

バイオインフォマティクス分野の アプリケーション開発における GitHubの活用とその実際

情報・システム研究機構 データサイエンス共同利用基盤施設
ライフサイエンス統合データベースセンター (DBCLS)

小野 浩雅

小野 浩雅 Hiromasa Ono
hono@dbcls.rois.ac.jp
https://twitter.com/h_ono



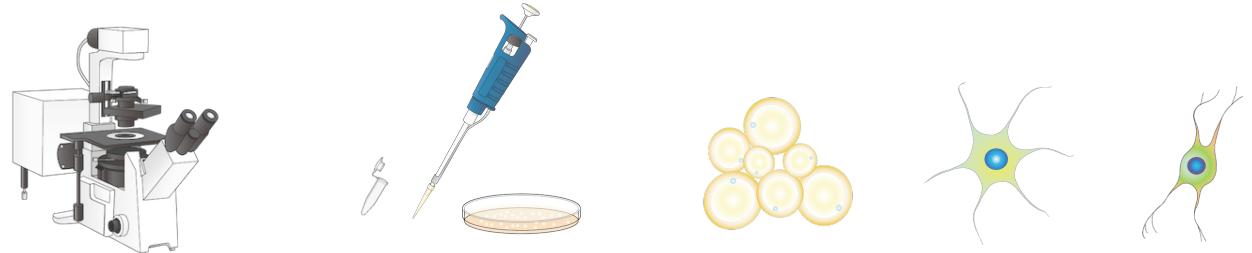
編集・管理者@TogoTV
生命科学分野の有用なDBやウェブツールの活用法を動画で紹介
<https://togotv.dbcls.jp/>

開発者@RefEx
正常組織や細胞株の遺伝子発現データを簡単に検索できるデータベース
<https://refex.dbcls.jp/>

講師@統合データベース講習会(AJACS)
生命科学分野のデータベース統合を担う人材育成(主に遺伝子発現解析を担当)
<https://github.com/AJACS-training/>



Wet と Dry



学部

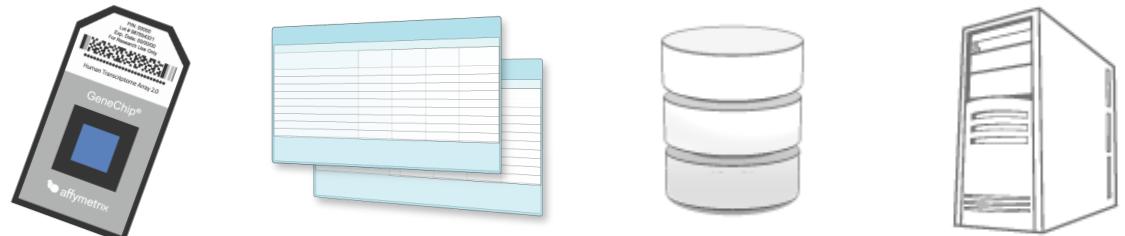
ブタ成熟脂肪細胞に由来する前駆脂肪細胞(DFAT-P)の神経細胞への分化転換に関する研究

修士

ブタ成熟脂肪細胞および顆粒膜細胞における脱分化機構の網羅的解析

博士

ブタ成熟脂肪細胞および卵胞顆粒層細胞における脱分化ならびに多能性獲得機構の網羅的解析



ライフサイエンス統合データベースセンター (DBCLS)

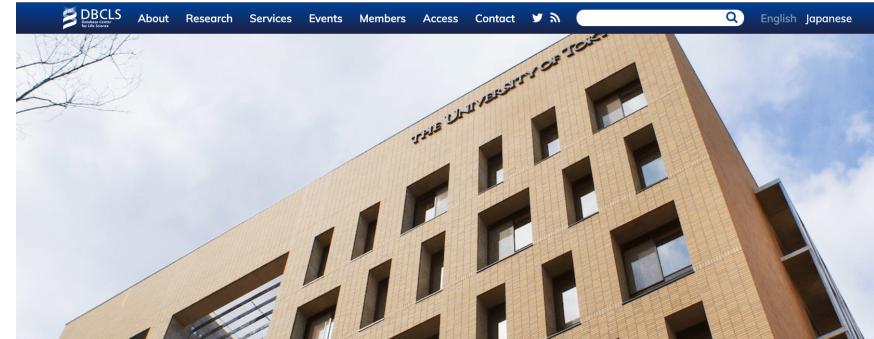


大学共同利用機関法人
情報・システム研究機構
データサイエンス共同利用基盤施設
ライフサイエンス統合データベースセンター
 DataBase Center for Life Science

生命科学分野におけるデータベース統合化の
拠点を形成することを目的に、2007年4月に
設立

生命科学分野のデータベースとサービスの
専門機関として、国内外のデータベース統合化と
利用者の利便性向上のための基盤技術
開発を行う

常勤で27名の研究員・技術員・事務員



News

2021. 10. 05 データベース探索のフレームワーク「TogoDX」を開発し、その応用事例として「TogoDX/Human」を公開しました
DBCLSは、NBDCと共にデータベース探索のためのフレームワーク「TogoDX (Togo Da...)
●広報 ●サービス

2021. 10. 01 Genomics & Informatics誌の2021年9月号が第7回Biomedical Linked Annotation Hackathon (BLAH7)の特集号として出版されました
Genomics & Informatics誌の2021年9月号が第7回Biomedical...
●広報

2021. 09. 30 [プレスリリース] カイコ遺伝子発現データを拡張・公開しました
当センターの坊農 雄雅 寄員教授が参加する、農研機構 横井 邦 主任研究員らの研究グループは、シルク...
●広報

2021. 09. 22 2021年日本バイオインフォマティクス学会年会・第10回生命医療情報学連合大会 (IIBMP2021) にブースを出展します
DBCLSは2021年日本バイオインフォマティクス学会年会・第10回生命医療情報学連合大会 (IIBM...
●イベント ●広報

2021. 09. 06 【オンライン開催】トーコーの日シンポジウム2021の参加者募集中です
バイオサイエンスデータベースセンター(NBDC)では、毎年10月5日を「トーコーの日」とし、ライフサ...
●イベント ●募集 ●広報

more

Site map

About
センターについて
沿革
よくある質問
サイトポリシー
ロゴとリボンについて

Research
研究開発
発表論文等一覧

Services
サービス一覧
引用文献一覧

Events
開催イベント一覧

News
Members
Access
Contact

DBCLS
Database Center for Life Science

大学共同利用機関法人 情報・システム研究機構
データサイエンス共同利用基盤施設
ライフサイエンス統合データベースセンター

Address
〒277-0871
千葉県柏市若柴178-4-4
東京大学柏の葉キャンパス
駒場サテライト6階
TEL. 04-7135-5508 (代表)
FAX. 04-7135-5534 (代表)

Links
統合ホームページ
DBCLS on YouTubeLive
データサイエンス共同利用基盤施設
情報・システム研究機構
文部科学省

大学共同利用機関法人 と 情報・システム研究機構

全国の研究者に共同利用・共同研究の場を提供する、日本で4つの中核的研究拠点です。

[HOME](#) | [研究所所在地一覧](#) | [お問い合わせ](#) | [サイトポリシー](#)

大学共同利用機関法人

Inter-University Research Institute Corporations

大学共同利用機関法人とは

機関・研究所一覧

シンポジウム

活動報告・評価

関連リンク

機関・研究所一覧

人間文化研究機構 (NIHU)

- ▶ 国立歴史民俗博物館
- ▶ 国文学研究資料館
- ▶ 国立国語研究所
- ▶ 國際日本文化研究センター
- ▶ 総合地球環境学研究所
- ▶ 国立民族学博物館

人間文化研究機構要覧

▶ 要覧2015 pdfファイル

自然科学研究機構 (NINS)

- ▶ 国立天文台
- ▶ 核融合科学研究所
- ▶ 基礎生物学研究所
- ▶ 生理学研究所
- ▶ 分子科学研究所
- ▶ 岡崎バイオサイエンスセンター
- ▶ 新分野創成センター

自然科学研究機構パンフレット

- ▶ 機構パンフレット2013(和文)pdfファイル
- ▶ 機構パンフレット2013(英文)pdfファイル

高エネルギー加速器研究機構 (KEK)

