

大学・研究機関等におけるデータ共有について

参考資料 2
情報委員会（第11回）
令和2年8月21日

機関・プロジェクト名	公開データの種別 (適用範囲)	オープン・クローズの考え方 (制限段階)	公開のタイミング (公開までの猶予期間)	新型コロナウイルス感染症対策で 講じた措置・活用事例
東京大学地震研究所	関東甲信越地域、紀伊半島とその周辺地域、瀬戸内海西部とその周辺地域の地震活動地図、地震波形データ、及び国立大学微小地震観測網データを用いた震源カタログ。 【データベース】 ・地震活動に関するデータベース	①オープン(公開) ②セミオープン(制限公開) 地震波形データの内、オンライン提供期間以外のデータについては、共同研究として提供可能。 の2分類	データベースに登録したものを即時公開。	—
大阪大学蛋白質研究所・Protein Data Bank Japan(PDBj)	生体高分子の立体構造に関するデータ。 【データベース】 ・Protein Data Bank Japan(PDBj)	①オープン(公開)	データバンク方式であり、データベースへの登録者が登録時に下記の中から公開時期を選択する。 1 データベースに登録したら即時公開。 2 論文発表(bioRxiv等のPublic Preprintを含む)があったら公開(最長1年) 3 期日を指定(登録から4週間後、6週間後、8週間後、6か月後、1年後)して公開。	・新型コロナウイルス感染症(COVID-19)に関連した蛋白質構造情報データベース(PDB)をインターネット上で一般に3/11に公開した(毎週水曜日9:00amに世界同時更新)。 ・通常、PDBデータは、該当の情報を含む研究が論文として発表されるまで非公開とされるが、新型コロナウイルスの構造情報がデータベースに登録された場合、個別に登録者(研究者)と連絡をとり、論文発表を待たずに即時公開することを強く勧めている。 ・PDBに登録されたウイルスタンパク質(メインプロテアーゼ、RNA依存型RNAポリメラーゼなど)の3次元構造をベースにして、結合する低分子化合物(既存薬)をインシリコスクリーニングにて探索し、ヒット化合物を得た。(BINDSインシリコユニット)
一橋大学経済研究所	近代日本経済の歴史統計を経済活動の諸分野にわたって推計、加工して体系的に集めた統計書である『長期経済統計』全14巻(1965-1988年)のデータ。 【データベース】 ・(日本)長期経済統計データベース(LTES)	①オープン(公開) 引用に際しては、その出所の明記を依頼している。	公開済み。	—

