

Towards Reliable Agile Iterative Planning via Predicting Documentation Changes of Work Items

Jirat Pasuksmit

jpasuksmit@student.unimelb.edu.au

The University of Melbourne

Australia

Patanamon Thongtanunam

patanamon.t@unimelb.edu.au

The University of Melbourne

Australia

Shanika Karunasekera

karus@unimelb.edu.au

The University of Melbourne

Australia

19TH International Conference On Mining Software Repositories (MSR 2022)

発表者：井町 瑠貴

背景 | アジャイル反復開発

アジャイル反復開発

- 短期間の反復でソフトウェアの機能を実装・提供するソフトウェア開発プロセス
- 一般的に「スプリント」と呼ばれる時間枠ごとに作業を行う

アジャイル反復開発の工数見積もりでは、文書を使用する

- 最近の調査研究では、開発者の97%が文書を使用し、文書化された情報は重要と考えていることが報告されている^[38]

[38] Jirat Pasuksmit, Patanamon Thongtanunam, and Shanika Karunasekera. 2021. Towards Just-Enough Documentation for Agile Effort Estimation: What Information Should Be Documented? Proc. of the ICSME (2021), 114-125.

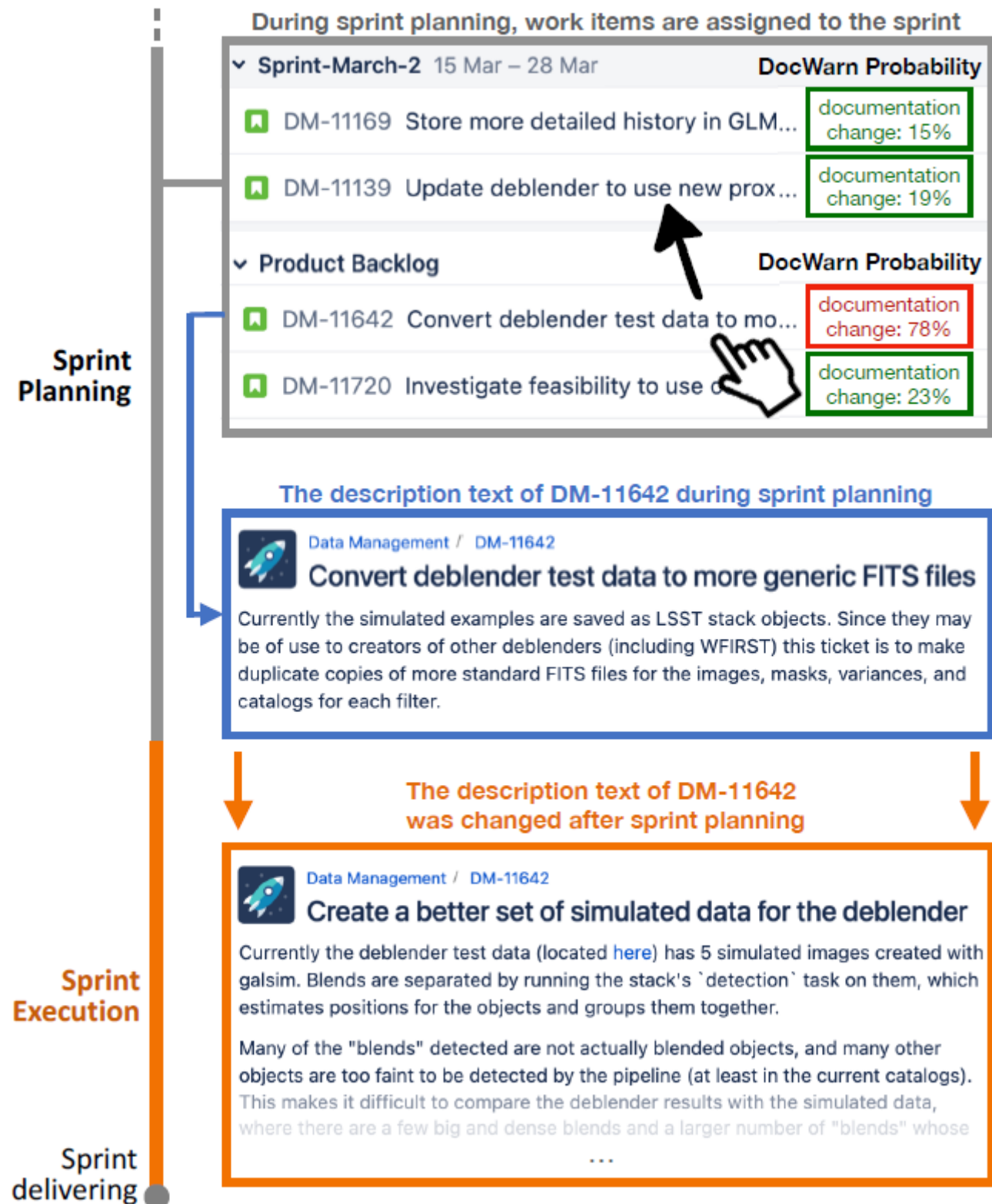
背景 | 動機となる例

DM-11642

- スプリントに割り当てられた後変更された
- 見積もり工数が1から10に増え、納期遅延の原因となった

DocWarn（本研究で開発）

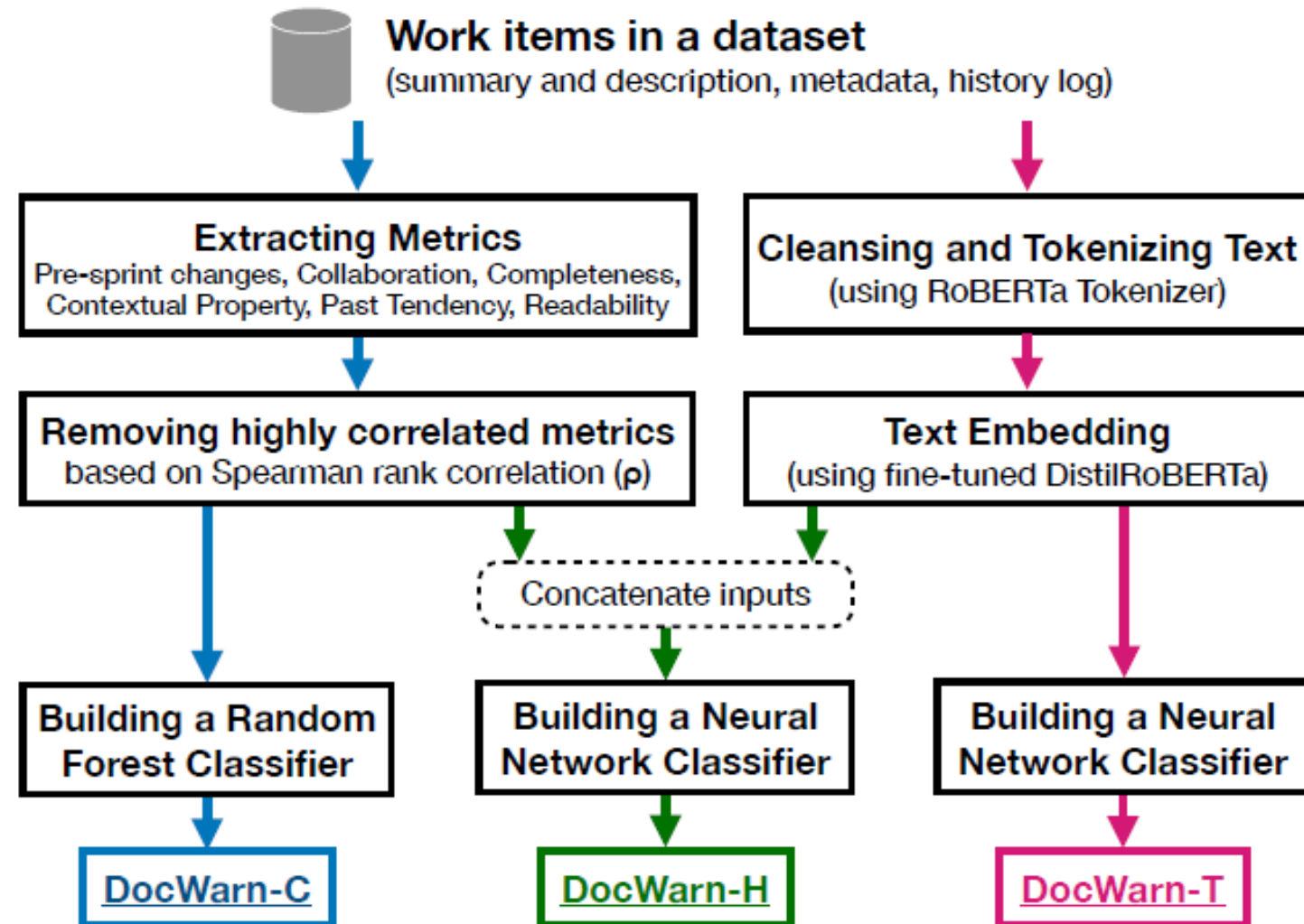
- 各作業項目の情報が変更される確率を提供
- チームは潜在的な不確実性を認識することができる