- ▶ 素粒子原子核研究所
- ▶ 物質構造科学研究所
- ▶ 加速器研究施設
- ▶ 共通基盤研究施設

高エネルギー加速器研究機構要覧

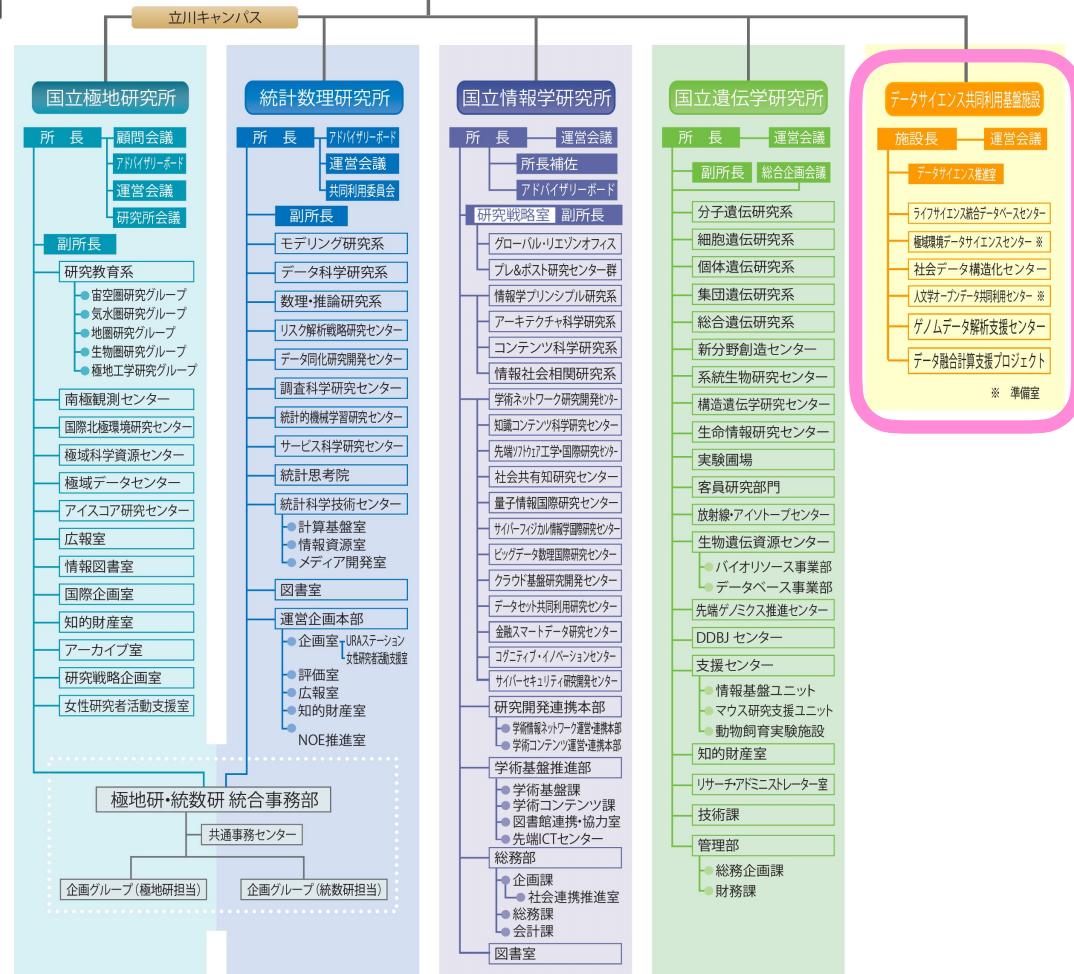
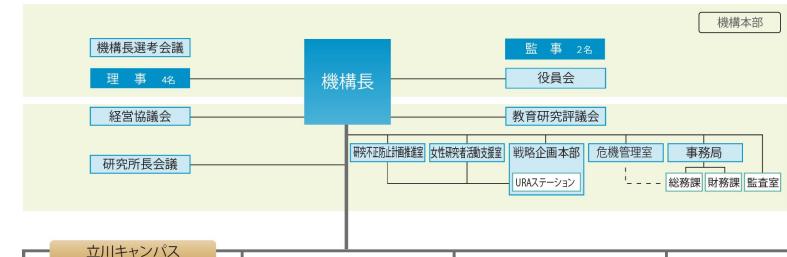
- ▶ KEK2015要覧 pdfファイル
- ▶ KEKパンフレット(英文) pdfファイル

情報・システム研究機構 (ROIS)

- ▶ 国立極地研究所
- ▶ 国立情報学研究所
- ▶ 統計数理研究所
- ▶ 国立遺伝学研究所
- ▶ ライフサイエンス統合データベースセンター
- ▶ 新領域融合研究センター

情報・システム研究機構要覧

- ▶ 2013年度要覧(和文)pdfファイル
- ▶ 2013-2014年度概要(英文)pdfファイル



Licensed under CC-BY 4.0 © 2021 Hiromasa Ono (DBCLS)

(左図) <http://www.4kikou.org/list.html>

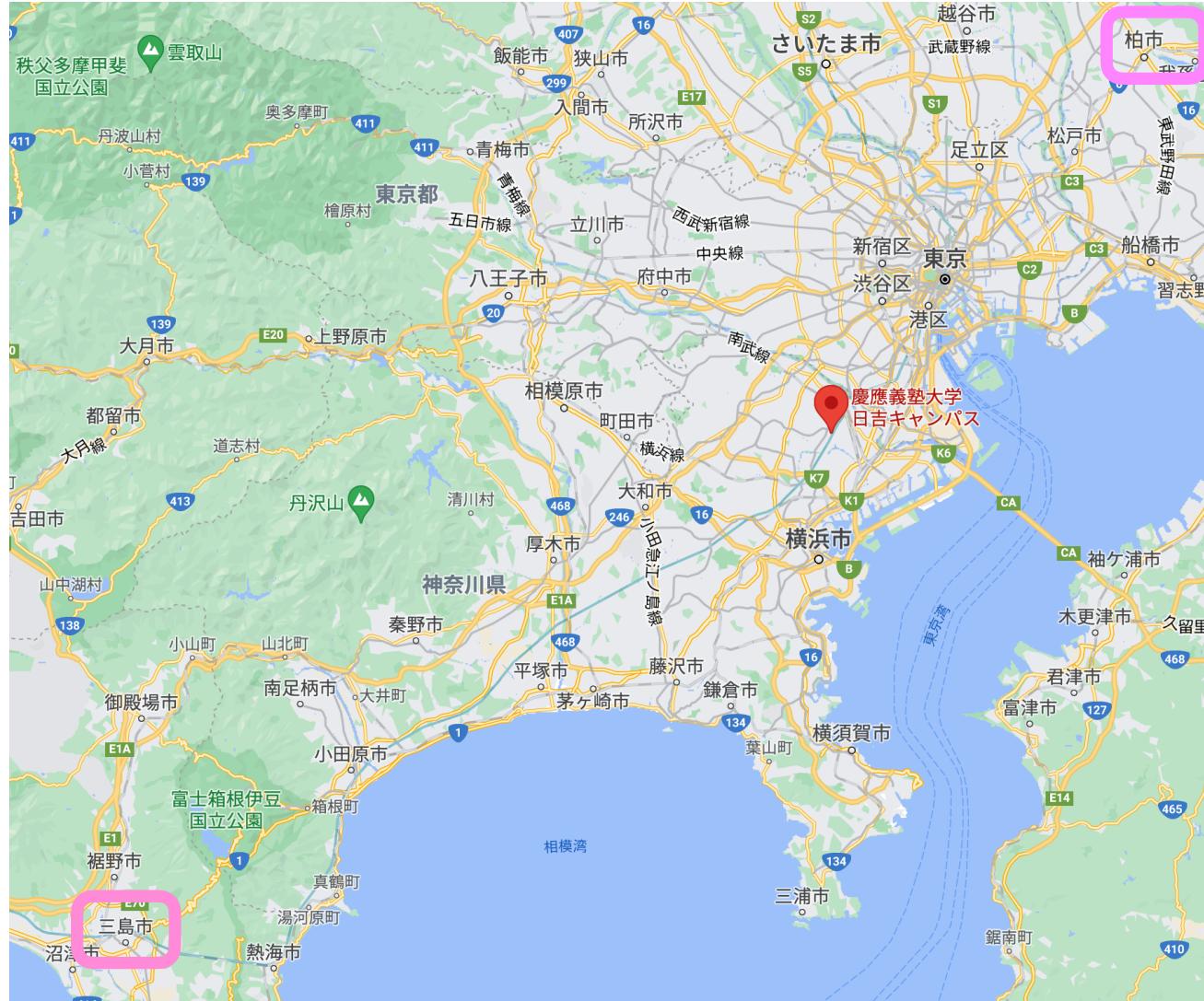
(右図) <http://www.rois.ac.jp/outline/index04.html> より

