

AIナビゲーション型フリースタイルテストの提案

- テスト技術不足を補う, 仕様の詳細化不十分に起因するバグの検出手法 -

有森 勇一（株式会社両備システムズ）

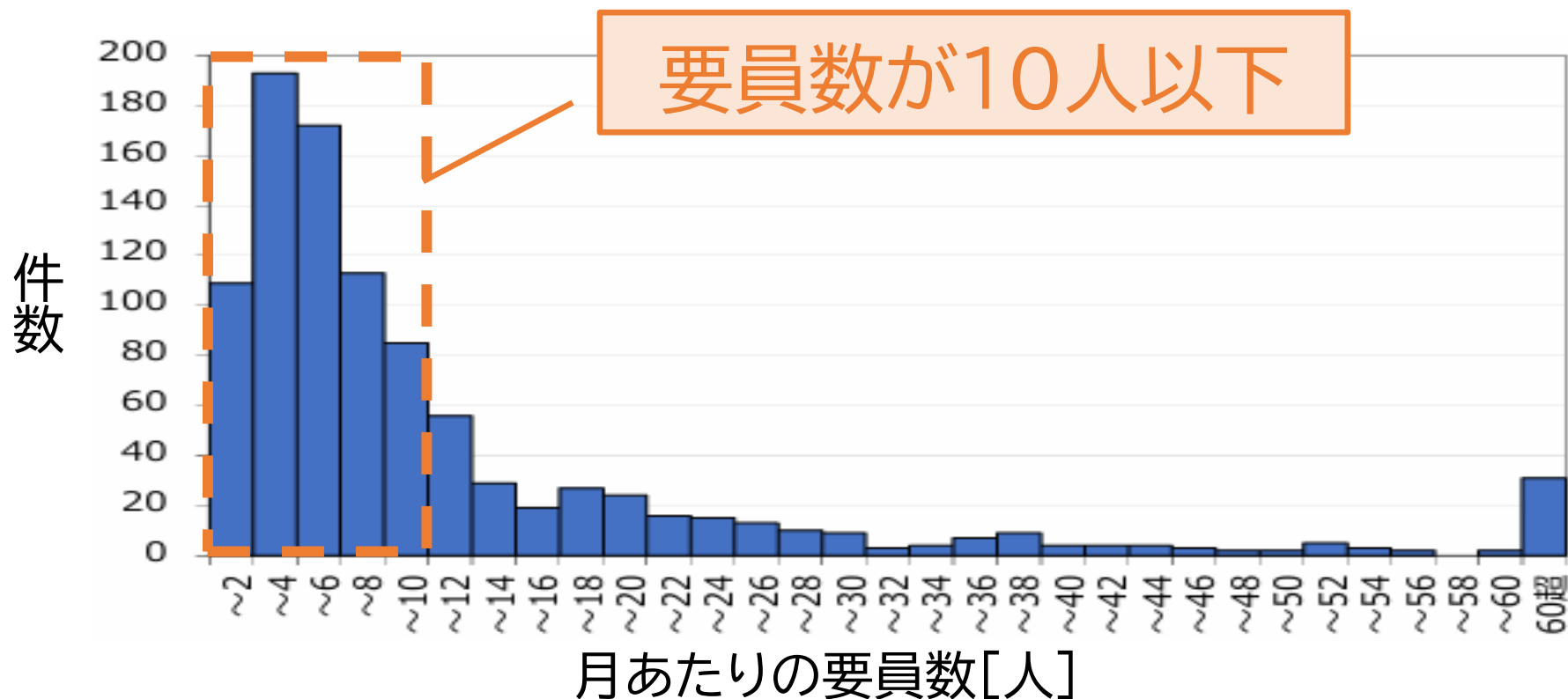
津梅 眞嗣（株式会社ゼネラル）

本日の内容

- 背景・問題
- 課題
- 提案
- 実験検証
- 実システム検証
- まとめ

ソフトウェア開発では、10人以下の小規模体制が多い

一般的な開発の要員数



出典：ソフトウェア開発分析データ集 2022 プロジェクトの月あたりの要員数

小規模体制の役割分担

少人数の体制では、1人が設計・開発・テストを兼任しやすい

限られたメンバーで開発すると...



