

○大波 純一¹、阪口 幸治²、有田 正規³、山地 一禎¹

1. 国立情報学研究所オープンサイエンス基盤研究センター(RCOS)
2. 国立情報学研究所学術基盤推進部
3. 国立遺伝学研究所生命情報・DDBJセンター

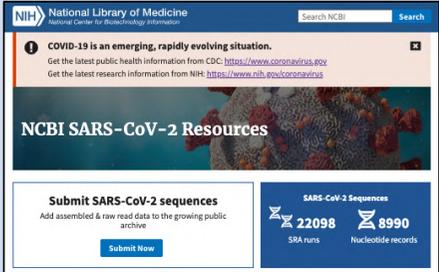
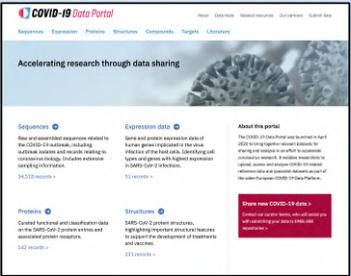
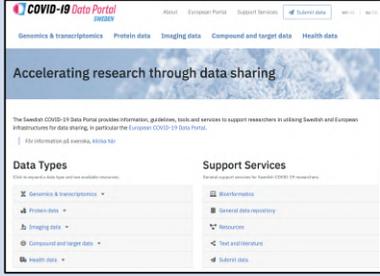
■概要

新型コロナウイルス感染症（COVID-19）の感染拡大に伴い、研究データ共有の重要性が高まりつつある。欧州では早期からCOVID-19の研究データに関するポータルサイト（COVID-19 Data Portal）を、EMBL-EBIが中心となって構築している。スウェーデンでもこの枠組みに倣い、COVID-19 Data Portal Swedenを公開している。一方、日本のCOVID-19に関するポータルサイトのほとんどは市民向けの情報を集約したもので、研究者が必要とするデータは機関ごとに散在していた。

今回、国立情報学研究所オープンサイエンス基盤研究センター（RCOS）でも欧州の枠組みに賛同し、国内の研究機関と協力して「COVID-19 データポータルJAPAN」を構築した。本サイトでは、研究者が国内外のCOVID-19の研究情報に迅速にアクセスできるよう、適切な分類とリンク先の選定、解説の付与を実施した。本サイトは日本初の組織横断的なCOVID-19に関する研究情報のポータルサイトとして、幅広い研究者に利用されることを目指し、コロナ禍の早期解決に繋げていく。

■ 海外・国際組織のCOVID-19研究データ共有

COVID-19の研究データを国家や国際組織ごとに公開し、研究の推進に繋げることが世界的な潮流となりつつある。

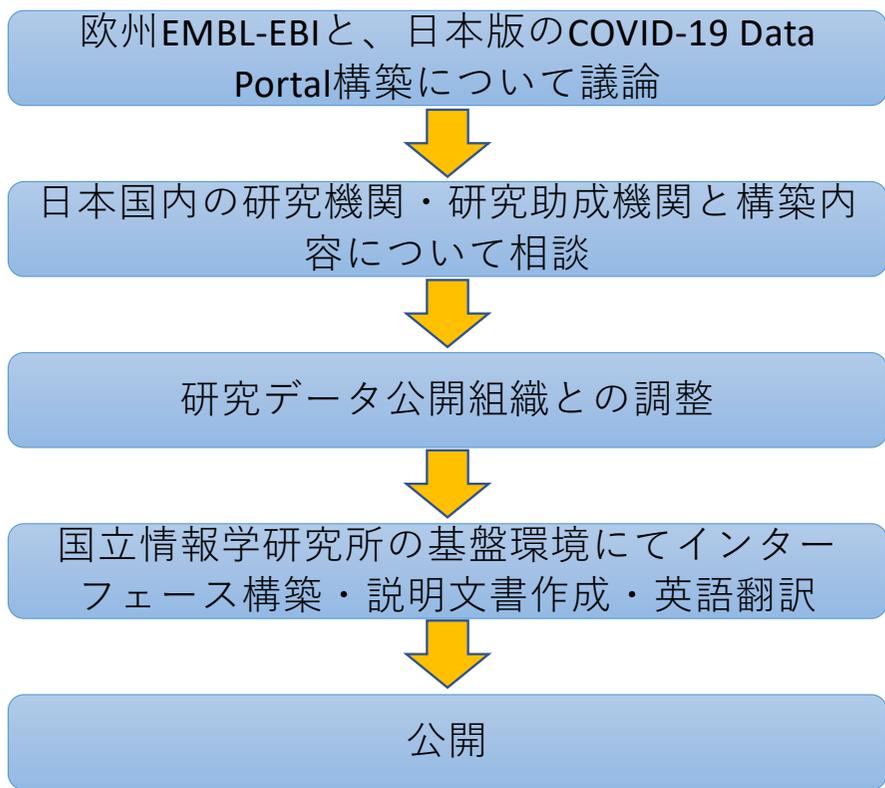
米国 NCBI SARS-CoV-2 Resources	欧州 COVID-19 Data Portal	スウェーデン COVID-19 Data Portal SWEDEN
 <p>https://www.ncbi.nlm.nih.gov/sars-cov-2/</p>	 <p>https://www.ncbi.nlm.nih.gov/sars-cov-2/</p>	 <p>https://www.covid19dataportal.se/</p>
<p>米国内のCOVID-19の研究結果を一箇所に集約し全てオープンに。迅速に閲覧可能。</p>	<p>欧州EMBL-EBIが主導し、欧州加盟国の研究データを公開。</p>	<p>欧州のCOVID-19 Data Portalに賛同し、スウェーデン国内のCOVID-19研究データを集約。</p>