機関・プロジェクト名	公開データの種別 (適用範囲)	オープン・クローズの考え方 (制限段階)	公開のタイミング (公開までの猶予期間)	新型コロナウイルス感染症対策で 講じた措置・活用事例
人間文化研究機構国文学研究資料館	国文学研究資料館が大型プロジェクト、調査研究等の活動により作成する日本古典籍ナショナルカタログと、国文研及び国内外の大学等が所蔵する日本文学を含む広範な分野・領域の古典籍の全冊画像のデータベース。 【データベース】 ・日本古典籍総合目録データベース(目録情報に主眼) ・新日本古典籍総合データベース(画像閲覧に特化)	①オープン(公開) 国文研所蔵古典籍画像データはCC BY-SAの条件で公開。国文研以外の機関が所蔵する古典籍については、転載などの利用に関しては各所蔵者ごとに条件を示し公開。また、国文研所蔵古典籍画像データの一部は、連携する情報・システム研究機構 人文学オープンデータ共同利用センター(CODH)からもCC BY-SAの条件でデータセットとして公開。	公開用データベースに登録したもののから即時公開。	—
自然科学研究機構国立天文台	国立天文台及び大学の観測装置で共同利用等により取得した天文観測データ。 【データベース】 ・SMOKA ・すばる望遠鏡アーカイブシステム(STARS) など	①オープン(公開) ②セミオープン(制限公開) SMOKAについて、観測データを取得する場合はユーザ登録が必要。 ③セミクローズ(制限共有) STARSについては、共同利用者に対して観測データを共有。 の3分類	観測装置によって、観測者の占有期間1~2年間を設定し、その後、当該観測データを公開。	アルマ望遠鏡の観測データアーカイブ(セミクローズ)について、新型コロナウイルス感染症によるユーザーの研究環境への影響を考慮して、観測者のデータ占有期間を1年から1年3ヶ月へ延長した。
情報・システム研究機構 国立遺伝学研究所	DNA塩基配列と機能注釈データ。 【データベース】 ・DNA Data Bank of Japan(DDBJ) ・Japanese Genotype-phenotype Archive(JGA) ・AMED Genome Group Sharing Database(AGD)	①オープン(公開) ② セミオープン(制限公開) JGAについてはデータ利用のためにJST NBDCの審査が必要。 ③セミクローズ(制限共有) AGDのデータは利用者間で共有。AGD利用についてはNBDCの審査等を経る。 の3分類	登録者は公開予定日を設定し延長することができる。	3極連携共同でコロナウイルスゲノムのオープン化を要求する声明を発表予定。
海洋研究開発機構 (JAMSTEC)	調査観測データ、シミュレーションデータ、画像、図面等の機構が保有する研究設備、施設を使った調査研究で得られた各種データ及びサンプル情報。計75のデータセットを http://www.godac.jamstec.go.jp/catalog/data_catalog/metadataList?lang=ja で公開。 【データベース】 ・航海・潜航データ・サンプル探索システム(DARWIN) ・深海映像・画像アーカイブ(J-EDI) ・地震研究情報データ提供システム(J-SEIS) など	①オープン(公開) 公開猶予期間後に公開。調査実施時の各種条件・取決め・契約等がある場合は個別対応。 ②クローズ(非公開) 個別に判断(事故調査に協力した映像や画像などのデータ) の2分類	取得した研究者が優先的に利用できる期間(公開猶予期間:原則2年、ただし共同利用については5年)を設定し、経過後に公開。航海に関する情報や航海報告等は2ヶ月で公開、定常的に取得する観測項目については準備でき次第公開。 なお、調査実施時の各種条件・取決め・契約等がある場合は個別対応。	—

機関・プロジェクト名	公開データの種別 (適用範囲)	オープン・クローズの考え方 (制限段階)	公開のタイミング (公開までの猶予期間)	新型コロナウイルス感染症対策で 講じた措置・活用事例
物質・材料研究機構 (NIMS)	無機材料や高分子材料、金属構造材料の物質構造等、NIMSの研究活動で得られたデータ、論文等から収集したデータに加え、論文等で発表した公知データをメタデータ(試料・装置・手法などの情報)を付与した形で収録したもの。 【データベース】 ・物質・材料データベース(MatNavi) 【リポジトリ】 ・材料データリポジトリ(Materials Data Repository: MDR)	①オープン(公開) ②セミオープン(制限公開) MatNaviの利用にはユーザ登録が必要。 の2分類	データベース等に登録したのから即時公開。	—
防災科学技術研究所 (NIED)	調査観測実験データ、シミュレーションデータ、画像、図面等の調査研究で得られた各種情報及びそれらを記録したもの 【データベース】 ・高感度地震観測網(Hi-net) ・基盤的火山観測網(V-net) ・日本海溝海底地震津波観測網(S-net)など	①オープン(公開) ②セミオープン(制限公開) Hi-net、V-net、S-netについて、データを取得する場合にはユーザ登録が必要。 の2分類	データベースに登録したのから即時公開。	—
宇宙航空研究開発機構 (JAXA)	A. JAXA衛星及び全球降水観測計画(GPM計画)等の海外機関との相互協力の枠組に基づく海外衛星の地球観測実績(観測エリア、観測日時等) B. JAXAの中・低分解能地球観測衛星プロダクト C. GPM計画等の海外機関との相互協力の枠組に基づく海外衛星プロダクト D. 科学衛星や探査機が取得した宇宙科学データ 【データベース】 ・地球観測衛星データ提供システム(G-Portal) ・宇宙科学データアーカイブ(DARTS) [*] など ※宇宙科学データについては、DARTSへの統合を推進。	①オープン(公開) 「A.地球観測実績」、「B. 中・低分解能地球観測衛星プロダクト」及び「C. 海外衛星プロダクト」は、G-Portalにて誰でも検索可能。 「D. 宇宙科学データ」は、宇宙科学研究所のデータポリシーに従ってDARTSにて公開し、用途を問わず誰でも利用可能。 ②セミオープン(制限公開) 「A.地球観測実績」、「B. 中・低分解能地球観測衛星プロダクト」及び「C. 海外衛星プロダクト」は、G-Portalにて利用者登録した者が利用可能。 の2分類	「A. 地球観測実績」、「B. 中・低分解能地球衛星プロダクト」及び「C. 海外衛星プロダクト」は、データ処理を行い、データベース登録したのから即時公開。 「D. 宇宙科学データ」は、データ取得から原則1年以内に公開。	2020年4月から、米国航空宇宙局(NASA)及び欧州宇宙機関(ESA)と協力して、地球観測衛星データを活用した新型コロナウイルス感染症による地球環境や社会経済活動等の変化の解析を実施。同年6月25日から、多様な専門家が社会経済・地球環境をはじめとする分野で活用できるよう、3機関合同ウェブサイト「Earth Observing Dashboard」及びJAXA単独ウェブサイト「JAXA for Earth」にて解析結果を公開中。 <ウェブサイト> ・「Earth Observing Dashboard」 http://eodashboard.org ・「JAXA for Earth」 http://earth.jaxa.jp/covid19/