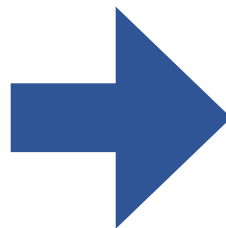
問題① ノウハウの属人化

有識者が開発に追われ、新人育成と知見共有が進まない

開発業務が忙しすぎると…



有識者が手一杯



新人の教育が不十分

組織内に **スキルのばらつき** が生じる

問題② 設計者視点の思い込み

設計時の思い込みを引き継ぎ、仕様バグを見逃しやすい

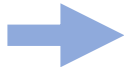
設計者がテストを兼任することの問題点



仕様書に、
・ 例外条件
・ 前提条件
漏れ/曖昧あり

(以降、仕様バグと呼ぶ)

要件定義・
設計



開発



テスト



考慮漏れ/勘違いで
仕様バグ見逃し



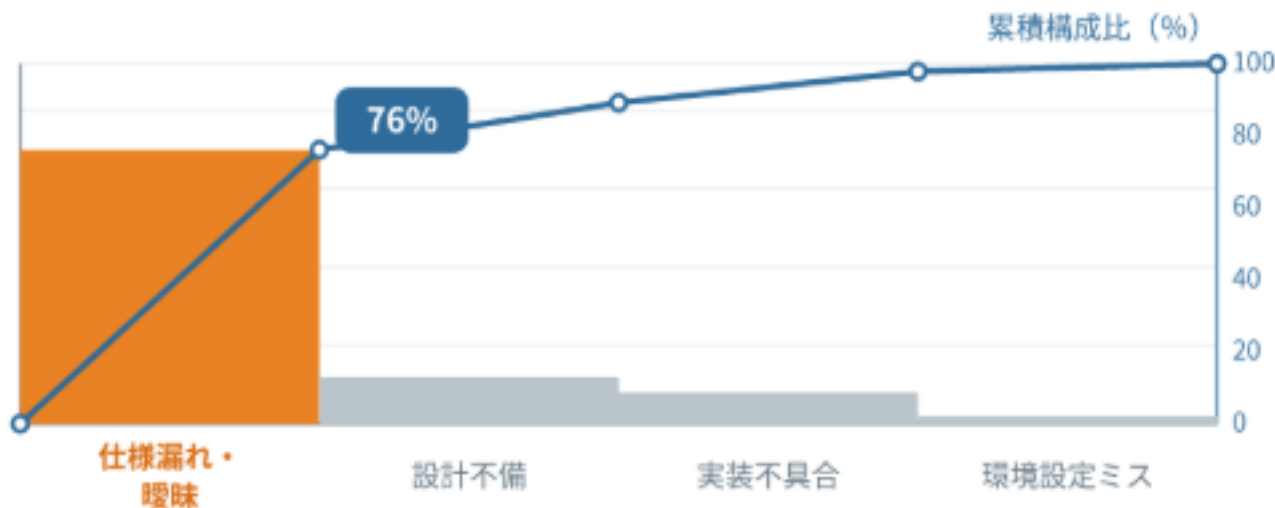
設計が正しいと **思い込んだまま**
テストを設計

リリース後に発見されたバグの76%は、仕様漏れ・曖昧さに起因していた

リリース後に発見されたバグの原因

76%

仕様漏れ・曖昧さが原因



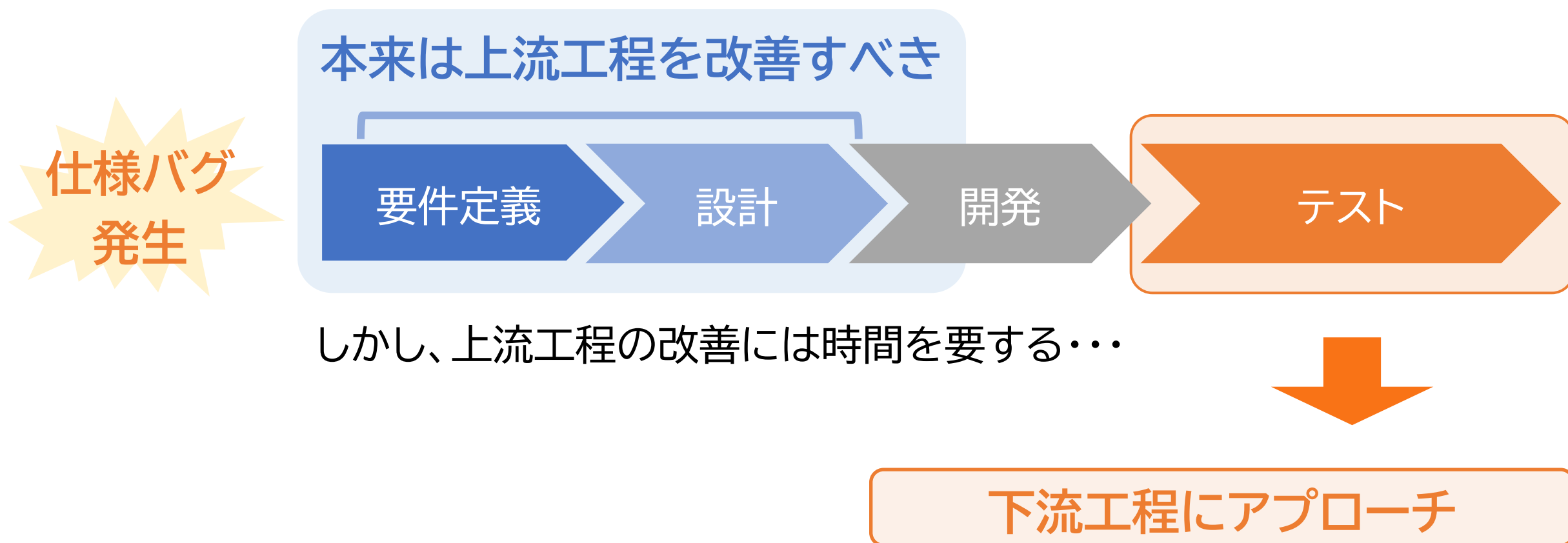