DBCLS は 千葉県柏市 と 静岡県三島市 にあります

(東京大学 柏の葉キャンパス駅前サテライト)(国立遺伝学研究所)



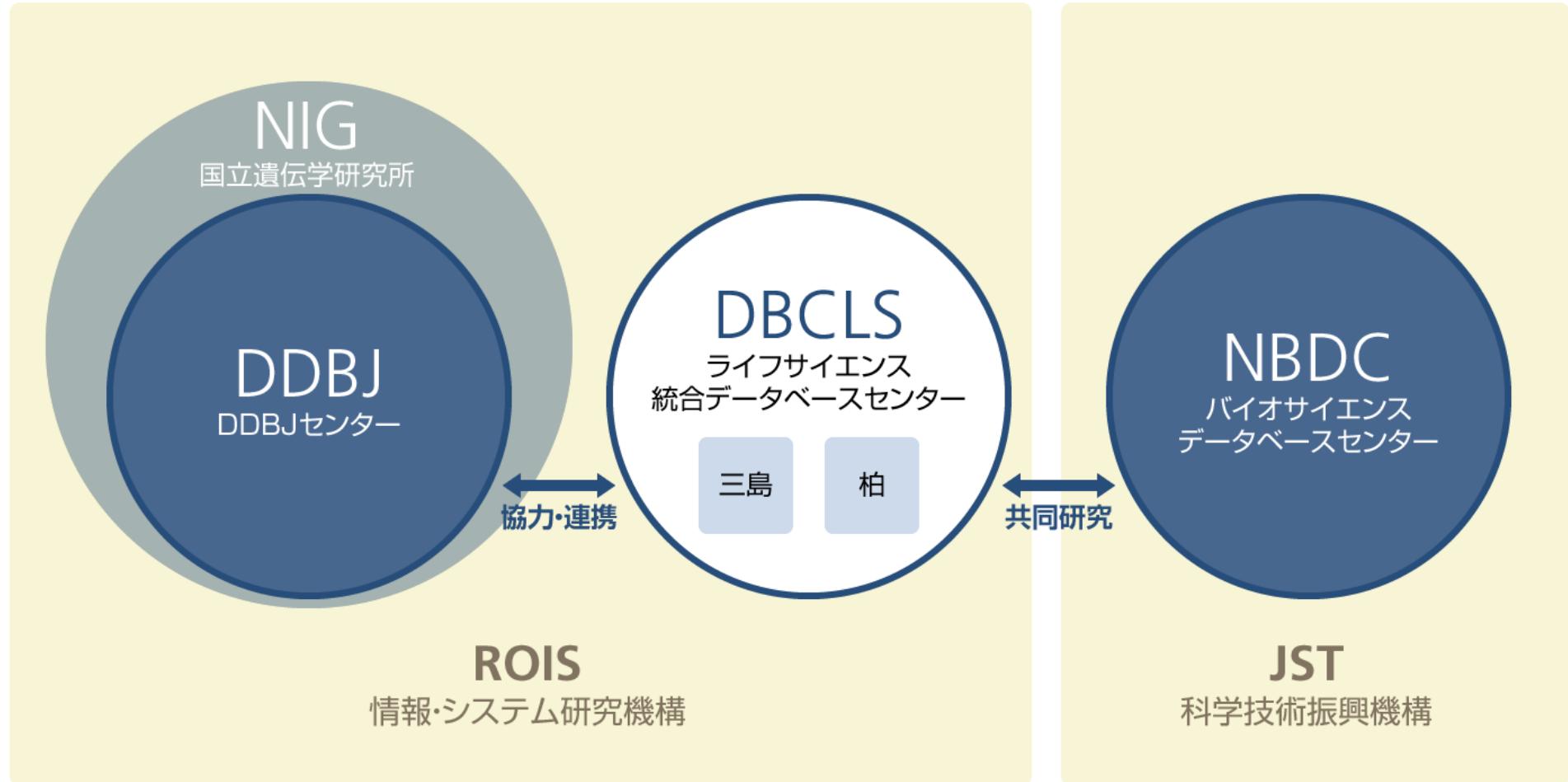
※現在はフルリモートワーク体制



Licensed under CC-BY 4.0 © 2021 Hiromasa Ono (DBCLS)

<https://www.google.co.jp/maps> より

DBCLS の立ち位置



NBDC/DBCLS 共同研究 のアウトライン

- ・ 多数あるデータベースのどれを使えばいいかわからない
- ・ データベースの有効な使い方を知りたい
- ・ 複数のデータベースを組み合わせて新たな知見を見出したい
→ これらに応える研究開発を行う

2. 統合化支援

DB構築ユーザ向け

統合化推進Pの支援
補完的有用DBのRDF化

国内の技術交流支援
国際連携による標準化

1. RDF統合化のための基盤技術開発

高度なRDF技術の実用化
分散環境・セキュリティ

標準化と国際・国内連携

W3C-HCLS など DB メタデータ
FALDO などの共通オントロジー
Identifiers.org などの URI

BioHackathon / SPARQLthon

DDBJ/EBI/NCBI/UniProt 等 RDF 連携
共通ソフトウェア・技術開発

3. エンドユーザ向け データベース利用技術 開発等

DB利用ユーザ向け

大規模データの活用
QAシステムを目標
日本語コンテンツ拡充
個人ゲノムや非公開データのアクセス制御

4. 既存・新規サービス/データベースの運用と拡張

インフラとしてウェブサービスやトリプルストアを持続的に運用 (NBDC/DDBJ連携)



GitHub(ギットハブ)ってなに

- ・ ソフトウェア開発プロジェクトのためのソースコード管理サービス
 - ・ Git (<https://git-scm.com/>) というプロジェクトの改訂履歴を管理し保存するバージョン管理システムを利用するためのハブ(Hub)がその名の由来
 - ・ コンピュータプログラムやソースコードの公開・共有・管理に利用されているウェブサービス
 - ・ コンピュータプログラムに限らず、Word文書や画像など、どんな種類のドキュメントでも、すべての段階の草案やアップデート履歴を保存し管理することができる
 - ・ GitHubは、グラフィカルなユーザインターフェースを提供しており、コマンドラインを使わずにウェブサイト上で様々な作業をすることができる
 - ・ アクセス制御機能や、タスク管理機能など、コラボレーションのための機能を多く備えているのが特長
- ・ 類似サービスに、BitbucketやGitLabがある
- ・ 2018年にMicrosoft社に買収され、傘下の一企業
 - ・ 2020年4月から、GitHubでも無料の非公開リポジトリを無制限に作成可能になった

私のGitHub <https://github.com/hiromasaono>

Search or jump to... Pull requests Issues Marketplace Explore

Overview Repositories 13 Projects Packages

Pinned

Customize your pins

- dbccls/RefEx** Public RefEx: a reference gene expression dataset of mammalian tissues and cell lines measured by different methods
- training** Public texts and scripts for training course
- utils** Public Shell
- wiki** Public
- brew-cask-upgrade-all** Public Forked from plasma0713/brew-cask-upgrade-all
- AJACS-training/AJACS-training.github.io** Public gituhub.io

Hiromasa Ono
hiromasaono
@dbcls , @AJACS-training

Edit profile

6 followers · 24 following · 14 contributions

@dbcls , ROIS
Mishima, Japan
hono@dbcls.rois.ac.jp

Achievements

Organizations

Contribution activity

482 contributions in the last year

Contribution settings ▾

Learn how we count contributions

Less More

2021

2020

2019

2018

2017

2016

Created 23 commits in 2 repositories
togodx/togodx-config-human 22 commits
dbccls/togoid-config 1 commit