機関・プロジェクト名	公開データの種別 (適用範囲)	オープン・クローズの考え方 (制限段階)	公開のタイミング (公開までの猶予期間)	新型コロナウイルス感染症対策で 講じた措置・活用事例
日本原子力研究開発機構(JAEA)	規格化重量減少量(NML)または規格化浸出速度(NLR)の数値データが記載されているもの(一部、試験条件と結果から算出できるものは算出している。)を対象として、各々の試験条件、試験結果および出典元についてのデータや緩衝材の基本特性試験によって得られたデータなど 【データベース】 ・ガラスの溶解に関するデータベース ・緩衝材基本特性データベース など	①オープン(公開) ②セミオープン(制限公開) 緩衝材基本特性データベースでデータを検索・利用するためには、ユーザ登録が必要。 の2分類	データベースに登録したものを即時公開。	—
理化学研究所	理化学研究所における研究活動において取得された生命科学系データやデータベースについて記述された、セマンティックウェブ準拠のメタデータを集約、統合し公開している。データベースカタログとして提供されるメタデータは理研が公開したデータベースを対象としている。さらに哺乳類、植物タンパク質、微生物等を含む主要なデータについては詳細な説明を含むメタデータを公共オントロジーを活用して編纂し、公開している。 【データベース】 ・理研メタデータベース	①オープン(公開) メタデータは編纂作業完了後は原則公開であるが、研究中あるいは編纂中のメタデータについては公開を猶予する。	公開猶予のメタデータを除き、データベースに登録したものを即時公開。公開猶予が解けたメタデータについてはその時点で即時公開。	新型コロナウイルス感染症関連の研究で新たに産出されたデータのカタログ情報(データベースのメタデータ)について公開作業を進めている。
	天然化合物の名称・構造、生物活性情報のほか、物性データや機器分析データなど。 【データベース】 ・NPEdia	①オープン(公開) ②クローズ(非公開) 独自のスクリーニング結果、共同研究で得られた評価結果については非公開。 の2分類	データベースに登録したものを即時公開。	—
	バイオリソース研究センターの保有するバイオリソースおよび、その特性情報。 【データベース】 ・(構築中/公開予定) COVID-19研究情報の整理統合および研究に必要なバイオリソースのデータベース	①オープン(公開) 提供可能なバイオリソースに関しては原則公開 ③クローズ(非公開) 提供準備、あるいは提供できない(保存のみ)のバイオリソースに関しては非公開 の2分類	バイオリソースとして提供可能となり次第公開	COVID-19に関する研究情報を、1.ウイルス感染・増殖、2. 発症・生体防御、3. 重症化、4. 治療・予防の4つの視点で整理統合し、各場面で役立つリソースを目的に応じて検索可能とする。