問題構造

- 1 ノウハウの属人化
- 2 設計者視点の思い込み



仕様バグの見逃し

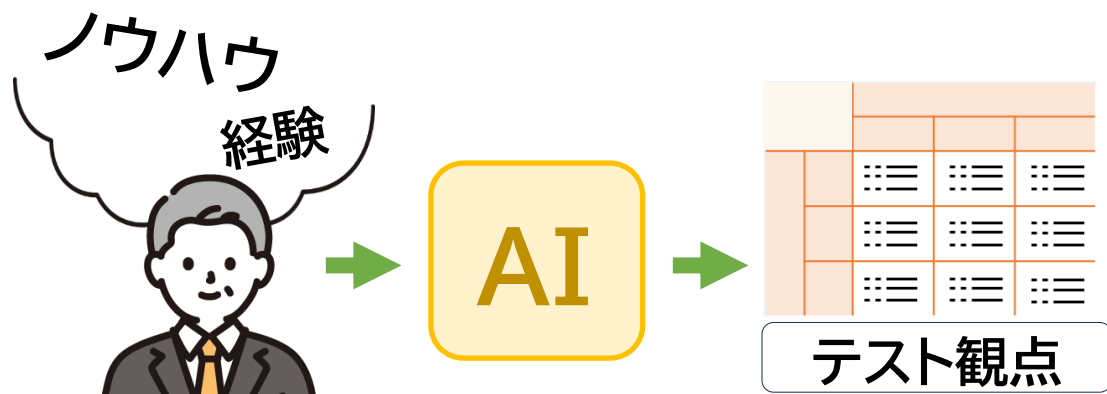
上流改善を進めながら、即効対策としてテスト強化による品質向上を図る



先行研究の支援範囲

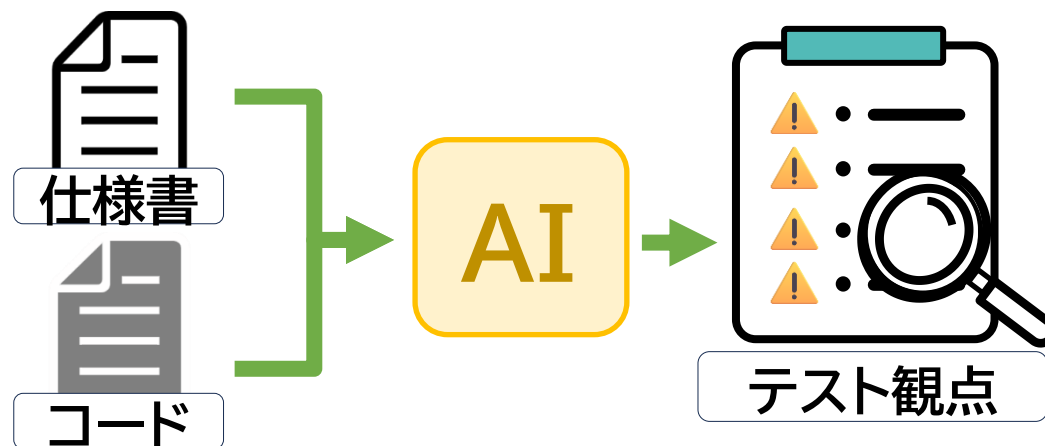
先行研究は、テスト実行前の観点提示を中心に支援している

属人化に対する先行研究



AIによるテスト観点抽出ノウハウの可視化

仕様バグに対する先行研究



LLMによる重要テスト観点の抽出

テスト観点抽出 に着目

課題 1 : 経験に依存せず仕様バグを検出する

課題 2 : テスト実行中の具体的な操作を支援する

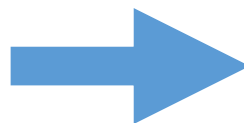
先行研究に残る課題

テスト観点を提示



【テスト観点】

- ・ テキスト入力欄のセキュリティを確認



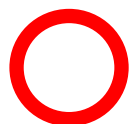
何を、どう操作する？



テスト実行者