DocWarn

本研究では以下の3つのDocWarnを開発

- DocWarn-C : 作業項目の特性に注目
- DocWarn-T : 作業項目の説明文に注目
- DocWarn-H : CとTで用いたベクトルを連結



リサーチクエスチョン

RQ1：ドキュメントの変化をどの程度予測できるのか

- DocWarnの精度を評価するために
2つのベースラインと比較

RQ2：ドキュメント変更の判断に最も影響する
特性はなにか

- ドキュメントの変更が起こりやすい作業項目の
特性を調査

RQ3：正しく予測できるドキュメント変更の理由は
なにか

- DocWarnが予測できるドキュメントの変更について、
定性的な評価を実施

調査対象プロジェクト

Jira

- アジャイルチームに最適な機能を提供する
開発者向けプロジェクト管理ツール

以下の2つの基準でプロジェクトを選択

- スプリントでソフトウェア開発している
- 積極的にドキュメントを整備している
 - 作業項目に関する長い説明文を文書化
 - 作業項目を頻繁に更新する傾向がある

選択の結果, 9つのプロジェクトを
調査対象とする

文書の変更

本研究での作業項目の文書変更

- 作業項目がスプリントに割り当てられた後、説明文が意味的に10%以上変更されること

説明文の意味的な違いの測定

- コサイン類似度を使用
- スプリント割り当て時と最終版の意味的な差異を測定

RQ1 | アプローチ

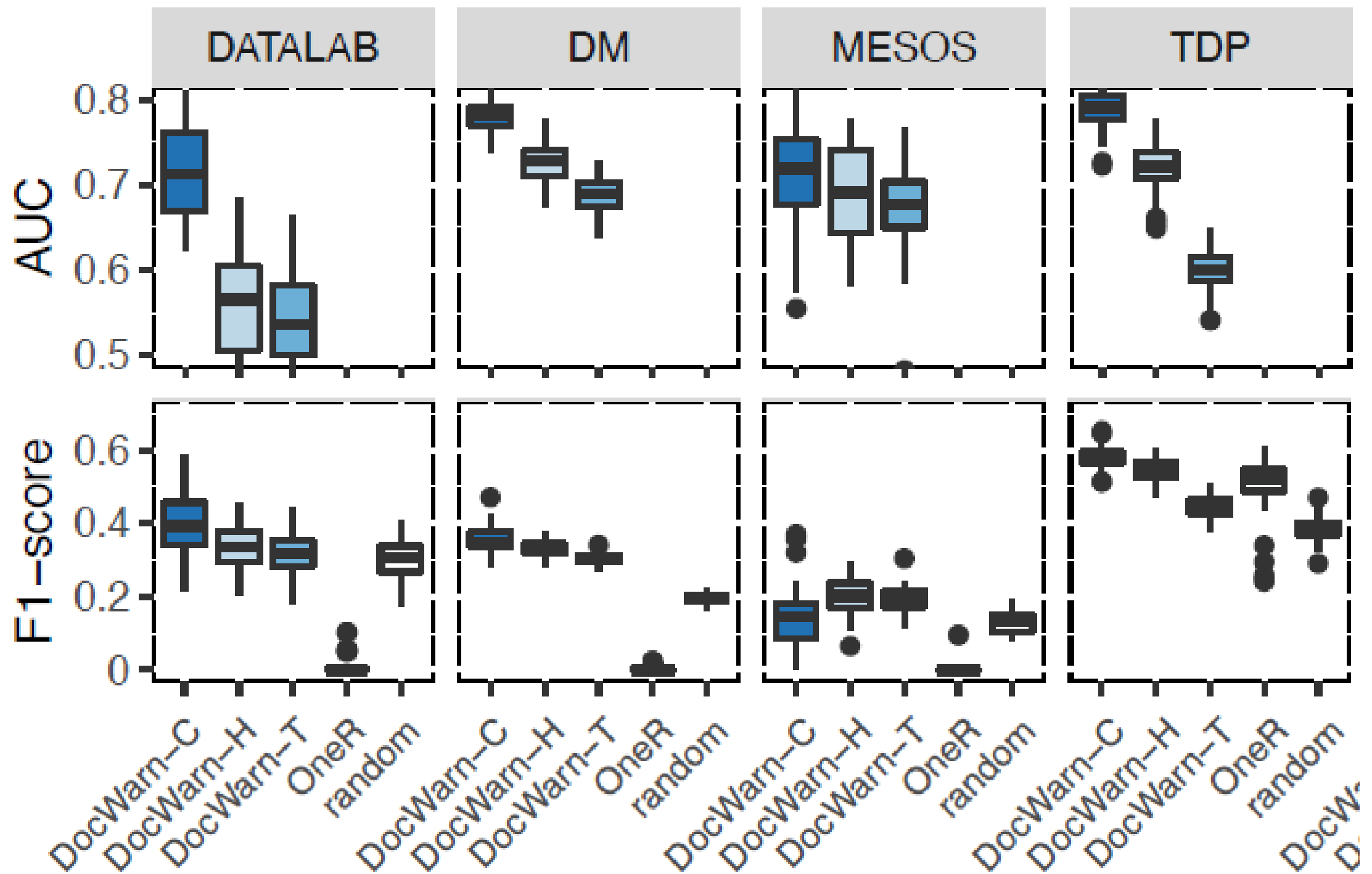
2つのベースライン（ランダム推測とOneR）と DocWarnのパフォーマンスを比較

評価指標

- AUCとF1スコアを利用

評価指標5分割交差検証を10回繰り返し、50個のモデルの評価結果を作成

RQ1 | 結果



RQ1 | 結果

- | DocWarnが2つのベースラインよりも優れた性能を発揮する傾向がある
- | DocWarnの中では、DocWarn-Cが最も優れた結果を示している
 - 作業項目の特性が、説明文より文書変更の予測因子として優れていることを示唆

