

DDBJが提供するグループ共有サービス

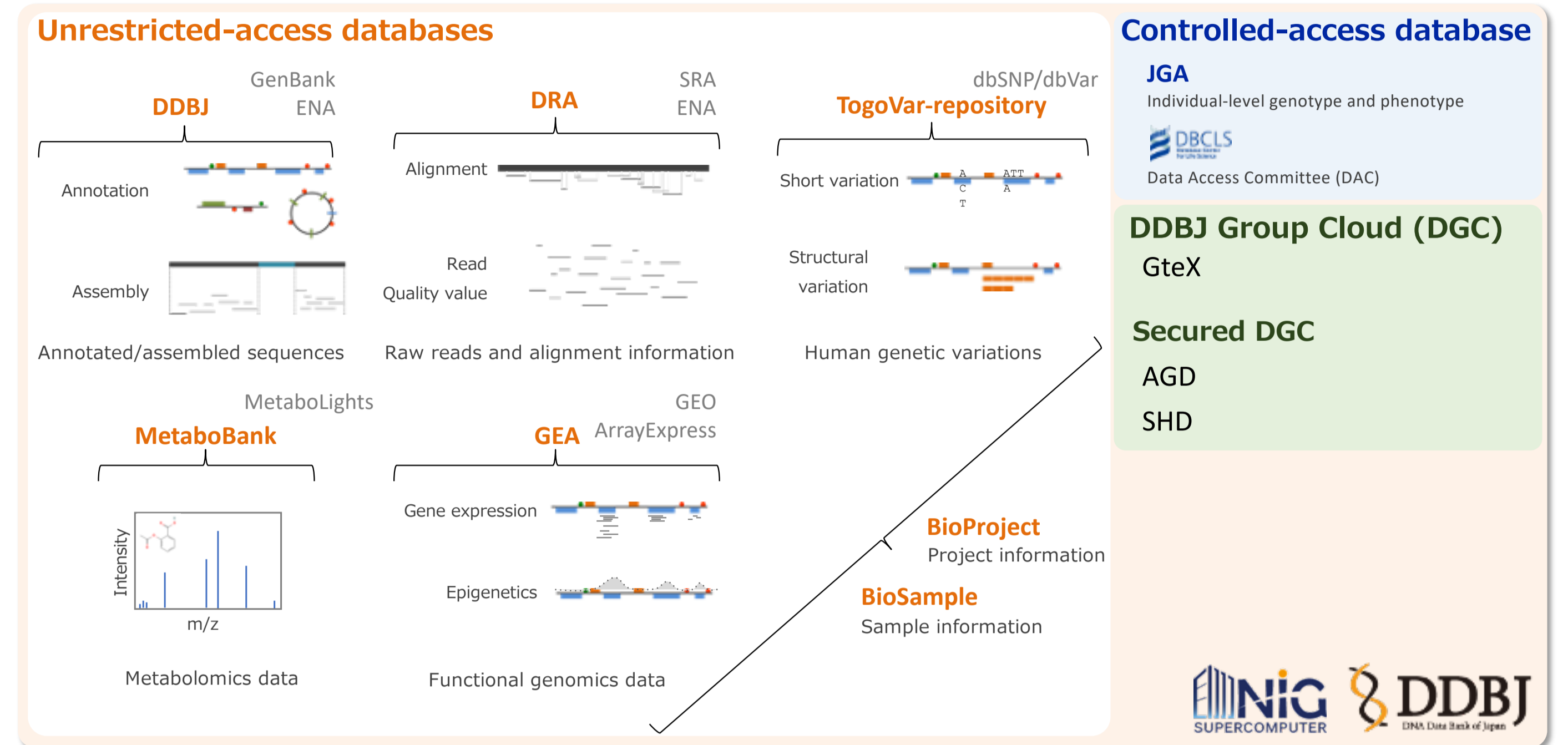
藤澤貴智、荒武、児玉悠一、福田亜沙美、谷澤靖洋、丹生智也、小笠原理、中村保一、有田正規
(国立遺伝学研究所 生命情報・DDBJセンター)

背景

- 研究データの種別や規模およびデータ解析に必要な計算機リソースは、データ共有プラットフォーム毎に異なる
- 研究データは計算ツールを備えた研究機関の計算機センターに保存されているが、データ公開については後回しになりがち
- オープン・アンド・クローズ戦略に基づくより良い研究データの管理・利活用の実現のためには様々な課題