テスト技術不足

ドメイン知識不足



観点はわかるが...



操作手順がわからない

既存のテストに、仕様バグを狙って検出するテストをプラスする

提案手法の適用イメージ



仕様書

要件定義

設計

実装

テスト

仕様バグ検出テスト

仕様バグ検出のためのテスト実行を追加。
そのテストの **観点・操作はAIが提示。**



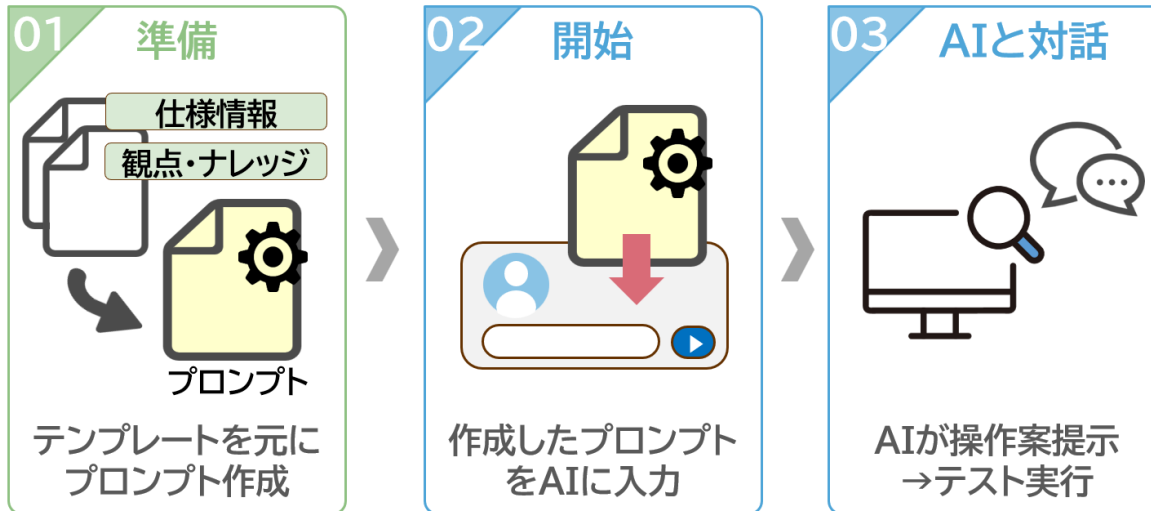
観点・操作提示

テスト実行

概要

- 生成AIがテスト実行を支援（観点・操作例を提示）
→ テスターは提案を選び、判断しながら試す！

実施プロセス



※実験では、2025年12月時点のMicrosoft Copilot 無料版(Think Deeperモード)を使用

AI対話イメージ

テスト対象の機能を選択してください。
1. 受注登録
2. 出荷処理

1

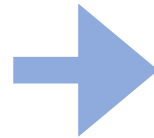
以下を参考にテストを実施してください。

観点	操作例	期待される挙動
...	...	...
...	...	...