Created a pull request in togodx/togodx-config-human that received 2 comments

Oct 1



DBCLS サービス <https://dbcls.rois.ac.jp/services.html> と GitHub レポジトリ



TopoID 生物のID こと ●データベース利用者 ▲アフィリエーション関係者 △大規模データ分析者	TogoWS 国内統一 イヌ ●データベース利用者 ▲アフィリエーション関係者 △大規模データ分析者	TogoDB 生物データ 整理す ます。 ●データベース利用者 ▲アフィリエーション関係者 △大規模データ分析者
BodyParts3D/notomography 頭でもうちょっとアシム解剖図を香港 し交換できるサービスです。作成した図 は、論文用いなど自由に公開することができます。 ●データベース利用者 ▲アフィリエーション関係者	新着論文レビュー 日本人を著者とする生命科学分野の論文 について、論文名と各自の評価による 日本のレビューアーを公開するサービスです。 ●データベース利用者 ▲アフィリエーション関係者	領域融合レビュー 生命科学分野における最新の研究成績につ いて、第一級の研究者や執筆による日本語のレ ビューを提供するサービスです。 ●データベース利用者 ▲アフィリエーション関係者
統合TV データ動画 いる ●データベース利用者 ▲アフィリエーション関係者	TogoGenome 生物学 情報 ●データベース利用者 ▲アフィリエーション関係者	GGGenome 統合微生物学 研究 ●データベース利用者 ▲アフィリエーション関係者
CRISPRdirect 入力 システム ●データベース利用者 ▲アフィリエーション関係者	統合遺伝子検索GGRNA 統合 配列 ●データベース利用者 ▲アフィリエーション関係者	Gendo 遺伝子、疾患について、関連する疾患、 薬理、生命現象などの知識をキー ワードでリスト表示するツールです。 ●データベース利用者 ▲アフィリエーション関係者
RefEx 遺伝子 検出 ●データベース利用者 ▲アフィリエーション関係者	AOE 公表 複数 アンド オート ●データベース利用者 ▲アフィリエーション関係者	DBCLS SRA 公表 アンド リード ●データベース利用者 ▲アフィリエーション関係者
PubCaseFinder 論著 少 に ●データベース利用者 ▲アフィリエーション関係者	PubAnnotation 文献アノテーションのレジストリ。特に 文献に登場する用語とその正式名称の 統合を行なうためのソースコレクションを統 合している ●自然言語処理 ▲アフィリエーション関係者 ●データベース利用者	PUBDICTIONARIES 論著 キーワード ●データベース利用者 ▲アフィリエーション関係者
inMeXes 生物 を ます ●データベース利用者 ▲アフィリエーション関係者	Allie 文献に登場する用語とその正式名称の 統合を行なうためのソースコレクションを統 合している ●自然言語処理 ▲アフィリエーション関係者 ●データベース利用者	Colil 文献に登場する用語とその正式名称の 統合を行なうためのソースコレクションを統 合している ●自然言語処理 ▲アフィリエーション関係者 ●データベース利用者
NEDC RDF Portal 日本学術会議によって貯蔵された 生物学データのRDFを集約しています。 ●データベース利用者 ▲アフィリエーション関係者 ●アカデミック研究者 ●大規模データ分析者	Togovar 日本ゲノム配列の個人によ る通達(パリティ)とそれに関連す る生物学データを収集・整備したデータ ベースです。 ●データベース利用者 ▲アフィリエーション関係者 ●アカデミック研究者 ●大規模データ分析者	

38サービス中、24サービスで
GitHubを利用中 (63%)



Licensed under CC-BY 4.0 © 2021 Hiromasa Ono (DBCLS)

GitHub DBCLS repository

<https://github.com/dbcls/>



Organization のレポジトリは簡単に作れる

組織単位
プロジェクト単位
研究室単位
など

The screenshot shows the GitHub organization page for 'DBCLS'. The top navigation bar includes links for Overview, Repositories (65), Packages, People (23), Teams (4), Projects, and Settings. Below the navigation is a section titled 'Pinned' which displays the message 'dbcls doesn't have any pinned public repositories yet.' To the right of this message is a 'Customize your pins' button. On the far right, there is a 'People' section showing a grid of user icons and a 'View all' link. Below the people section is a 'Top languages' section with Ruby, JavaScript, TypeScript, HTML, and Java represented by colored dots. At the bottom left is a 'Most used topics' section with 'rdf' and 'sparql' highlighted.

Repositories

- rdf-config** [Public]
Ruby ★ 2 MIT 7 0 1 Updated 3 hours ago
- togoid-config** [Public]
Ruby ★ 2 MIT 8 6 1 Updated 14 hours ago
- orthology** [Public]
JavaScript ★ 0 0 0 0 Updated yesterday
- graphdbs** [Public]
Shell ★ 0 1 0 0 Updated yesterday
- website** [Public]
HTML ★ 2 4 0 3 Updated 2 days ago
- togodb** [Public]
JavaScript ★ 0 MIT 1 0 3 Updated 9 days ago



Licensed under CC-BY 4.0 © 2021 Hiromasa Ono (DBCLS)

本日紹介するGitHubで運用・開発されているサービス



- DBCLS ウェブサイト <https://dbcls.rois.ac.jp/>
- TogoTV <https://togotv.dbcls.jp/>
 - 生命科学分野における有用な情報(動画)を紹介するウェブサイト
- TogоАD <https://togoid.dbcls.jp/>
 - 生命科学分野におけるデータベース(DB)のID間の対応関係を検索および変換することができるウェブツール
- TogoDX <https://togodx.dbcls.jp/human/>
 - 国内外のデータベースから収集・統合した、ヒトに関する遺伝子、タンパク質、化合物、疾患などの情報をワンストップで探索することができるウェブツール



DBCLS ウェブサイト

- 「GitHub Pages」を使って運用
 - GitHubの静的サイトのホスティングサービス
 - 高速・メンテナンスコスト低
- GitHub リポジトリにコンテンツを集約・管理

The screenshot shows the GitHub repository page for 'dbcls / website'. It displays a pull request with 8 reviews, 2 forks, and 1 star. The commit history shows a recent update from 'sakumakeiko' on October 5, 2021. A yellow banner at the top indicates potential security vulnerabilities in dependencies. Below the commit history is a table of contents for a file named '2021-10-05-post1.md'. The main content area contains a table with columns 'layout', 'published', 'tags', 'category', and 'title'. The 'published' column has a value of 'true'. The 'tags' column contains 'public_relations' and 'services'. The 'category' column has a value of 'ja'. The 'title' column contains the text: 'データベース探索のフレームワーク「TogoDX」を開発し、その応用事例として「TogoDX/Human」を公開しました'.

The screenshot shows the DBCLS website homepage. At the top right is the DBCLS logo and navigation links for About, Research, Services, Events, Members, Access, Contact, English, and Japanese. The main content area features a large image of a modern building with the text 'THE UNIVERSITY OF TOHO' visible. Below the image is a 'News' section with several articles listed by date. A blue arrow points from the GitHub screenshot to the news section on the website. The footer contains links for Site map, About, Research, Services, Events, News, Members, Access, and Contact. It also includes an Address section with the DBCLS building's location and contact information, and a Links section with logos for JBI portal, DDBJ, NBDC, and PDB.



TogoTV <https://togotv.dbcls.jp/>

- 生命科学分野における有用な情報を紹介するウェブサイト

- データベースやツールの動画マニュアル(698本)
- 講演(948本)や講習会動画(273本)
- イラスト(921本)

2021年9月末現在

- 動画マニュアル

- ウェブサイトへのアクセスの仕方から結果の解釈まで、**操作の一拳手一投足**がわかる
- 各動画はYouTubeで公開されており、**環境に応じた解像度、倍速表示等**で快適に閲覧可能
- 動画の概要を示す「見出し」から視聴箇所に移動可能
- 動画ファイルはダウンロード可能
- 1900本を超える動画が公開されており、**のべ200万回以上(月間3万回以上)**再生(2021年5月末現在)
- 自主学習、新人・後輩指導、講義・勉強会の教材として誰でも自由に利用できます



TOGO TV [トップページ](#) [TogoTVについて](#) [動画を探す](#) [画像を探す](#) [講習会資料を探す](#) [お問合せ](#) [キーワードから動画を探す](#)

新着動画

- TOGO TV TogoTVを使って生命科学分野のデータベースのID間の対応関係を検索し直す 2021-08-08 7分53秒
- TOGO TV クラウドを使ってゲノム編集のサイクルを絞り込む 2021-08-31 1時間38分51秒
- TOGO TV gnomAD v2.1を使ってヒトアソシエーションのアレル頻度を調べる 2021-08-23 11分27秒
- TOGO TV draw.ioを使ってペン図とフローチャートを作成する 2021-08-22 7分57秒
- TOGO TV クリエイティブ・コモンズ・ライセンスで著作権が許可された動画をYouTubeで見る方法 2021-08-21 3分15秒

見どころダイジェスト

バイオインフォマティクス超入門！ 統合TVを使い倒して必要なスキルを身につけよう！

この動画が再生できない場合は、YouTubeでご覧ください。 https://youtu.be/Twfrzg_wvI

2020.12.25 4535 作者: 小野 浩雅 / 製作: 小野 浩雅(ONO_Hiromasa)