DDBJグループクラウド (DGC)

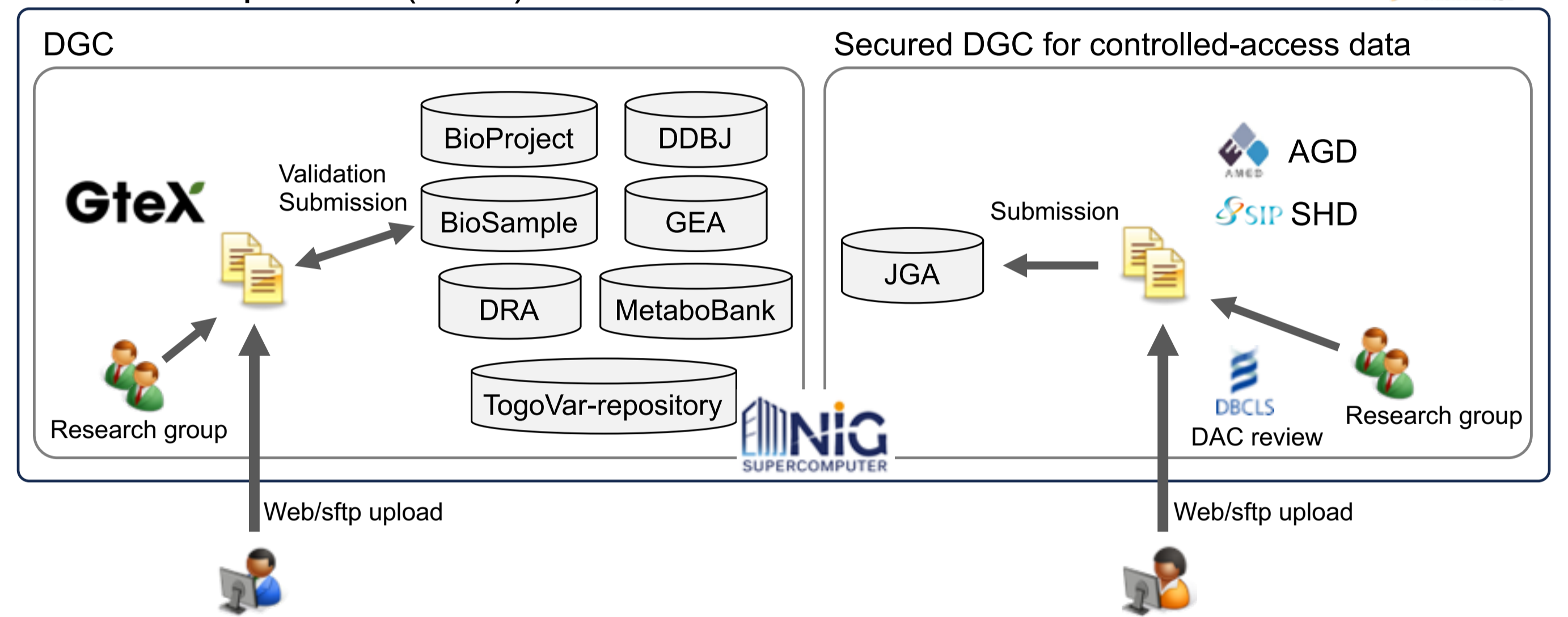
- 出版前データ共有のためのサービスを2017年から提供開始
- カスタマイズや運用にかかる費用についてはDGC加入者が負担の有料サービス
- ヒト由来試料を含むデータ共有においては、NBDCヒトデータ共有ガイドラインに従って、NBDC DAC (Data access committee) により、JGA制限公開と同等の事前審査を実施。(2024年4月よりDBCLS DACに移管)



DGCの運用実績

- AMED genome group sharing Database (AGD) 2017年～
- SIP Healthcare group sharing Database (SHD) 2021~2023年
- 革新的GX技術創出事業 (GteX) バイオものづくり領域 先端的植物バイオものづくり基盤の構築 2024年4月～
- グリーンイノベーション基金事業 (GI) 「バイオものづくり技術によるCO2を直接原料としたカーボンリサイクルの推進」 2024年～

