーニング実施 (1 回目)
実験-2	トレーニング実施 (2 回目)

表 6：シナリオの観点

レベル	観点
1	・設計手法や開発言語に属した一般的な欠陥 ・開発するソフトウェア固有といえない欠陥
2	・ドメイン知識を必要とする欠陥や製品固有の欠陥 (単一機能部や特定機能など影響範囲や観点の限られた範囲の欠陥)
3	・ドメイン知識を必要とする欠陥や製品固有の欠陥 (複数機能部や製品全体など影響範囲や観定の範囲が広い欠陥)

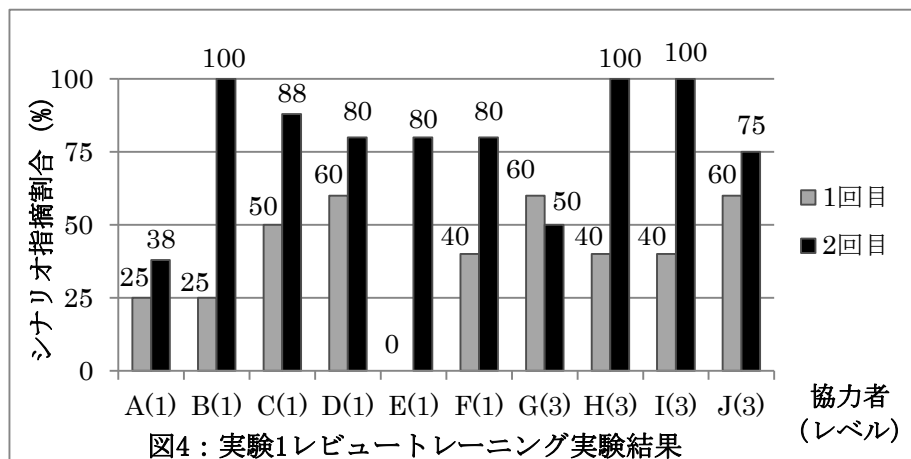
表 5：実験 2 手順表

	実施内容
準備-1	欠陥情報まとめ (シナリオ)
準備-2	シナリオの観点分け (表 6)
準備-3	レベル毎の教材作成 (3 種類)
準備-4	実験協力者のレベル分け (表 3)
実験-1	レベル 1 教材によるレベル測定
実験-2	レベル 2 教材によるレベル測定
実験-3	レベル 3 教材によるレベル測定

### 3.3 実験結果

レビュー実験の結果を以下に示す。実験 1 のトレーニング前後の欠陥指摘割合を図 4 に示す。縦軸は、レビューで指摘できたシナリオ件数の割合を表し、横軸は、実験の協力者で括弧内の数字は定義したレビューアのレベルを表している。若手レビューア、レベル 1 レビューア、レベル 3 レビューアに着目する。

- ・若手レビューア：若手に該当する協力者は、B, C, D, E, F, H, I さんであり、全員トレーニング前後の指摘が増えた。とりわけ B, H, I さんは、全てのシナリオを指摘した。
- ・レベル 1 レビューア：レベル 1 の協力者は A-F さんであり、全員トレーニング前後の指摘が増えた。
- ・レベル 3 レビューア：レベル 3 の協力者は G-J さんで、若手の方が伸びていることがわ



かる。また、Gさんはシナリオに記載された指摘が減少した。

実験2の結果を表7に示す。レベル2とレベル3を使った教材に関しては、実験前の想定レビューレベルに近い実験結果が出た。しかし、レビューアCに関してはレベル2とレベル3の結果が逆転しているという想定と異なる結果が出た。また、本来であれば最も高得点となるレベル1の指摘割合が全般的に低い点も想定と異なっていた。

表7：実験2教材によるレビューレベル測定実験結果

レビューア	シナリオ指摘割合(%)	教材レビューレベル		
		レベル1	レベル2	レベル3
A	実験前の想定レビューレベル 1	33%	83%	0%
B	1	67%	50%	0%
C	1~2	67%	33%	40%
D	2	50%	67%	20%
E	2~3	67%	83%	20%
F	3~4	100%	100%	100%

### 3.4 考察

本研究では、レビューアのレベルを定義し、レベルに合った知識をトレーニングすることで、効率よくレビューアを育成できないか検討した。以下に1章で挙げたRQ1-3について考察する。

*RQ1. 若手レビューアを（短時間で、かつ、実害のある失敗はさせずに）レビューを任せられる人材に成長させるにはどのようにしたらよいか*

実験1の結果より、シナリオによりレビューアに役割を持たせたEIDeR-Training法を活用することで知識を身に付けられると言える。実務において若手レビューアにレビューを任せる場合は、若手レビューアのドメイン知識量は少ないため、1回のレビューを集中して行える量のシナリオを渡し、レビューを複数回行い、最終的に全てのシナリオを確認するレビューのやり方をするとよいと考える。