バイオインフォマティクス超入門！ 統合TVを使い倒して必要なスキルを身につけよう！

本日の前回TVは、2020年12月23日に開催された**Informatics Career勉強会**からライフサイエンス統合データベースセンター(DBCLS) 小野 浩雅 様による「バイオインフォマティクス超入門！ 統合TVを使い倒して必要なスキルを身につけよう！」をお送りします。約45分です。

11月にリリースされた「[統合TV](#)」の機能や使い方を実演を交え紹介します。バイオインフォマティクスを学ぼうとする初学者のみなさんが必要とするバイオインフォマティクスのスキルを身につけるための方法を知り、独習できるようになりますことを目指します。[講義スライドPDF](#)は[こちら](#)からご覧いただけます。

動画ファイルのダウンロード https://2020.12.25_informatics.togotv.mov

DOI <https://doi.org/10.7875/togotv.2020.096>

再利用時のライセンス クリエイティブ・コモンズ CC-BY-4.0

関連動画

- TOGO TV コード、データ、計算環境をまとめて公開するケースCode Oceanの紹介
- TOGO TV draw.ioを使ってペン図とフローチャートを作成する
- TOGO TV ゲームデータベースツールを使ってGN5データの構造と部分GN5データを理解する
- TOGO TV 関連動画



生命科学研究のためのGitHubの使い方 基本編

<https://togotv.dbcls.jp/20180621.html>



DBCLS Research Services Contact About

TOGO TV

トップページ TogoTVについて 動画を探す 画像を探す 講習会資料を探す お問合せ キーワードから動画を探す

生命科学研究のための GitHub の使い方 基本編

180620 版

この動画が再生されない場合は、YouTubeでご覧ください。 <https://youtu.be/bNAaG0NNCCk>

2018.06.21 07:17 作者: hiuchi 編集: 小野 浩雅(ONO, Hiromasa)

生命科学研究のためのGitHubの使い方 基本編

GitHubは、"Git"というプロジェクトの改訂履歴を管理し保存するバージョン管理システムを利用するためのハブ(Hub)がその名の由来で、コンピュータプログラムやソースコードの公開・共有・管理に利用されているウェブサービスです。

一方で、コンピュータプログラムに限らず、Word文書や画像など、どんな種類のドキュメントでも、すべての段階の草案やアップデート履歴を保存し管理することができます。GitHubは、グラフィカルなユーザインターフェースを提供しており、コマンドラインを使わずにウェブサイト上で様々な作業をすることができます。また、アクセス制御機能や、タスク管理機能など、コラボレーションのための機能を多く備えているのが特長です。

見どころダイジェスト

- 00:11 1. GitHub の紹介
- 00:53 2. アカウントを作成する
- 01:51 3. リポジトリを作成する
- 02:50 4. コミットする
- 04:23 5. ブランチを作成する
- 05:23 6. ファイルをダウンロードする
- 05:39 7. プルリクエストを送信し、マージする

この動画のタグ

ウェブツール 環境設定

動画ファイルのダウンロード
180621_GitHub_basic.mov

DOI

<https://doi.org/10.7875/togotv.2018.114>

再利用時のライセンス

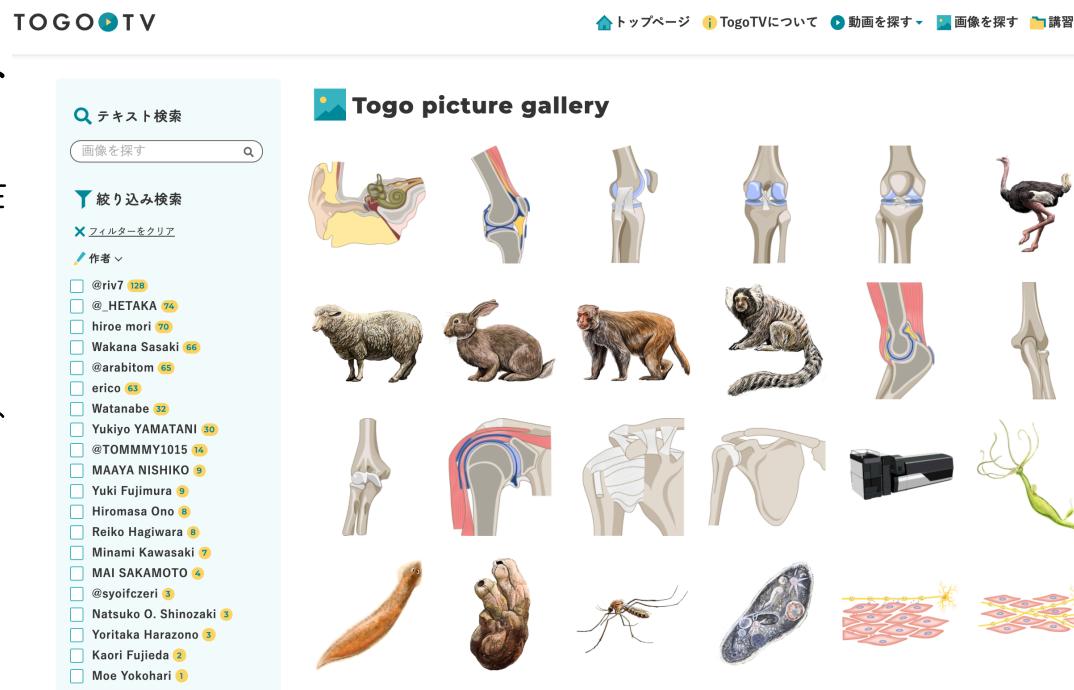
クリエイティブ・コモンズ CC-BY-4.0 ?



Licensed under CC-BY 4.0 © 2021 Hiromasa Ono (DBCLS)

画像を探す <https://togotv.dbcls.jp/pics.html>

- Togo picture gallery
 - 生命科学分野のイラストが、誰でも自由に利用可能 (CC-BY-4.0)
 - 研究発表のスライド・ポスター作成、資料作成等に、ぜひご活用ください
 - 統合TVのコンテンツを再利用したいのですが、著作権の扱いはどうなっていますか?
 - <https://togotv.dbcls.jp/faq.html>
 - 2021年9月末で48件の引用論文



画像のご利用にあたっては、出典元 (© 2016 DBCLS TogoTV) あるいは該当ページのURLを必ず明記してください。出典の記載のみが条件で、論文の図表や発表スライド、ポスター、プレスリリース、販促資料など内容を問わず、自由にご利用いただけます。出典元の記載箇所については、デザインの体裁を損なわない箇所・表現で構いません。[出典元表示用の画像も用意](#)しておりますので、目的に応じて適宜ご自由にご利用ください。論文等 ([引用例](#)) では、Acknowledgementや図のlegend等にご記載いただいています。(例: "The image of XX is from TogoTV (© 2016 DBCLS TogoTV).")

