

Python APIドキュメントにおける 非推奨ドキュメントの検出

オープンソースソフトウェア工学研究室

学籍番号 60230010

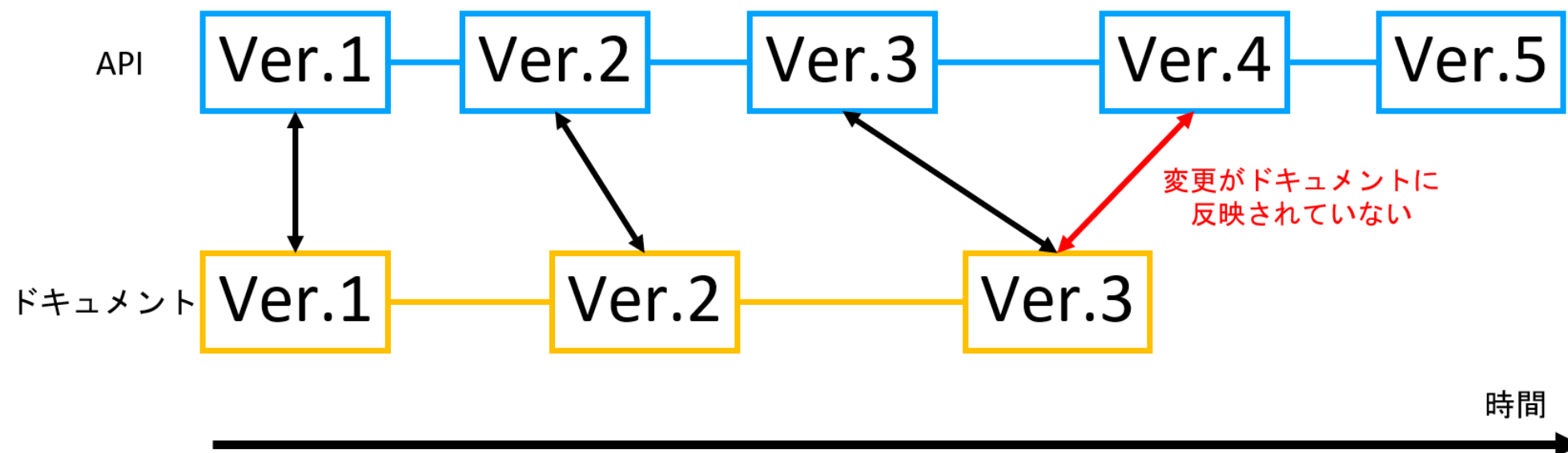
氏名 井町 瑠貴

背景 | ドキュメント

ドキュメント

- プログラマ向けに書かれた説明書・チュートリアル
 - 機能, 仕様, 履歴, 使い方, 注意事項 など
 - 本研究ではAPIドキュメントを指す

APIに変更がある際, ドキュメントの更新が必要な場合がある



背景 | APIの更新

APIの更新

- 以下のような場合，APIは更新される必要がある
 - 新機能の追加，機能の拡張
 - バグや脆弱性の修正

一般的に既存のAPIはしばらくの間保持される

- 後方互換性維持のため

APIの更新の際，推奨されること

- 新しいAPIの導入
- 既存のAPIは非推奨とする（新しく利用しない）

関連研究 | コード・ドキュメントの対応

| ドキュメントとは？

- APIドキュメント
 - ソースコードとAPIドキュメント間の不整合を検出
 - APIドキュメントの欠陥の自動検出
 - APIドキュメントの自動更新（クラス・メソッド名）
- コメント（JavaDoc）
 - コード・コメントの不整合
 - コードの修正に伴って、既存のコメントを自動更新

| 対象言語

- 大半がjavaプロジェクトを対象
 - Pythonプロジェクトではどうなのか？

関連研究 | python API

pythonを対象とした研究

- 非推奨のライブラリAPIがどのように扱われるか
 - APIの廃止はうまく処理されていない
(→適切に処理するために対処が必要)
- APIScanner
 - ライブラリ内の非推奨APIを自動的に検出
- Tensorflow
 - API進化の傾向を分析
- ライブラリマイグレーション
 - 使用中のライブラリを代替ライブラリに置き換えるプロセス

関連研究 | 非推奨APIの調査①

| pythonの非推奨のライブラリAPIがどのように扱われるか

- どのように対応するのか？
 - 推奨された方法を見捨て、アドホックな実装を行う
- 非推奨のAPIはどのようにドキュメント化されるか
 - 公式ドキュメントに記載されていない非推奨のライブラリAPIが相当数存在
 - 記載されていても、代替案が提供されていないものばかり
- 非推奨のライブラリAPIをどのように使用しているか
 - 非推奨APIの使用状況はプロジェクトの進化の過程でほとんど変更されない

関連研究 | 非推奨APIの調査②

- | 非推奨のAPIはどのようにドキュメント化されるか
 - 非推奨APIの約1/4は文書化されていない
 - よくて8%, 悪くて47%
 - 文書化されたものの約1/3は代替手段を提供していない
 - Javaで行われた調査と同等の割合
 - 非推奨ドキュメントに以下のような記述も見られた
 - 明確でない
 - 曖昧である
 - 間違っている

関連研究を踏まえて

| APIが更新された際に既存のAPIが非推奨となる

| 非推奨なAPIに対応した新しいAPIがあると考えられる

- 機能そのものがなくなる場合も一部あると考えられる

| 非推奨なAPIに対応した新しいAPIを検出することが有効ではないか？

- 関連研究では手動で対応付け

• 変更履歴, リリースノート

- API更新→非推奨になったAPIの検出？

検討が必要なこと

非推奨とは？

- 関連研究では, 「リリースノートに非推奨と記載されているAPI」を非推奨なAPIとしている
- 今後廃止される予定であるAPI？
- コードの更新だけではわからない
 - 今後の開発予定なども関わる

非推奨は誰が決める？

- API開発者が決めている？

検討が必要なこと

非推奨≠ダメ（API利用者目線）

- 人によっては非推奨なAPIを使ったままでもいい（困らない）場合もある

そもそもドキュメントに注目する必要があるのか

- 目指していること
 - （非推奨な）APIを利用して不利益を被ることがない
- 「APIの更新に伴って、API利用者のコードが更新させる」ことも手段のひとつではないか？
 - 今後の開発予定なども関わる

検討が必要なこと

研究目的

- APIを使う人が，APIドキュメントを参照して不利益を被らない状態にすること
 - ドキュメントに推奨された情報がある状態にする
 - 非推奨であることを提示するだけでは足りない？

現在行っていること

関連研究の調査

- Pythonを対象とした非推奨APIに関する研究
- 以下についてどのように考えられているか
 - 非推奨とは（判定方法）
 - ドキュメント化とは（なにを指している）