*RQ2. レビューアのレベルを定義し測定することは可能か*

実験2の結果より、レベル別の教材を使うことで、各レビューアのレベルを測定できた。レビューアDとEは想定通りレベル2の実力があり、今後ドメイン知識を身につけることでレベル3になれることを示している。レビューアCは、プロジェクトに早期に投入し、OJTをメインにしたためドメイン知識はレビューアD、Eよりも高いが、基礎知識をきちんと身につけていない部分がありレベル2とレベル3のシナリオ指摘割合が逆転した結果となったと考える。

実験2の結果で述べた想定よりもレベル1の指摘割合が低い結果について述べる。実験2の教材は、レベル1の教材のページ数が10ページ、レベル2と3が6ページとなっている。想定よりもレベル1の指摘割合が低いことから教材の作り方、特に教材の分量を調整する必要があると考える。教材の分量によっては難易度が想定したレベルにならない。このため、レベル測定用の教材は分量を少なめにし、埋め込んだ不具合の難易度がそのまま教材の難易度になるように調整する必要がある。

また、レビューアによっては教育の過程で知識に偏りがあり、実験の結果が想定通りにならなかったケースがあった。これは、レビューアの弱点を洗い出せたと捉えることができる。弱点部分のトレーニングをEIDeR-Training法で補うことで、偏りを是正できると考

える。このことから、教材を使用したレベル測定を行う際にはレビューアの想定レベルの教材だけを使用するのではなく、レベル 1~3 の教材を一通り実施したほうが良いと考える。

**RQ3. レベル別に観点を絞った教材を用意することで、トレーニングの効果をあげられないか**

はじめに、レベル別の教材という点に着目する。表 3 で定義したレベルに対して、表 6 の観点に基づき教材を作成した。図 4 から、2 回目のトレーニングの方が指摘率は増加した協力者が多いことがわかる。ただし、増加の要因がシナリオの使い方がわかったことによるのか、1 回目のトレーニングでドメインを理解したことによるのかが判断できなかった。レベル定義の妥当性を測定するには、シナリオの使い方による指摘ではなく、ドメイン知識を理解したことによる測定が必要だと考える。

次に、観点を絞った教材という点に着目する。図 4 の C さんは、昨年度の研究においても実験に協力してくれた方である。実験後に C さんから、今回はやるべき項目を絞って指摘できたというコメントを頂いた。

コメントからも、観点を絞った教材でトレーニングすることは効率であると考ええる。また、観点を渡さずにトレーニングを実施することでトレーニングの定着度やトレーニーの実力を測ることができるのではないかと考える。

以上から、レベル別に観点を絞った教材を用意することで、トレーニングの効果をあげることができると思う。

#### 4. おわりに

本研究では、昨年度の研究で考案した EIDeR-Training 法にシナリオレビューのシナリオを組み合わせることで、レビューアのレベル毎に効率的なトレーニングが可能であることを示せた。欠陥知識、ドメイン知識をシナリオに蓄積しておけば、EIDeR-Training 法の教材をすぐ作れるだけでなく、シナリオをそのままレビューにも活かすことができ非常に有効である。

しかし、本研究ではシナリオのまとめ方や作り方を手法としては確立できていない。今後シナリオレビューの研究が進み、シナリオの効率的な作成方法が確立することでシナリオレビューの普及が進むことを期待したい。また、シナリオレビュー自体のトレーニングに本研究で提案した EIDeR-Training 法は有効であり、本研究がシナリオレビューの普及の一助になれば幸いである。

#### 5. 参考文献

- [1] 細川宣啓, 永田敦, 森崎修司, 西村英俊, 小田部健, 添田健太郎, 中谷一樹「レビューオリエンテーションキットを用いた育成によるレビュー文化の醸成」, 日本科学技術連盟 SQiP 研究会, 2011
- [2] 細川宣啓, 永田敦, 藤原雅明, 森崎修司, 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」, 日本科学技術連盟 SQiP 研究会, 2012
- [3] 竹下千晶, 山路厚「技術者のスキル向上につながるピア・レビューに進化させる仕組み」, (株)デンソークリエイティブ プロジェクトセンター現場改善推進室, SQiP シンポジウム 2012
- [4] 中谷一樹, 原佑貴子, 上田裕之, 森崎修司, 原山拓実, 藤田延介, 山田将貴「レビューアのドメイン知識を飛躍的に向上させるトレーニングの提案~実成果物の利用により実践的なレビュースキルを向上~」, 日本科学技術連盟 SQiP 研究会, 2014
- [5] 森崎修司 著, 「間違いだらけの設計レビュー」, 日経 SYSTEMS, 2014