3階層のプロンプト設計により、テスト対象に合わせた提案を提示

単純なプロンプト指示では…

- ① AI提案の内容が理解しづらい
- ② テスト対象の仕様に合わない
観点・操作例が提示される
- ③ 重要度の低い観点が提示される



テストナビプロンプト

① **AIの役割・進行ルール**
検出目的・提案形式・対話の進め方

固定

② **テスト対象の仕様情報**
画面・機能・制約・業務ルール

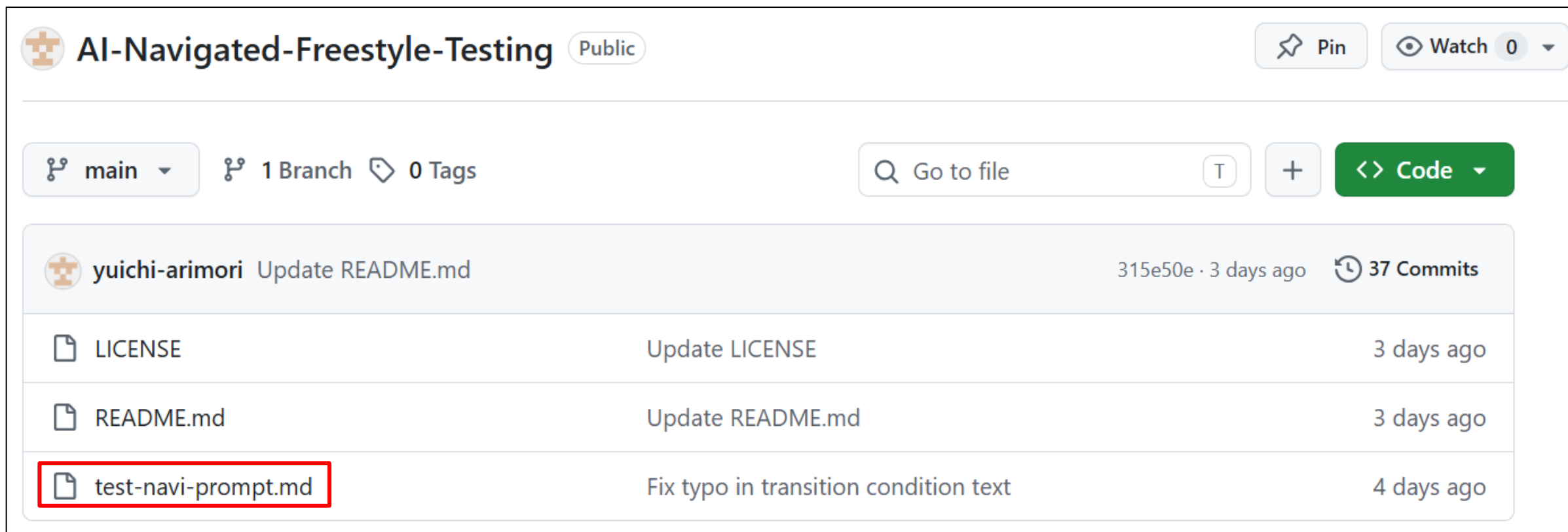
対象ごとに
差し替え

③ **観点・ナレッジ**
一般的な観点・過去バグ情報

テストナビプロンプトの公開先

テストナビプロンプトは、GitHubで公開しています。

<https://github.com/yuichi-arimori/AI-Navigated-Freestyle-Testing>



The screenshot shows the GitHub repository page for "AI-Navigated-Freestyle-Testing" by user "yuichi-arimori". The repository is public. The interface includes a header with the repository name, a "Public" badge, and buttons for "Pin" and "Watch". Below the header, there are buttons for "main" branch, "1 Branch", and "0 Tags". A search bar "Go to file" and a "Code" button are also present. The commit history table shows three commits: "Update README.md" (3 days ago), "Update LICENSE" (3 days ago), and "Fix typo in transition condition text" (4 days ago). The file "test-navi-prompt.md" is highlighted with a red box.

