ソフトウェアアップサイクルのためのプロジェクトカタログの主要属性の検討

1-8

Preliminary Study of Key Attributes on Project Catalog for Software Upcycling

陳　思楠
Sinan Chen

中村 匡秀
Masahide Nakamura

（神戸大学）
(Kobe University)

【研究の背景】ソフトウェアアップサイクル
・時代におけるニーズの変化や技術の進歩
→　ソフトウェア中の**技術的負債**が増加
・過去のソフトウェア資産を利活用する方向性

項目	リユース	リサイクル	アップサイクル
【例】メガソーラーの再利用	解体して，家庭用途にリユース	ガラス，金属に分解し，リサイクル	おしゃれなテーブルにアップサイクル
ソフトウェアの再利用	「そのまま」使えるのは原始的な部分のみ．多様なコンテキストで「そのまま」使うのは無理難題なのでは？	無形物なので，資源に分解して再利用ということ自体できないし，意味がない．	既存SWの特徴・コードを生かしつつ，別用途に流用するために手を加える．そんなことができるのか？

【課題】ソフトウェアの説明はReadmeに依存
・Readmeの**書き方がバラバラ，一覧性が悪い**
・Readmeが**管理されていない**SWも大量に存在

【先行研究】クラス名単語に基くSWの概要推測
・Readmeが欠如でも，概要を効率に把握 2-3

例：チャットボット？　町人口の統計？　健康管理？質問？

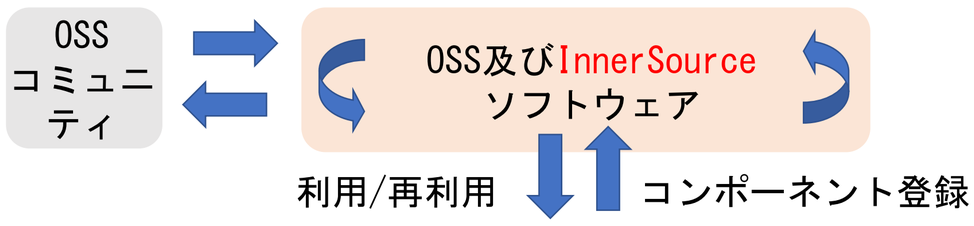
【本研究の着目】SWプロジェクトのカタログ化
・関連情報をカテゴリー別に抽出・可視化
→　素材発見やSW理解の促進，管理の容易化

一般商品のカタログ

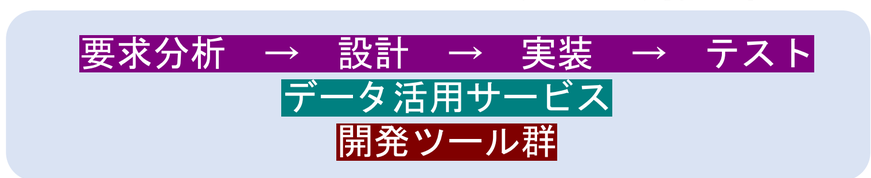
SWプロジェクトのカタログ

【関連研究】SWコンポーネントの管理と再利用

(1) 東芝グループの**ソフトウェア資産管理システム**　[参考1]



(2) プロジェクトごとの**ソフトウェア開発管理システム**



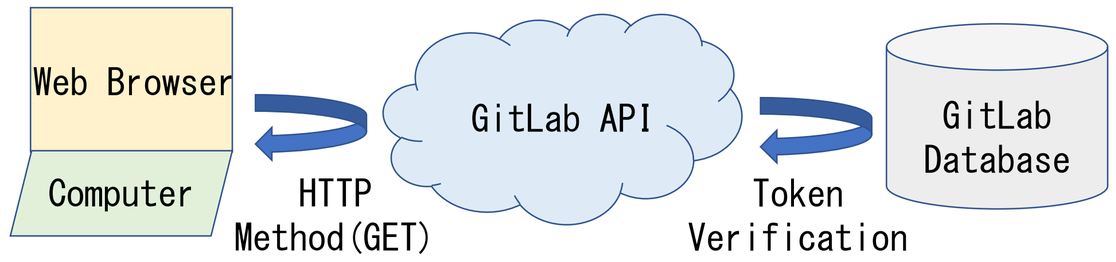
→　不足：基盤層の部品や特定用途が多かった

【本研究の目的】SWアップサイクルのための**プロジェクトカタログの主要属性**を予備検討すること
A1：SWプロジェクトカタログの属性を整理．A2：GitLabからプロジェクト属性の抽出手法を調査

【A1】カタログ属性に必要なカテゴリーと項目

カテゴリー	基本情報	アップサイクルの文脈	機能情報	非機能情報
項目	名前， 作成者， 概要， 作成日時， 利用形態， 言語， 階層構造， サイズ	プログラムコード， 設計， コンセプト， プロジェクト全体	主要機能， 入力データ， 出力データ， UI画面	データベース， ライブラリー， 性能効率性， 互換性， 使用性， 信頼性， セキュリティ， 保守性， 移植性

【A2】属性抽出のための概念図と調査した手法



項目	調査した手法	備考
名前，作成者，概要，作成日時	GitLab API [参考2]	../api/v4/projects
階層構造	同上	../api/v4/projects/{id}//repository/tree
プログラムコード	Python: import gitlab [参考3]	Projects → File → Decode
言語	Python: from guesslang import Guess [参考4]	guess.language_name(code)
設計，入力データ，出力データ	コード実行可視化のWebツール [参考5]	Python，Java，C，C++，JavaScript，Rubyをサポート

【まとめ】SWアップサイクルのためのプロジェクトのカタログ属性整理と既存手法調査を行った

【今後の課題】属性抽出手法の調査を深めつつ，中村研のSW資産のカタログ化の実装を行いたい

参考1: https://www.global.toshiba/content/dam/toshiba/jp/technology/corporate/review/2021/02/1-5.pdf
参考2: https://docs.gitlab.com/ee/api/
参考3: https://python-gitlab.readthedocs.io/en/stable/gl_objects/projects.html
参考4: https://github.com/yoeo/guesslang
参考5: https://pythontutor.com/