国内の研究データ共有

機関ごとに特設ページの設置や、研究データの公開を実施しているが、ポータルなどで組織横断的にCOVID-19の研究データを公開している例は無い。

遺伝研DDBJセンター	国立感染症研究所	大阪大学 蛋白質研究所	国立医薬品食品 衛生研究所	JST / NBDC
 <p>INSDC Statement on SARS-CoV-2 sequence data sharing during COVID-19</p> <p>http://www.insdc.org/documents</p>	 <p>https://www.niid.go.jp/niid/ja/diseases/ka/corona-virus/2019-ncov.html</p>	 <p>https://pdbj.org/featured/covid-19</p>	 <p>https://www.nih.gov/sars-cov-2/index.html</p>	 <p>https://biosciencedbc.jp/blog/20200303-01.html</p>
<p>米国・欧州と共同で、新型コロナウイルス関連の配列情報公開に関する共同声明を発出。</p>	<p>特設ページにて各種レポート、感染者の統計情報、検査方法、コロナウイルスの電顕写真を公開。ゲノム配列等はこのサイトから公開されていない。</p>	<p>特設ページにて、構造情報を104件表示。</p>	<p>医薬品や検査情報についてまとめ、特設ページから公開。</p>	<p>日本国内の研究データをまとめ、ブログで公開。更新継続中。</p>

「COVID-19データポータルJAPAN」の開発手順と画面



研究ツール（バイオインフォマティクス）

研究データ（遺伝子・配列情報）

研究データへ直接アクセスできるハイパーリンク、詳細な説明文を設置。

研究データを利用するためのツール・基盤情報も掲載。

研究ツール（データの投稿）

研究者が研究データを投稿するための基盤情報 →さらなるデータ公開の循環へ繋げる

トップページ <https://covid19dataportal.jp>



■リンクデータ一覧

13機関、25種類のコンテンツを掲載し、特に国内のCOVID-19研究データを網羅的に集約した。
COVID-19のデータ駆動型研究が推進できるよう、今後もデータの追加やコンテンツの拡大を実施していく。

提供機関	コンテンツタイトル	概要（件数は2020年7月時点）
ライフサイエンス統合データベースセンター	DBCLS Service	バイオインフォマティクスやデータベース統合に役立つツールやサービスの一覧
国立遺伝学研究所 DDBJセンター	DDBJ ARSA	10082件の塩基配列
	DDBJ SRA	71件の次世代シーケンサによる解析
	DDBJ 検索・解析	配列データの検索や解析ツール
	DDBJ 登録	三大国際データバンクのひとつの登録ページ
国立医薬品食品衛生研究所	新型コロナウイルスに関する情報提供のページ	新型コロナウイルスに関する医薬品・PCR検査・抗体検査・食品安全に関する情報を提供
国立感染症研究所	感染研 画像・映像アーカイブ	コロナウイルスの電顕写真など
国立情報学研究所	CiNii Articles	1504件の文献情報を新しい順に。
	CiNii Books	41件の書籍情報を新しい順に。
	JAIRO Cloud	クラウド型の機関リポジトリ環境提供サービス
国立国会図書館	国立国会図書館 リサーチ・ナビ	13件のCOVID-19に関する調べ方とコンテンツ一覧のページ
日本学術会議	Information on important infectious diseases including COVID-19 in Japan	政府や自治体から出された統計情報や提言について公開

提供機関	コンテンツタイトル	概要（件数は2020年7月時点）
京都大学化学研究所 附属バイオインフォマティクスセンター	KEGG Virus	11件の部分配列情報
	KEGG Genome	1件のゲノム配列全長情報
	KEGG Disease	新型コロナウイルス感染症の定義情報
	KEGG: Kyoto Encyclopedia of Genes and Genomes	KEGGデータベースシステム全体の統括ページ
科学技術振興機構 バイオサイエンスデータベースセンター	TogoProt	数十件の蛋白質DBの統合検索
	NBDCブログ	COVID-19に関するコンテンツ一覧を記したブログ記事
	NBDC	データベースのカタログ、横断検索、アーカイブなど
	LSDB Archive	国内のライフサイエンス研究者が生み出したデータセットのアーカイブ
大阪大学蛋白質研究所	PDBj: Protein Data Bank Japan	104件の代表構造と解説
米国国立衛生研究所	PubMed	31054件の文献情報。
	PubMed Central: PMC	米国NIHの収集するCOVID-19関連のOAジャーナルへのリンク。
UniProt Consortium	UniProt: The Universal Protein Resource	53件のタンパク質データ
Europe PMC funder group	Europe PMC	114401件のOAジャーナルとプレプリント