File	Commit Message	Time
LICENSE	Update LICENSE	3 days ago
README.md	Update README.md	3 days ago
test-navi-prompt.md	Fix typo in transition condition text	4 days ago

研究課題 1 : 経験に依存せず、仕様バグを検出できるか

研究課題 2 : AIが示す操作例や入力例は、テストに実用的か

検証方法

実験参加者(23名)



新人 9名
中堅以上 14名

1 自由にテスト



60分

2 提案手法でテスト



60分

3 アンケート



実用性を評価

今回の研究では、疑似SNSアプリを題材に検証を実施した

テスト対象

FindingBugs(https://jam0824.github.io/finding_bugs/)



仕様バグ検出件数



自由にテストで検出



提案手法で検出



新人
9名

平均検出数(件/人)

2.7件

4.2件

累計

6.9件



中堅以上
14名

平均検出数(件/人)

3.8件

2.1件

累計

5.9件

提案手法の併用により 新人の累計検出数は中堅以上と同程度 になった

発見例1 絵文字を1文字入力すると、2文字分カウントされる

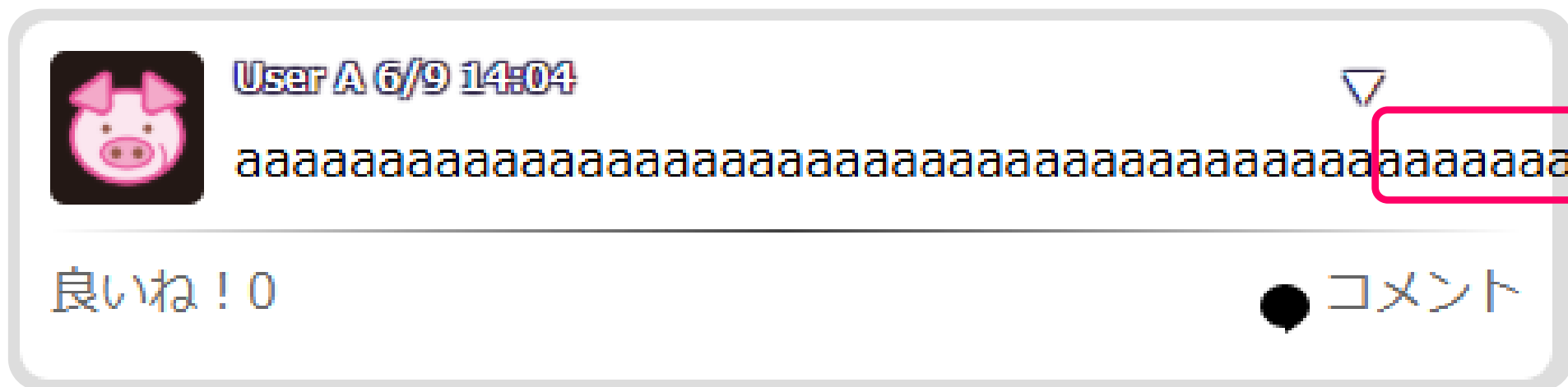
1 投稿入力画面(入力前)

2 投稿入力画面(入力後)

「140文字投稿可能」の仕様違反

発見例2 半角英字を140文字入力すると、折り返されず見切れる

1 投稿表示画面



表示制御の考慮漏れ

Q. AIから提案された操作手順は、すぐにテスト実行できましたか？

肯定
57%

中立
30%

否定
13%

はい
13%

ややはい
44%

どちらでもない
30%

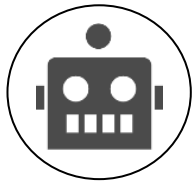
ややいいえ
13%

AI提案には **一定の実用性** がある一方、**改善の余地** も確認された

アンケートから、以下のような改善点も見えてきた

改善の余地があるAI提案例

例1: 実施できない操作手順を提案する

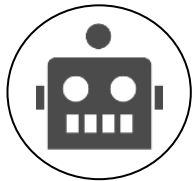


2つのボタンを同時にクリックしてね

どうやって！？



例2: テスト観点に優先順位を提示しない



テスト観点①、②、③…

時間ないから大事なテストだけ教えて！



実システムでの検証

環境

項目	実システム適用	(参考) 実験検証
種類	経路検索・予約Webシステム	疑似SNSアプリ
開発規模	約140KLOC	約0.9KLOC
機能数	30機能	6機能
システム構成	外部連携、DB等を含む	単一のHTML、DBなし