機関・プロジェクト名	公開データの種別 (適用範囲)	オープン・クローズの考え方 (制限段階)	公開のタイミング (公開までの猶予期間)	新型コロナウイルス感染症対策で 講じた措置・活用事例
理化学研究所	<p>生命科学分野の画像データと生命システム動態の定量データ。</p> <p>【データベース】 ・SSBD:database (Systems Science of Biological Dynamics database) 生命科学分野の画像データと生命システム動態の定量データを共有する統合データベース 2020年1月現在687セットの動態情報、4,162セットの画像情報を公開</p> <p>【リポジトリ】 ・SSBD:repository (Systems Science of Biological Dynamics repository) 生命科学分野の画像データと生命システム動態の定量データを共有するリポジトリ</p>	①オープン(公開)	SSBD:database データベースへの登録作業が完了したものが即時公開。 SSBD:repository 論文発表時に公開	—
	<p>タンパク質の量子化学計算値データベース。理化学研究所が中心であるが、星薬科大学、神戸大学など、FMO創薬コンソーシアム(https://fmodd.jp/)のメンバーが計算したデータも含まれる。</p> <p>【データベース】 ・FMOデータベース (タンパク質の量子化学計算値データベース)</p>	①オープン(公開)	データベースに登録したものから順次公開。論文公開前のものについては、論文公開後に登録。	新型コロナウイルス(SARS-CoV-2)タンパク質と治療候補化合物の分子間相互作用をFMO法を用いて計算したデータを公開。治療候補化合物がどのようなメカニズムで標的タンパク質と結合するかを精密に評価し、より強く標的に結合する効果の高い治療薬を設計する際に役立つと期待される。 https://www.riken.jp/pr/news/2020/20200417_1/
	<p>理化学研究所が運用する全天X線監視装置MAXIの観測データ。</p> <p>【データベース】 ・全天X線監視装置MAXIデータベース 417個のX線天体の11年間のX線画像、X線光度曲線などを提供。 MAXIオンデマンド解析(指定された任意の時間の任意の天体の、X線画像、光度曲線、スペクトルを提供)</p>	①オープン(公開)	データ処理が終了次第、即時公開。早いもので観測4時間後、通常は観測1日後。	—

機関・プロジェクト名	公開データの種別 (適用範囲)	オープン・クローズの考え方 (制限段階)	公開のタイミング (公開までの猶予期間)	新型コロナウイルス感染症対策で 講じた措置・活用事例
理化学研究所	革新脳プロジェクト(革新的技術による脳機能ネットワークの全容解明)で得られ、論文化したマーマセット脳の構造・機能に関するデータを公開している。主な内容は、マーマセットの脳のMRI画像、マクロレベルの結合、遺伝子発現マップ、ECoG電極アレイによる機能データなどである。	①オープン(公開) ③クローズ(非公開) の2分類。一般公開することにより有用な研究基盤になると思われるデータを、論文化が終わった後に公開している。	論文が出しだい公開する。	—
日本医療研究開発機構(AMED)	「ゲノム医療実現のためのデータシェアリングポリシー」適用のAMED研究開発課題における、研究成果に紐づくゲノムデータ及び臨床情報や解析・解釈結果等を含めたゲノム情報等。 【データベース】 ・AGD(AMEDゲノム制限共有データベース)	①オープン(公開) ②セミオープン(制限公開) ③セミクローズ(制限共有) データの共有は原則的に研究者間の合意に基づく。 AGDの制限共有データを利用する場合は、研究者間の合意後、データ利用申請が必要。 の3分類	原則として、「生データ生成後2年」又は「研究成果の公表時(論文採択、特許出願公開等の日)」のいずれか早い時点までに登録。 AGD登録データは、公開(セミオープンまたはオープン)を前提としたデータ。	—
	東北メディカル・メガバンク計画のゲノム・コホート調査に参加した日本人の血漿の、網羅的メタボローム及びプロテオーム統合解析を行い、各代謝物やタンパク質の分布や頻度情報などをまとめたデータベース(日本人多層オミックス参照パネル)。 【データベース】 ・jMorp	①オープン(公開)	データベースに登録したものを即時公開。	—
	東北メディカル・メガバンク計画のゲノム・コホート調査由来の、健康調査及び全ゲノム配列情報を含む生体試料の解析情報(基本属性情報、調査票情報、生理学検査情報、検体検査情報、ゲノム・オミックス情報、診療情報、MRI画像情報等)を統合したデータベース。 【データベース】 ・dbTMM	①セミオープン(制限公開) 所定の登録・審査の手続きを経て利用可能。	データベースに登録したものを即時公開。	—

