06: DDBJ Workflow Execution Service

大田達郎 (ライフサイエンス統合データベースセンター)

Licensed under a Creative Commons 表示4.0 国際ライセンス (c) 2021 大田達郎 (ライフサイエンス統合データベースセンター)

謝辞

- The Sapporo team: Hirotaka Suetake (main developer), Manabu Ishii, Tomoya Tanjo
- The DDBJ Human Genome working group
- Prof. Takeo Igarashi and Dr. Tsukasa Fukusato (UTokyo)
- Open-source communities
  - Workflow Meetup Japan
  - BioHackathon Europe and Japan
  - Workflow language and framework communities

Licensed under a Creative Commons 表示4.0 国際ライセンス (c) 2021 大田達郎 (ライフサイエンス統合データベースセンター)

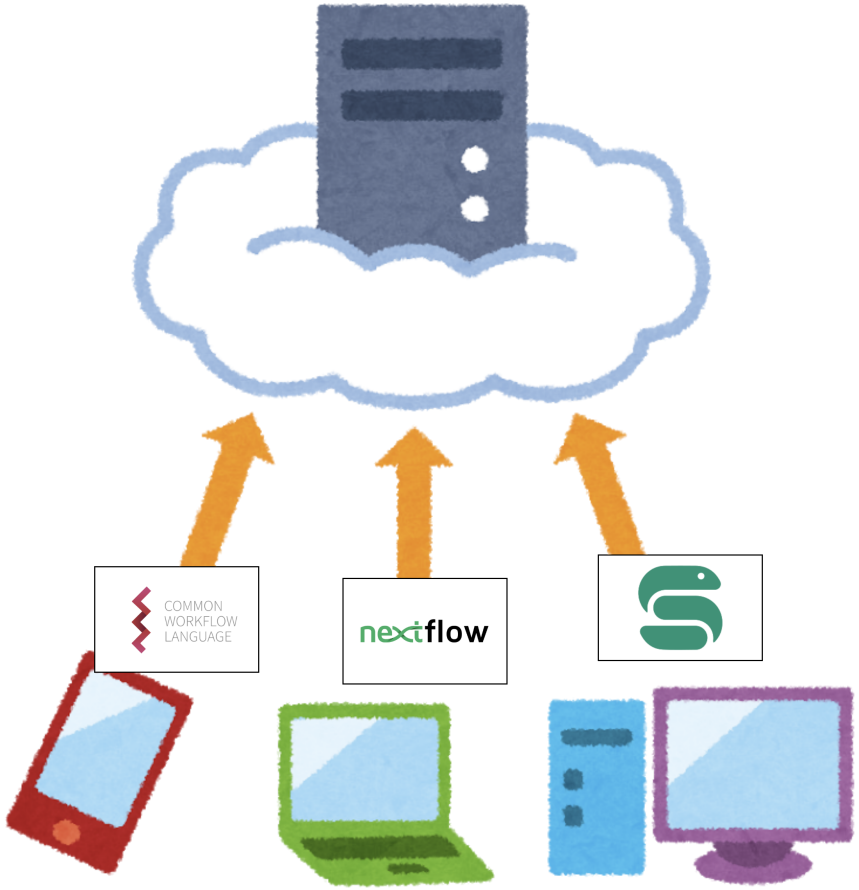
発表要旨

DDBJでは国内の生命科学データ解析をサポートするため、共用のスーパーコンピューターシステム（DDBJスパコン）を提供している。さらに、センシティブなデータを取り扱うための個人ゲノム解析環境を有償で提供している。DDBJスパコンでは、生命科学データ解析でよく用いられるコマンドラインソフトウェア（ツール）を Singularity コンテナで提供するなど、各ユーザがなるべく負担をかけずに解析を実行できるようサポートしている。しかし、データ解析では目的に応じてツールを組み合わせたワークフローを構築することが一般的であり、この作業はユーザらに委ねられている。そのため、ワークフローを共有する仕組みを作ることで、ユーザの負担はさらに軽減できるものと考えられる。DDBJ ではワークフローの共有と実行を簡易に行うことのできる DDBJ WES を開発している。DDBJ WES を用いることにより、一般的なワークフロー言語を用いて構築されたワークフローを使ったデータ解析を容易に行うことができる。

Licensed under a Creative Commons 表示4.0 国際ライセンス (c) 2021 大田達郎 (ライフサイエンス統合データベースセンター)