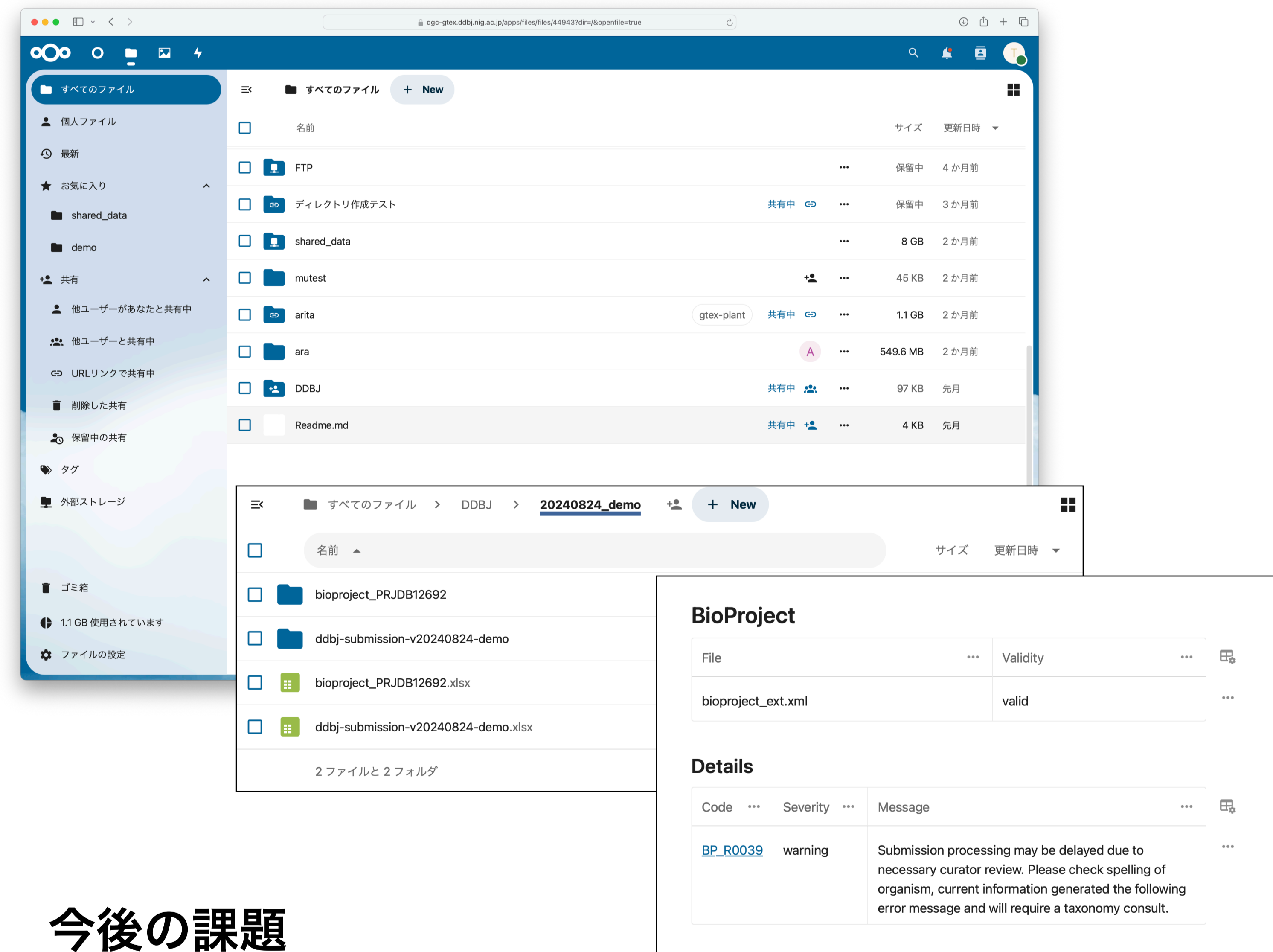
DDBJ Group Cloud (DGC) Service



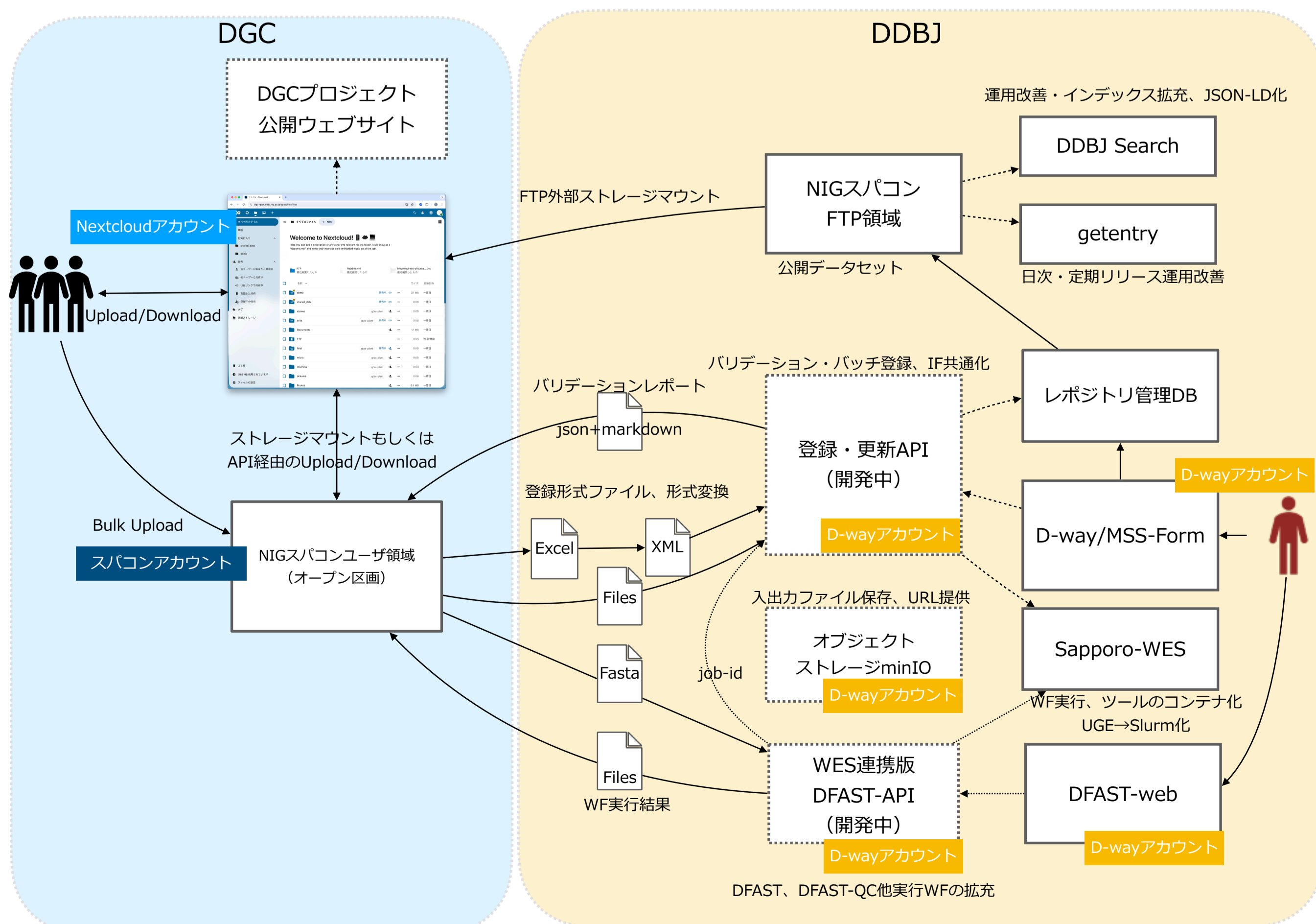
DGCの最近の取り組み

ヒト由来試料以外のデータ共有サービスの提供を開始

- nextcloudを用いて出版前データなどをグループアクセス制御に柔軟に対応
- BioProject/BioSample形式変換およびバリデーションサービスを利用可能
- INSDC公開データセットなどローカルアクセスが可能
- NIGスパコンで提供されるバイオインフォマティクスツールが利用可能
- データ公開を条件にアクセッション番号を付与を支援



DGC周辺の開発計画



今後の課題

- バリデーション実行自動化とレポート提供の自動化
- BioProject/BioSample以外のバリデーション、DB登録対象DBの拡充
- DFAST他データ解析WF連携と拡充
- DDBJ登録形式ファイル作成支援
- nextcloudアカウントとD-wayアカウント連携