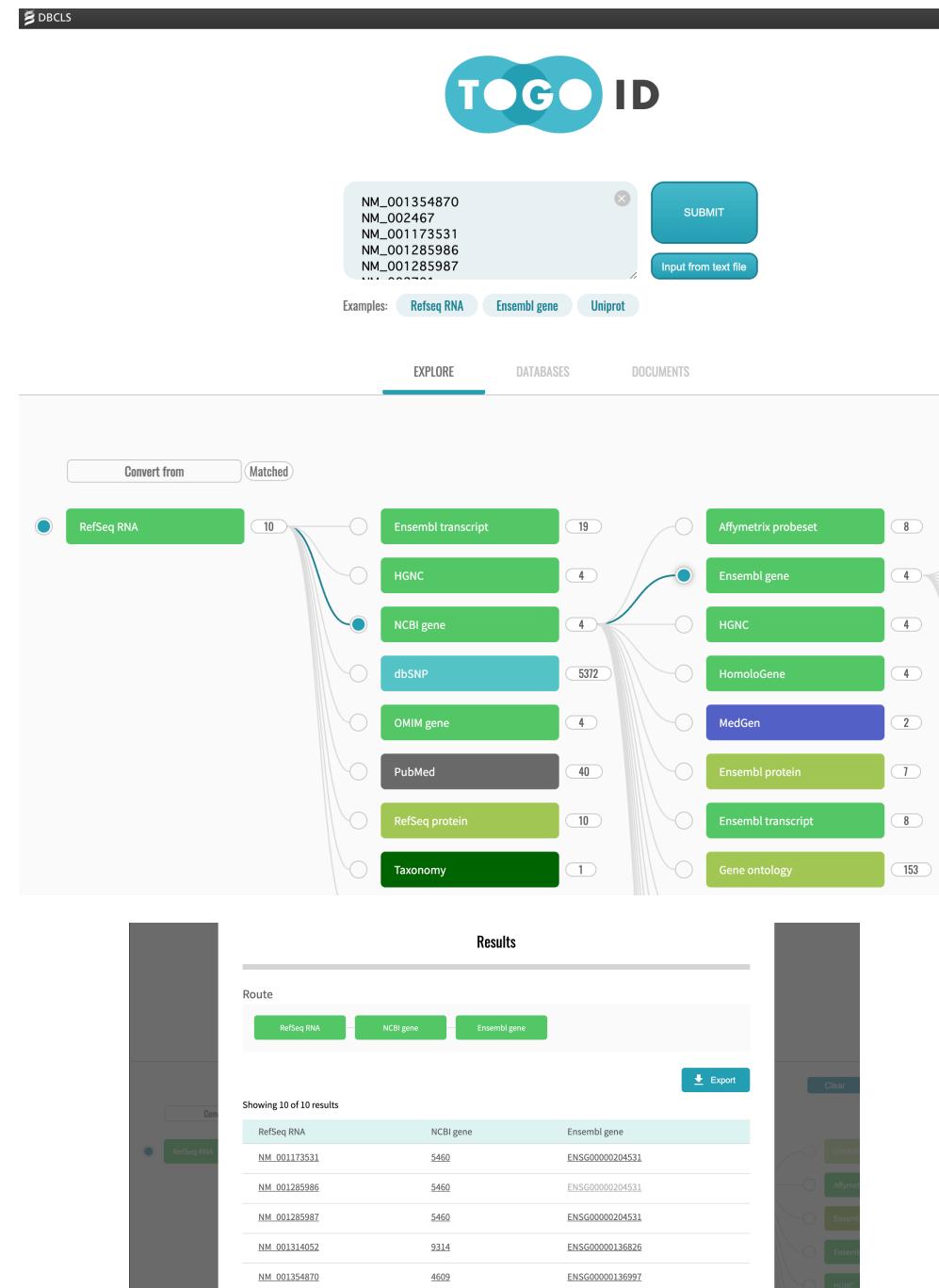
Google Spreadsheet + GitHub actions

A	B	C	D	E	F	G	H	I	J
	To go TV _Video_ID	id	name	url	uploadDate	description	keywords	embedUrl	duration(ISO 8601)
1									
2		1920	https://doi.org/10.7875/togotv.2021.068	[英字幕有/EN sub] TogоДを使って生命科学分野のデータベースを検索する @ AJACCS	https://togotv.dbcls.jp/20210910.html	2021-09-10	ウェブツール,情報収集,ゲノム,遺伝子,xxkVEtJMW2k ウェブインターフェイスだけでなく、<a href= TogoID has a function to copy the converted II		
3		1919	https://doi.org/10.7875/togotv.2021.067	ウェブツールを使ってゲノム編集の標的サイトを検索する @ AJACCS	https://togotv.dbcls.jp/20210831.html	2021-08-31	本日の統合TVは、2021年8月19日に開催され ウェブツール,塩基配列,ゲノム,CRISPR,x-v50Deb6NQ 本講習では、だれでも無償かつ自由に使うこ		PT7M53S
4		1918	https://doi.org/10.7875/togotv.2021.066	gnomAD v2.1.1 を使ってヒトバリアントの集団中のアレル頻度を	https://togotv.dbcls.jp/20210823.html	2021-08-23	ゲノム,遺伝子,構造解析,バリアント解釈,XGGSBs-fGo 重篤な小児病を持っている患者およびその近親者		PT11M27S
5		1917	https://doi.org/10.7875/togotv.2021.065	draw.io を使ってベン図とフローチャートを作成する	https://togotv.dbcls.jp/20210822.html	2021-08-22	環境設定,可視化,ウェブツール,情報収集,9aIJWrEze_M		PT7M57S
6		1916	https://doi.org/10.7875/togotv.2021.064	クリエイティブ・コモンズ・ライセンスの付与されたコンテンツ	https://togotv.dbcls.jp/20210821.html	2021-08-21	情報収集,クリエイティブ・コモンズ,rLWwiBxTEa4 今回は、CCライセンスの付与されたコンテン		PT3M15S
7		1915	https://doi.org/10.7875/togotv.2021.063	トランスマキス解析のための多階層パスウェイ可視化の自動化	https://togotv.dbcls.jp/20210806.html	2021-08-06	本日の統合TVは、2021年6月10日(金)にオンラインで講演,メタボローム VAyjQNTr_OI		PT14M27S
8		1914	https://doi.org/10.7875/togotv.2021.062	MetaboAnalystを用いた主成分分析とエンリッチメント解析 @ 第	https://togotv.dbcls.jp/20210805.html	2021-08-05	本日の統合TVは、2021年6月10日(金)にオンラインで講演,メタボローム vePMd9xEdV8		PT15M50S
9		1913	https://doi.org/10.7875/togotv.2021.061	質量分析インフォマティクスによる未知化合物のアノテーション	https://togotv.dbcls.jp/20210804.html	2021-08-04	本日の統合TVは、2021年6月10日(金)にオンラインで講演,メタボローム Qcoqfud-LSI		PT24M54S
10		1912	https://doi.org/10.7875/togotv.2021.060	メタボロミクスに資する質量分析データ処理プロトコール @ 第	https://togotv.dbcls.jp/20210803.html	2021-08-03	本日の統合TVは、2021年6月10日(金)にオンラインで講演,メタボローム p_jecCwE0yY		PT20M53S
11		1911	https://doi.org/10.7875/togotv.2021.059	定量メタボロームデータ取得に向けた分析技術およびソフトウエ	https://togotv.dbcls.jp/20210802.html	2021-08-02	本日の統合TVは、2021年6月10日(金)にオンラインで講演,メタボローム 9K63qNY4uGU		PT37M43S
12		1910	https://doi.org/10.7875/togotv.2021.058	脂肪細胞におけるインスリン依存性糖代謝の速度論的トランスポ	https://togotv.dbcls.jp/20210801.html	2021-08-01	本日の統合TVは、2021年6月10日(金)にオンラインで講演,メタボローム Do0bn6nzn_U		PT36M21S
13		1909	https://doi.org/10.7875/togotv.2021.057	次世代シーケンスデータベースを使って公開NGSデータの検索と	https://togotv.dbcls.jp/20210730.html	2021-07-30	本日の統合TVは、2021年7月15日に開催され 講演,AJACS,NGS,データベース,ウェブ,zM9AQ8AN-oA 本講習では、NGSデータベースの種類とSRA		PT1H40M04S
14		1908	https://doi.org/10.7875/togotv.2021.056	ゲノムデータベース/ゲノムプラウザを使って配列解析と遺伝子	https://togotv.dbcls.jp/20210729.html	2021-07-29	本日の統合TVは、2021年7月15日に開催され 講演,AJACS,ゲノム,塩基配列,遺伝子,50fkJbWYiaoA 本講習では、ゲノムデータベースとゲノムア		PT1H35M08S
15		1907	https://doi.org/10.7875/togotv.2021.055	protocols.io を使って公開プロトコールと公開ワークスペースを	https://togotv.dbcls.jp/20210728.html	2021-07-28	ウェブツール,情報収集,hjiNc09lVc0		PT5M02S
16		1906	https://doi.org/10.7875/togotv.2021.054	protocols.ioを使って研究プロトコールを作成/共有する	https://togotv.dbcls.jp/20210727.html	2021-07-27	ウェブツール,情報収集,9TstPox8kzg		PT8M35S
17		1905	https://doi.org/10.7875/togotv.2021.053	PubMedをフル活用して論文執筆を効率化する @ AJACSオンライン	https://togotv.dbcls.jp/20210715.html	2021-07-15	本日の統合TVは、2021年6月15日に開催され 講演,AJACS,文献検索,文章執筆,辞書 本講習では、PubMedの使い方が身につき、Ex_X5TRrQeo		PT1H33M47S
18		1904	https://doi.org/10.7875/togotv.2021.052	NBDCのサービスを使って基本的なDBの使い方を知る @ AJACS	https://togotv.dbcls.jp/20210714.html	2021-07-14	本日の統合TVは、2021年6月15日に開催され 講演,AJACS,NBDC,データベース 本講習では、NBDCのサービスを使って、研ZeUHgwUORsc		PT1H30M22S
19		1903	https://doi.org/10.7875/togotv.2021.051	UCSC Table Browserを使って多様なアノテーショントラックを	https://togotv.dbcls.jp/20210707.html	2021-07-07	ゲノム,配列解析,UCSC,遺伝子,発現解析 今回の動画では、Table Browserを用いる例と		PT8M58S
20		1902	https://doi.org/10.7875/togotv.2021.050	RNA-seqを使ってウェブブラウザ上でRNA-seqデータを解析す	https://togotv.dbcls.jp/20210531.html	2021-05-31	発現解析,遺伝子,NGS,RNA-seq,可視化,tVnGIQvfws FASTQファイルの定量化、品質チェック、遺伝子の可視化 今回は、ヒトの培養細胞(コントロール)とH2O2を晒した培養細胞のRNA-seqデータ(Stu		PT11M24S
		1901	https://doi.org/10.7875/togotv.2021.049	AnnotationHub によるR/Bioconductorアノテーションパッケージ	https://togotv.dbcls.jp/20210522.html	2021-05-22	本日の統合TVは、2021年3月10日に開催され 講演,プログラミング,環境設定,情報収集,3ZcpVSHThzs 本講習では、R/Bioconductorでの最新のアノテーションパッケージ R言語で作成したライフサイエンス系のパッケージ 今後データをBioconductorに登録する人は、この講演動画は、<a href="https://sites.google.com		