RQ2 | アプローチ

DocWarn-Cにおける各メトリクスの影響力を平均減少精度（MDA）に基づいて測定

- RQ1で作成した50個の交差検証モデルすべてについて MDA を推定

各メトリクスの影響度をランク付け

- Scott-Knott ESD 検定を実施

スピアマン順位相関を用いて各メトリクスと確率の関係を測定

- 確率の間の関係の方向性を調査

RQ2 | 結果

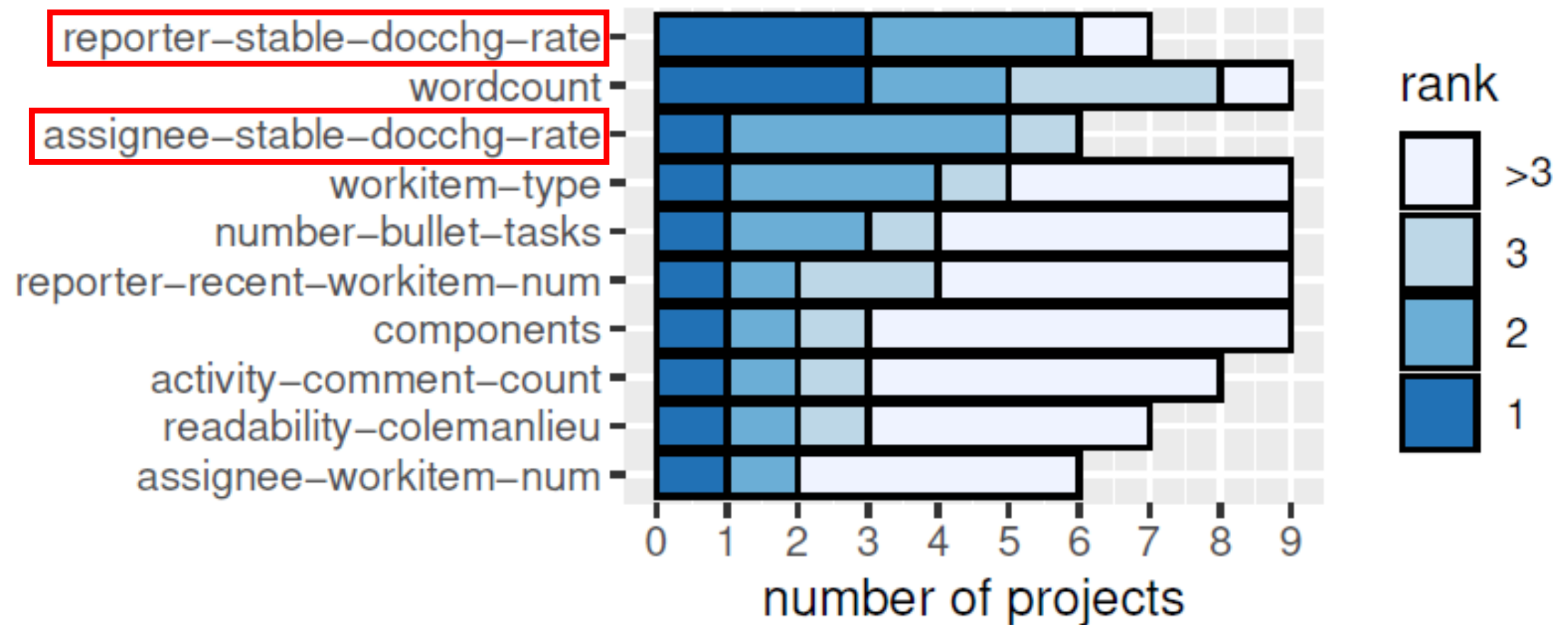


Figure 5: The ranks of the 10 most influential characteristics (metrics) in DocWARN-C.