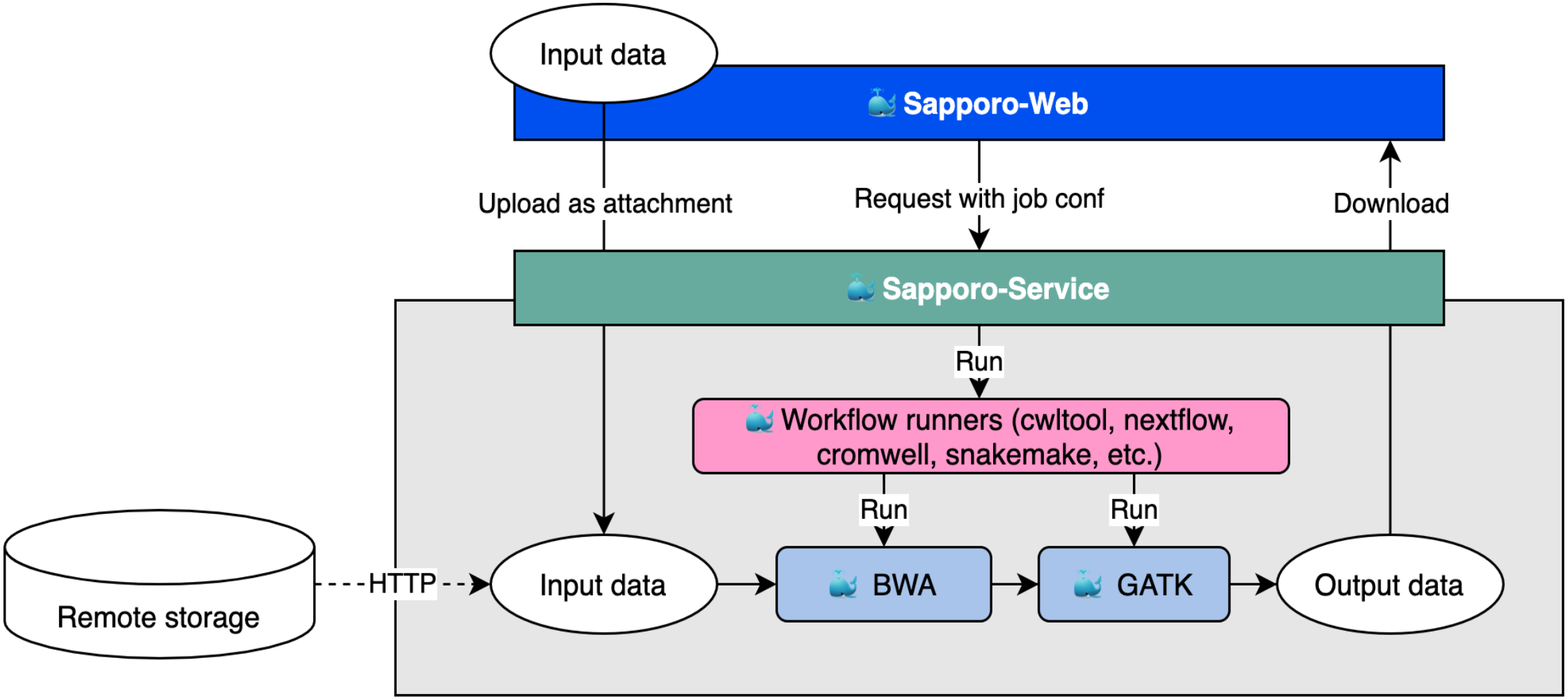
Sapporo のコンセプト

- GA4GH WES standard の実装
- 2つのコンポーネント: API サーバ と ブラウザベースの GUI
- **複数のワークフロー実行系をサポート**
  - `cwltool`, `nextflow`, `cromwell`, `snakemake`