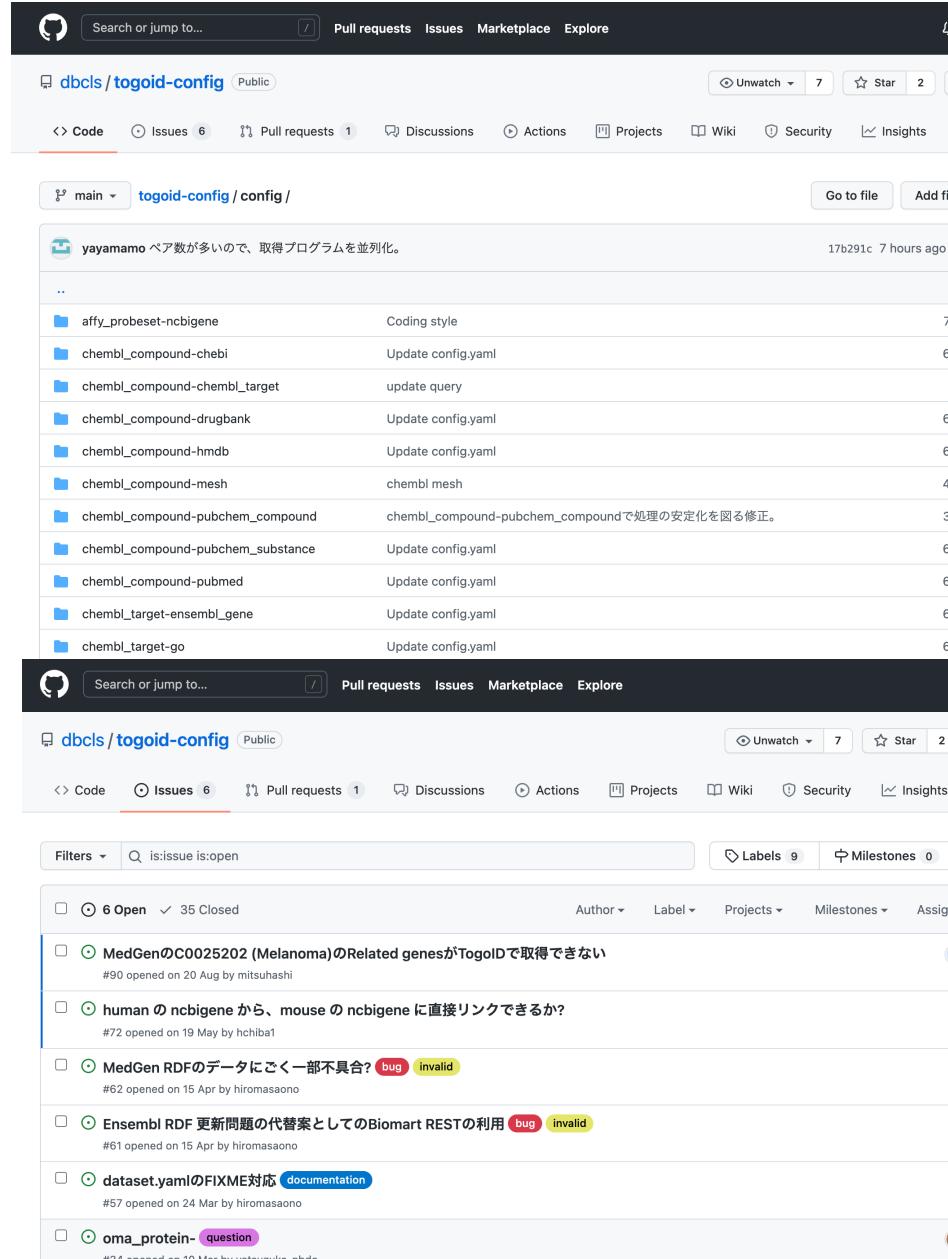
Togoid <https://togoid.dbcls.jp/>

- データベース統合の基盤となるID変換サービス
- バイオインフォマティクスで様々なDBを活用するには**データベースID間のリンク(対応関係)**が重要
 - 等価なものに付けられたID間の変換
 - 遺伝子DB1↔遺伝子DB2
 - 関連する情報の取得
 - バリアント→遺伝子
 - 遺伝子→トランスクリプト
 - トランスクリプト→タンパク質
 - タンパク質→立体構造
 - 立体構造→相互作用
 - 相互作用→化合物・医薬品
 - 化合物・医薬品→パスウェイ
 - パスウェイ→疾患
- ウェブ上で対話的に操作して変換し結果をダウンロード
 - 始点となるIDから探索的に接続先のDBをたどる



togoid-config <https://github.com/dbcls/togoid-config/>

- TogоДデータ取得・更新の自動化
 - 変換元DBと変換先DBのペア毎に、IDの対応関係を抽出するプログラムを作成
 - 63DB、156DBペア(2021年9月末現在)
- Rakefile 自動更新手順
- bin/ 各種取得・変換スクリプト群
- config/ db1-db2ごとの変換規則群
 - dataset.yaml データベース一覧
 - db1-db2/config.yaml 更新手順
- input/ 共通の前処理入力データ置き場
- output/ 生成される出力IDペア置き場
 - tsv/db1-db2.tsv タブ区切りファイル
 - ttl/db1-db2.ttl RDF版ファイル
- 参考: Issuesで議論、バグ報告、タスク管理をする



The screenshot shows two views of the GitHub repository dbcls/togoid-config.

Top View: Shows the 'config' directory listing various configuration files for different databases. The files include: affy_probeset-ncbigene, chembl_compound-chebi, chembl_compound-chembl_target, chembl_compound-drugbank, chembl_compound-hmdb, chembl_compound-mesh, chembl_compound-pubchem_compound, chembl_compound-pubchem_substance, chembl_compound-pubmed, chembl_target-ensembl_gene, and chembl_target-go. Each file has a brief description and a timestamp indicating it was last updated 7 hours ago.