テスター

対象システムの開発経験なし

検証方法

AIが示した操作例のみをもとにテストを実施

結果

仕様バグを8件検出

【課題1】 経験に依存せず仕様バグ検出できるか

▶ **YES** (テスターのノウハウ不要)

【課題2】 AIが示す入力・操作例が実行できる

▶ **YES** (仕様の知識不要)

実システムでも **研究課題への有用性を確認** できた

実システムでのバグ発見例 — 特殊文字

発見例 1 領収書画面で、特殊文字が表示されない

1 会員情報の設定画面

会員情報変更

お名前

山田 吉太

フリガナ

ヤマダ ヨシタ



2 領収書(プレビュー)

お客様名 山田 太様

発行日 202X年04月01日
領収書番号 XX2026040100000000

発行会社 XXXX株式会社
登録番号 T12345678

住所 岡山市XXXX1-1-1
XXビル 5F

TEL 0120-123-456
FAX 0120-123-457

領 収 書

会員情報の名前を「**吉**（吉の異体字）」で登録

領収書では、「**吉**」が表示されない

AIが **特殊文字の入力** を提案したことで、発見することができた

実システムでのバグ発見例 — 最大値

発見例 2 クレジットカード登録画面で、カード情報の保存に失敗する

1 会員情報の設定画面

会員情報変更

お名前

BounthaviXayChanthal VongsanakhonPhommasa

フリガナ

ブンタヴィサイチャントル ヴォンサナコンボンマサ

姓・名を最大値の20文字で登録（架空の外国人名）

2 クレジットカード情報の設定画面

クレジットカードの登録

カード名義

BounthaviXayChanthal VongsanakhonPhommasa

カード番号

0000 0000 0000 0000

有効期限(月/年)

12 / 30

セキュリティコード

123

カード情報の保存に失敗しました。暫く経ってから再度お試しください。

カード情報の保存に失敗する

AIが **最大値での入力** を提案したことで、発見することができた

提案手法の有効性と実用性

背景

小規模体制に伴う、
複数工程の兼任と新人教育時間不足

問題

ノウハウの属人化と思い込みによる
仕様の漏れ・曖昧さの見逃し

提案

AIがテストの操作・入力例を指示し、
テスターに新たな視点を与える

研究課題1

経験に依存せず、
仕様バグを検出できるか



確認できたこと

経験の浅いテスターの
仕様バグ検出を支援

研究課題2

AIが示す操作例や入力例は、
テストに実用的か



確認できたこと

一定の実用性を確認

- 1 現場への適用を広げる**
 - 異なる業務・システムへ適用し、有効性と適用条件を検証
- 2 プロンプトを改善し、AIの提案をより実用的にする**
 - 実行可能性や優先度を踏まえた観点・操作例・入力例を提示
- 3 知見を上流工程へ活用し、品質改善サイクルを構築する**
 - 検出した仕様の漏れや曖昧さを、要件定義・設計へ反映
 - 有効なテスト観点をナレッジとして蓄積し、次のテストに活用

謝辞

本研究は、2025年度SQiP研究会「ソフトウェアテストコース」での議論を通じて、検討し、洗練してきたものです。

研究活動において、以下の方々にご指導を賜りました。

- 喜多義弘 主査
- 秋山浩一 副主査
- 西田尚弘 アドバイザー

また、早坂淳氏、丹羽優斗氏と共に研究に取り組みました。

実験には、松谷峰生氏が開発・公開されたテスト練習用Webアプリケーション「FindingBugs」を利用させていただきました。

皆さまに深く御礼申し上げます。

ご清聴ありがとうございました。