RQ2 | 結果

| RSDRやASDRにおいて、文書変更が生じる確率に逆相関がみられた

- 不安定な項目の担当経験のある開発者が報告者や担当者になった場合、不安定になる傾向がある

| 説明文の長さにおいて、文書変更が生じる確率に逆相関がみられた

- 説明文が短いほど、作業項目に文書変更が発生する確率が高い
- チームは、説明文が短い作業項目を再検討し、十分な情報が含まれていることを確認する必要がある

RQ3 | アプローチ

| DocWarnが正しく予測できる作業項目の
ドキュメント変更の理由を手動で調査

- DM, TDQ, TUPのサンプルに対してオープン
コーディングを実施

| オープンコーディングで得られた理由に
基づいて文書変更の理由を分類

- 9つのプロジェクトから合計394の作業項目を分類

RQ3 | 結果

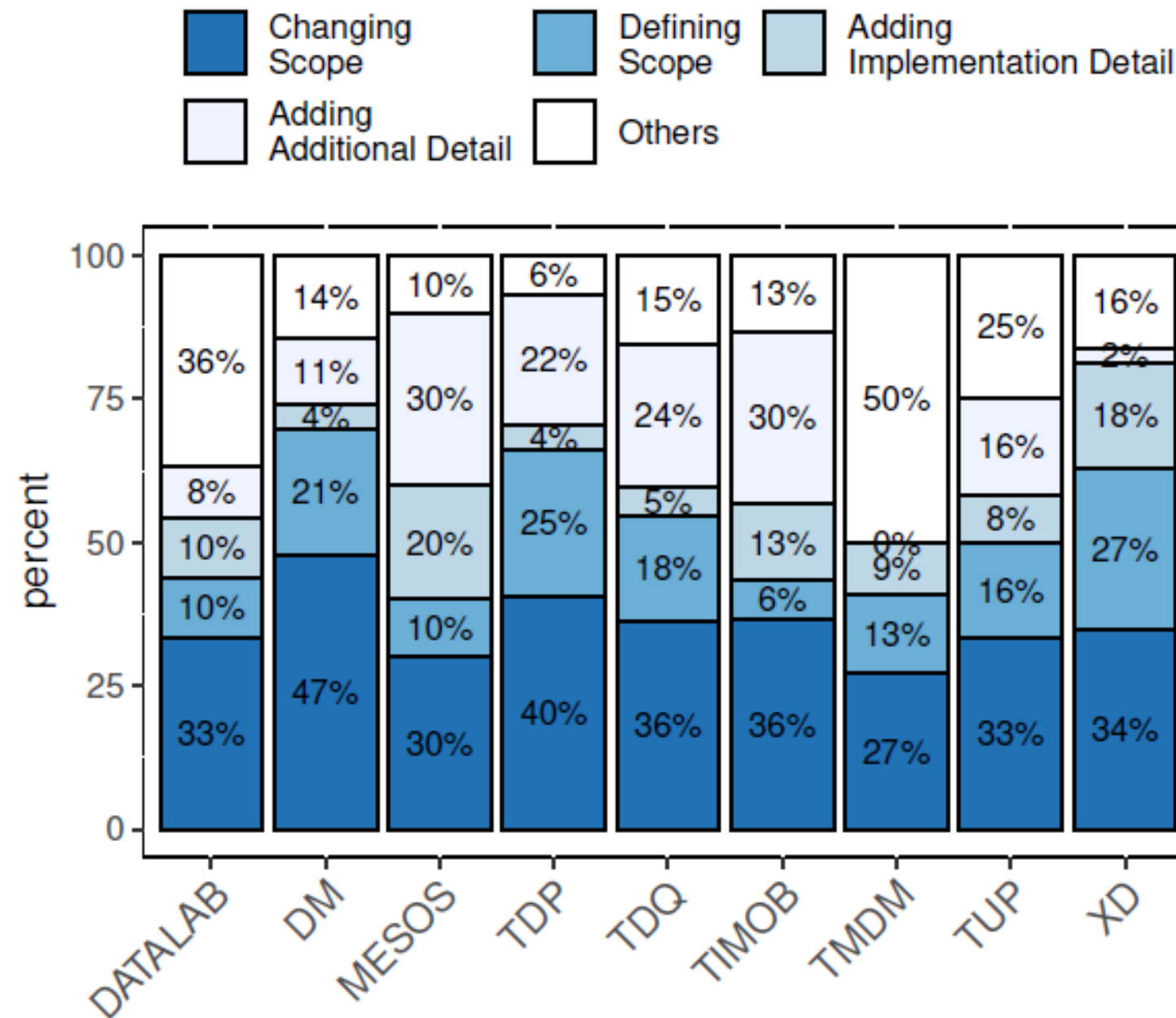


Figure 6: The reasons for the documentation changes that DocWARN-C can predict correctly.