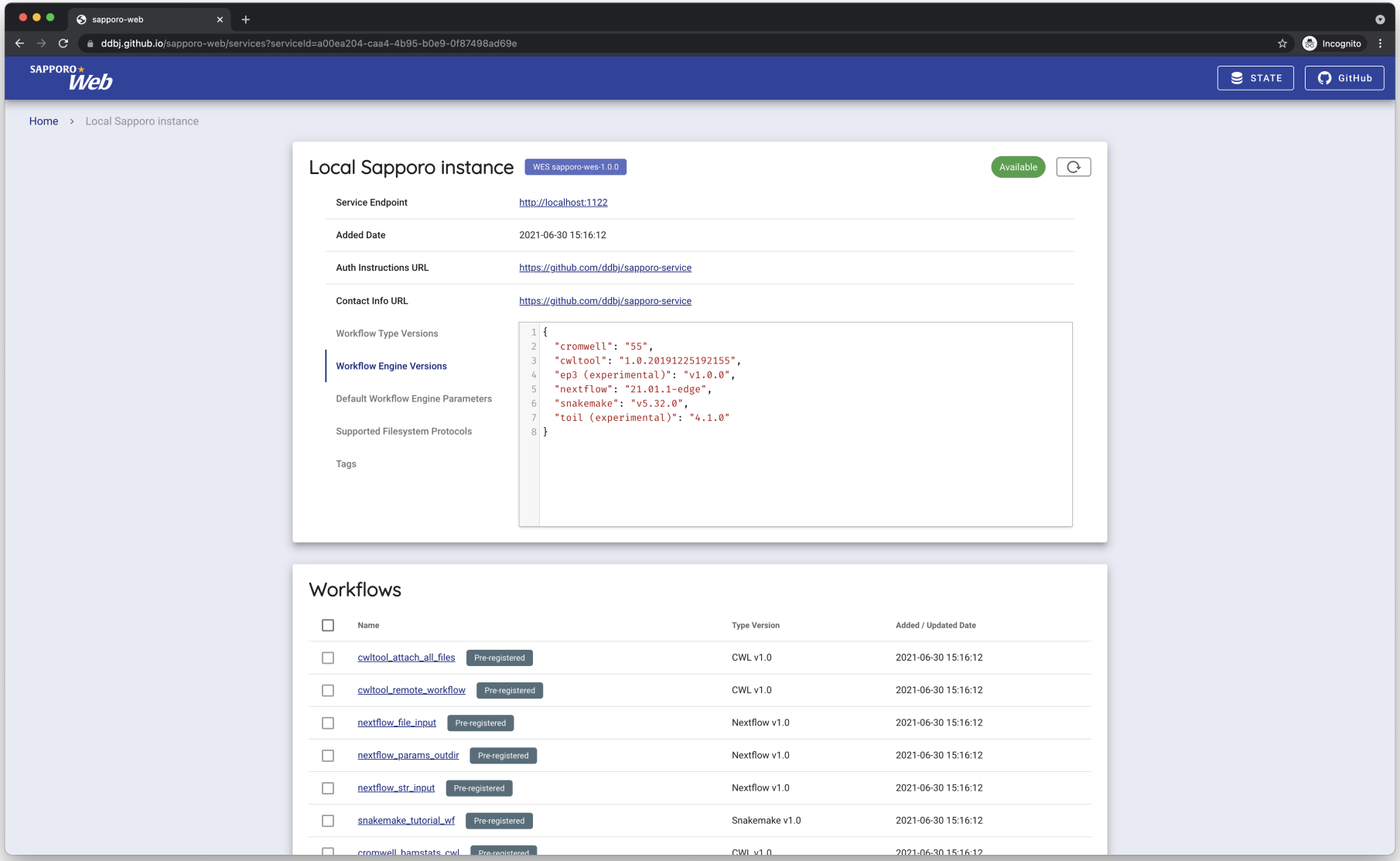
Licensed under a Creative Commons 表示4.0 国際ライセンス (c) 2021 大田達郎 (ライフサイエンス統合データベースセンター)

Sapporo のアーキテクチャ



Licensed under a Creative Commons 表示4.0 国際ライセンス (c) 2021 大田達郎 (ライフサイエンス統合データベースセンター)

GUI (Sapporo-web) によるワークフロー実行



Licensed under a Creative Commons 表示4.0 国際ライセンス (c) 2021 大田達郎 (ライフサイエンス統合データベースセンター)

課題

同じ研究者が学習コストをあまりかけることなく複数の言語の資産を使うために:

- ✅ 実行系のインストールとデプロイをサポート
- ✅ ワークフロー実行のI/Fを共通化
- ⬜ ワークフローパラメータ設定のI/Fを共通化
- ⬜ ファイルやディレクトリのステージングをサポート

言語によってはWESに向いていないものも: 個別に作り込みが必要

Licensed under a Creative Commons 表示4.0 国際ライセンス (c) 2021 大田達郎 (ライフサイエンス統合データベースセンター)

外部リンク

- Project documentation: https://github.com/ddbj/sapporo
- WES server source code repository: https://github.com/ddbj/sapporo-service
- Web GUI source code repository: https://github.com/ddbj/sapporo-web
- Public web GUI (require running WES server): https://ddbj.github.io/sapporo-web
- Hands-on Tutorial: https://ddbj.github.io/sapporo/GettingStarted

Licensed under a Creative Commons 表示4.0 国際ライセンス (c) 2021 大田達郎 (ライフサイエンス統合データベースセンター)