機関・プロジェクト名	公開データの種別 (適用範囲)	オープン・クローズの考え方 (制限段階)	公開のタイミング (公開までの猶予期間)	新型コロナウイルス感染症対策で 講じた措置・活用事例
情報統合型物質・材料開発イニシア ティブ (MI ² I)	NIMSの研究活動で得られた物質・材料の基礎データ や、データ科学的手法等の解析ツールを格納。	①オープン(公開) 解析ツールのうち公開できるものに限る。 ②クローズ(非公開) プロジェクト独自のデータ、共有できないデータ・ ツールなどについては非公開。 の2分類	解析ツールを公開するか否かは、所有者の 判断に委ねている。 プロジェクト期間中(～2020.3.31まで)にプロ ジェクト参加者とコンソーシアム会員の範囲内 でのセミクローズ(制限共有)としていたデータ は、クローズ(非公開)に変更。当該データを用 いた研究論文発表を行った場合は、NIMS の材料データリポジトリ(Materials Data Repository:MDR)で公開することを推奨して いる。	—
データ統合・解析システム (DIAS)	地球環境情報のプラットフォームとして、国内外の各機 関が保有する衛星観測、大気・地上気象観測、海洋観 測、気候変動予測モデルの出力データなど	①オープン(公開) DIAS利用規約に同意することで利用可能 ②セミオープン(制限公開) データ利用者の申請を受け、データ提供者の許 可を得ることで利用可能 ③セミクローズ(制限共有) データ提供者があらかじめ設定したグループの 構成員のみ利用可能 ④クローズ(非公開) の4分類	データベースに登録したのから即時公開。	—
バイオサイエンスデータベースセン ター (NBDC)	公的資金を用いたプロジェクト等で産生された生命科 学分野の研究データ(ゲノム配列、タンパク質、代謝産 物、画像等)及び公的資金を用いたプロジェクト等で産 生されたヒトに関するデータ(DDBJと連携、PDBjを一部 支援)。	①オープン(公開) ②セミオープン(制限公開) NBDCヒトゲノムデータベースで制限公開デー タを登録/利用する場合には、登録/利用申請が 必要。 の2分類	論文等による成果公開や知的財産権取得等 の期間の後、公開。	新型コロナウイルス(SARS-CoV-2)および 新型コロナウイルス感染症(COVID-19)の 研究データ・研究リソースに関する情報紹介 ページを作成(2020年3月3日公開、6月1日 更新)したところ約3万4千件(R2年6月末時 点)の閲覧あり(NBDCポータルサイト全体と して、国内からのアクセスが増加した)。 https://biosciencedbc.jp/blog/20200303- 01.html
北極域研究加速プロジェクト (ArCS II)	北極域の調査観測データ、レーダー観測及び衛星観測 によるリモートセンシングデータ、生物、土壌、海水等 の調査研究等で得られた標本(サンプル)の分析データ 等。 なお、モデル系、工学系、人文社会科学系のデータに 関しては個別対応。 【データベース】 ・北極域データアーカイブシステム(ADS)	①オープン(公開) 北極域研究加速プロジェクトにて観測したデー タについては原則公開。 ②セミオープン(制限公開) ③クローズ(非公開) 非公開・制限公開が必要となるものについ ては、個別に判断。 の3分類	メタデータについては3ヶ月以内に登録し、準 備でき次第公開。 実データについては3ヶ月以内に登録し、2年 以内に公開。 なお、国際共同研究等により各種条件・取決 め・契約等がある場合は個別対応。	—