RQ3 | 結果

スコープ変更とスコープ定義という、スコープの変更に関する文書の変更が高い割合となった

- スコープ変更

- 作業項目がスプリントに割り当てられた後、作業範囲を変更するために文書を変更すること

- スコープ定義

- スプリント割り当て時に作業範囲が定義されていない場合の文書変更のこと

結論

- | アジャイルチームが不確実性をよりよく管理するために、スプリントに割り当てられた後の作業項目の文書変更を予測するDocWarnを提案
- | DocWarnはベースラインより大幅に改善され、スコアの修正に関連するものを正しく予測できる
- | 今後の研究では、より細かい文書に対する文書変更を予測し、変更の影響を推定できるアプローチを検討する必要がある

所感

- | 自分を取り扱いたいドキュメントとは違う種類のドキュメントを取り扱っており、興味深く感じた
- | 積極的な文書化を行わないプロジェクトではどのような結果になるのか少し気になった

補足資料

補足 | DocWarn-Cで用いた特性

Table 1: A list of metrics that we used to measure the characteristics of a work item at sprint assignment time.

Dimension	Metric	Description
Pre-Sprint Changes	textdesc-change-count, assignee-change-count, comment-count	The number of change activities (i.e., changing text, changing assignee, comments) occurred <i>before</i> the work item was assigned to a sprint.
Collaboration	in-degree, out-degree, total-degree, kcoreness, clustering-coefficient, closeness-centrality, betweenness-centrality, eigenvector-centrality	Degree of activity that the reporter of work item recently had with other developers in his/her collaboration network in the project.
Completeness	has-acceptance-criteria, has-code, has-attachments, has-testcases, has-stacktraces, number-bullet-tasks, number-hyperlinks	Whether the description text contained the technical information.
Primitive Attribute	workitem-type, priority, components, number-watchers, number-votes, has-subtasks, has-epic	The attributes of a work item that are directly extracted from JIRA metadata.
Past Tendency	reporter-workitem-num, reporter-recent-workitem-num, reporter-stable-docchg-rate, reporter-stable-sprintchg-rate, assignee-workitem-num, assignee-recent-workitem-num, assignee-stable-docchg-rate, assignee-stable-sprintchg-rate	The past tendency of the work items reporter and assignee in terms of the involved work items in the past, the recently involved work items, the rate of involved work items that did not have documentation changes, and the rate of involved work items that did not have sprint changes.
Readability	flesch, fog, lix, kincaid, ari, coleman-lieu, smog, wordcount	The scores of how difficult to understand the description text. To estimate the readability score, we used seven readability tests, i.e., Flesch [13], Fog [16], Lix [22], Flesch-Kincaid [26], Automated Readability Index [49], Coleman-Liau [9], and SMOG [36].

*Due to space limitation, we provided the description of each metrics and how we extract them in our replication package.¹

補足 | 調査対象としたプロジェクト

Table 2: An overview of the studied projects.

Project	Period	#work items		
		Total	Studied	Chg*
Apache DataLab (DATA LAB)	2018-2021	1,896	695	157
Data Management (DM)	2014-2021	30,649	6,831	830
Apache Mesosphere (MESOS)	2011-2021	10,190	1,258	92
Talend Data Prep (TDP)	2015-2021	6,060	1,780	554
Talend Data Quality (TDQ)	2008-2021	15,761	2,858	833
Titanium SDK/CLI (TIMOB)	2011-2021	22,446	1,165	108
Talend MDM (TMDM)	2009-2021	9,115	553	80
Talend Unified Platform (TUP)	2011-2021	19,430	1,150	84
Spring XD (XD)	2013-2018	3,707	1,441	160

Chg*: work items that have documentation changes after sprint assignment time

補足資料 | OneR アルゴリズム

OneR アルゴリズム

- 最も単純な規則を導出するアルゴリズムの1つ
- 各特徴量に対して以下のように処理を行う
 - 特徴量の値と出力間でクロステーブルを作成
 - 各特徴量の値ごとに、クロステーブルから読み取れる特定の特徴量を持つインスタンスの、最も頻度の高いクラスを予測するための規則を作成
 - 特徴量に対して、規則の誤差の合計を計算
 - 誤差の合計が最小となる特徴量を選択
- 詳細は以下のリンクの4.5.1を参照
 - <https://hacarus.github.io/interpretable-ml-book-ja/rules.html>