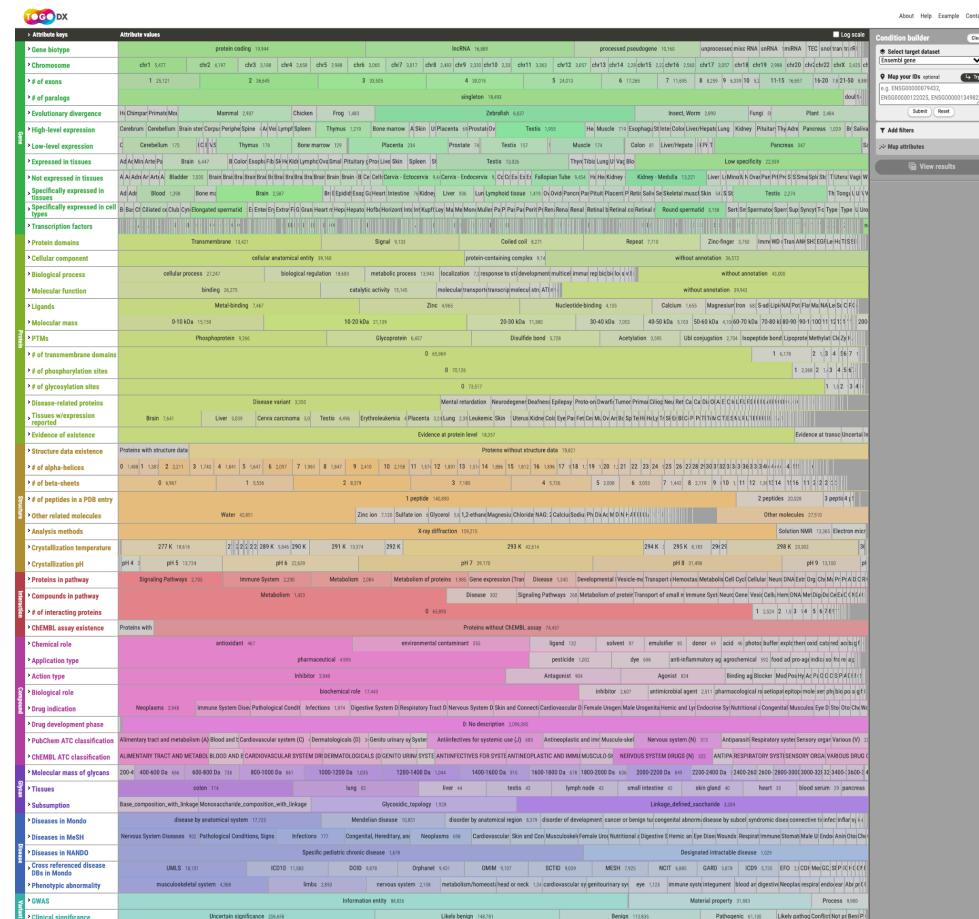
Bottom View: Shows the 'Issues' tab with 6 open issues. The issues are:

- MedGenのC0025202 (Melanoma)のRelated genesがTogоДで取得できない (#90) - opened on 20 Aug by mitsuhashi
- human の ncbigene から、mouse の ncbigene に直接リンクできるか? (#72) - opened on 19 May by hchiba1
- MedGen RDFのデータにごく一部不具合? (#62) - opened on 15 Apr by hiromasaono
- Ensembl RDF 更新問題の代替案としてのBiomaRt RESTの利用 (#61) - opened on 15 Apr by hiromasaono
- dataset.yamlのFIXME対応 (#57) - opened on 24 Mar by hiromasaono
- oma_protein- question (#33) - opened on 10 Mar by yateiuruka-nbdc



TogoDX/human <https://togodx.dbcls.jp/human/>

- 国内外のデータベースから収集・統合した、ヒトに関する遺伝子、タンパク質、化合物、疾患などの情報をワンストップで探索することができるサービス
 - 2021年10月公開



The screenshot shows the TogoDX/human search results page. At the top, there are several dropdown menus for filtering results by target dataset (e.g., ENSEMBL, HGNC), species (e.g., Human, Mouse, Zebrafish), and specific attributes like 'Protein domain' or 'Cellular component'. Below these are two main sections of tables:

- Attribute keys**: A table showing the count of entries for various attributes such as Gene biotype, Chromosome, # of exons, # of paralogs, Evolutionary divergence, High-level expression, Low-level expression, Expressed in tissues, Not expressed in tissues, Specifically expressed in tissues, Specifically expressed in cell types, Transcription factors, Proteins domains, Cellular component, Biological process, Molecular function, Ligands, and Molecular mass.
- Attribute values**: A large table showing the count of entries for specific values under each attribute key, such as protein-coding (13344), chromosome 1 (1,312), exon 1 (1), and so on, up to molecular mass 100 kDa (1,139).

- TogoDX(Data eXplorer) は、生命科学分野における様々なデータベースを統合的に探し、俯瞰するためのフレームワークです。膨大な情報を多様な属性(attribute)によって柔軟に絞り込み、必要な情報を抽出できる新しい仕組み

- TogoDX/humanでは、約20個のデータベースに由来する50個以上のattributeが利用可能

TogoDX/human の使い方

- ①肺でタンパク質として発現が確認され、②細胞膜表面に局在し、③タンパク質立体構造が明らかになっており、④対応する医薬品が開発されている、ヒトのタンパク質の一覧を取得する
 - 最終的に出力したいキーとなる dataset を **Select target dataset** で決める
 - UniProt
 - 調べたい属性(attribute)を **Add filters** で選び、絞り込む
 - ① Protein / Tissues w/expression reported / Lung
 - ② Protein / Cellular component / cellular anatomical entity / membrane / plasma membrane
 - ③ Structure / Structure data existence / Proteins with structure data
 - ④ Interaction / ChEMBL assay existence / Conf-score 9: Direct single protein target assigned
 - **Map attributes**機能を使って他の属性での分布を見る
 - ヒト遺伝子が保存されている最も遠縁の生物別に分類する
 - GTEx(V6)で測定された49組織中で、組織特異的に高発現する遺伝子を組織別に分類する
 - UniProtの各エントリーに付与されているタンパク質と関連する疾患で分類する

GitHub無しでは不可能だった(?) TogоД, TogоДX の開発



- DBCLS/NBDCメンバー(30数名) + 発注業者さん(10数名)
 - アプリケーションUI開発
 - アプリケーションが受け取るデータの仕様決め・設定・整備
 - データの取得、表示、連携
 - ドキュメント作成 (About, Help, Examples…)
 - デバッグ
- GitHubの恩恵
 - バージョン管理
 - ソースコードのバックアップ・更新差分の可視化(画像も可能)
 - アクセス制御
 - 編集後の更新時のレビューを必須に
 - ブランチ操作
 - 開発(実験)環境と公開環境
- GitHub + Slack + Zoom + Google Docs
 - バージョン管理ツール + チャット + オンラインミーティング
 - SlackにGitHubの更新通知
 - Zoomで議論・共同作業
 - Google Docsで議事録作成、データの整理



GitHubの最初の一歩

- 馴染みのない人には、最初はかなりとっつきにくい
 - Git独特の用語や操作
 - リポジトリ、コミット、プッシュ、クローン、マージ、ブランチ、プルリクエスト…
 - ちょっと詳しい人が近くにいると始めやすい(聞きやすい)
 - まずアカウントを作ってみる
 - 自分のファイルのバージョン管理などから始めてみる
- GUIツールや教材などは充実している
 - コマンドライン(黒い画面)でやらないといけないことはない
 - 公式ウェブサイト版だけでもかなり便利に使える
 - 【2021年版】ノンエンジニアにおすすめのGit GUIクライアント5選
 - <https://aizine.ai/git-gui-0201/>
 - Hiroshi WatanabeさんによるGitHub演習
 - <https://github.com/kaityo256/github>
 - CC-BY 4.0で提供されている

公募にもGitHubアカウント

<https://twitter.com/tonets/status/1448093299068211200>

大上雅史 | Ohue M
@tonets

大学の公募で必要書類に「githubのアカウント名」は初めて見たかもしれない。



待遇 Compensation	[基本年俸] 特任研究員（常勤）10,347,600円～16,908,600 (12分の1の額を月額基本給として毎月支給) 通勤手当（住居手当、扶養手当、退職手当及び賞与は支給しない) *今回はスキルを重視した人材の獲得が目的であるため、「特任研究員」という肩書ではありますが、給与は特任助教、特任准教授、よりも高いランク、1000万以上を想定しています。 [その他] 上記の他の労働条件については国立大学法人大阪大学任期付教職員就業規則等によります。 https://www.osaka-u.ac.jp/ja/guide/information/joho/kitei_shugyou.html 以上の労働条件等については、本採用情報掲載時点のものであり、変更の可能性があります。※敷地内原則禁煙
募集期間 Application period	2022年03月31日 必着 適任者の採用が決まり次第、募集を締め切ります

takanzai @takanzai · 8時間 jrecin.jst.go.jp/seek/SeekJorDe... 大阪大学量子ソフトウェア研究拠点、量子ソフトウェアエンジニア募集（年俸1000万～） 量子コンピュータに自身の知見やスキルが少しでも活かせると思われる方、量子コンピューティング分野に新規に取り組みたい方 特任研究員（常勤）10,347,600円～16,908,600 このスレッドを表示



Licensed under CC-BY 4.0 © 2021 Hiromasa Ono (DBCLS)

バイオインフォマティクス分野の アプリケーション開発におけるGitHubの活用とその実際



仮説構築 から (実験、データ解析、) 論文執筆 まで
生命科学系研究者の暮らしをささえます



<https://dbcls.rois.ac.jp/>



Licensed under CC-BY 4.0 © 2021 Hiromasa Ono (DBCLS)

