

資料1-2

国際的動向を踏まえたオープン
サイエンスの推進に関する検討会
(第11回)
令和元年8月20日(火)

研究データ基盤整備と国際展開

ワーキング・グループ

報告書

図表集

研究ワークフローと利用されるツール

BY: Jeroen Bosman, <http://dx.doi.org/10.6084/m9.figshare.1286826>



ベンチャー発の便利な研究ツールは大手出版社に買収

研究ワークフローとElsevier



Scopus **Engineering Village** 文献データベース
文献データベース

MENDELEY 文献管理ツール

Knovel ワークフローツール **hivebench** デジタル実験ノート

Mendeley Data platform 研究データ管理

WebShop 論文執筆・投稿支援サービス

SSRN **bepress™** Digital Commons
プレプリントサーバ 機関リポジトリ

ScienceDirect 出版プラットフォーム

bepress™ Digital Commons 機関リポジトリ

SciVal **PLUM ANALYTICS** **Pure** 業績データベース
研究分析ツール 研究成果評価分析ツール

研究データをめぐる世界的な動向

Ø ドイツ学長協会

- 2014年に、組織における**研究データの管理が大学執行部の戦略的課題**として提言

Ø エディンバラ大学

- 2011年に、エディンバラ大学がイギリスで最初に**機関の研究データ管理ポリシー**を作成

Ø UCバークレー

- 2013年に、**Data Intensive Scienceを促進**するためにBerkeley Institute for Data Science (BIDS)を設置

Ø オーストラリア：組織連携の強化

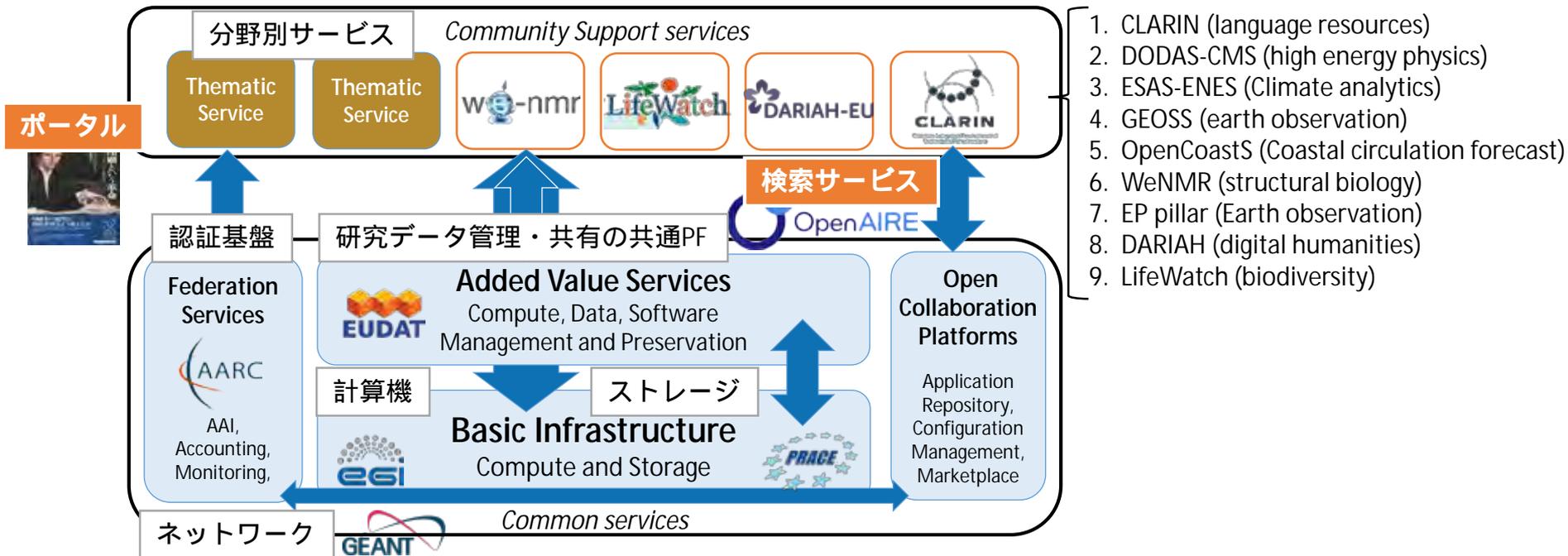
- **研究データインフラ提供や、ストレージ、サービス基盤のサポート**などを行う複数の連携プログラムからなるAustralian Research Data Cloud (ARDC) を設立

Ø G7/G8科学技術大臣会合

- G8科学大臣会合（英国 2013年）
共同声明において、オープンアクセスと研究データのオープン化について言及
- G7科学技術大臣会合（つくば 2016年）
オープンサイエンス・ワーキンググループの設置（EUと日本が共同議長）
- **オープンサイエンス・ワークショップ**（パリ 2019年）

European Open Science Cloud

- **従来**：個々の大学や研究機関でEインフラを構築
- **今後**：個々のインフラを繋ぎEU全体でEインフラを構築
 - 既存のEインフラをベースにネットワークから分野別DBまでを有機的に結合
 - ポータル（EOSC-hub）検索サービス（OpenAIRE）から具体化
 - 人社系へのサポートなどロングテールにおける活用も考慮
 - 産業界との連携も考慮



国際的動向を踏まえたオープンサイエンスに関する検討会（2015年） 我が国におけるオープンサイエンス推進のあり方について（抜粋）

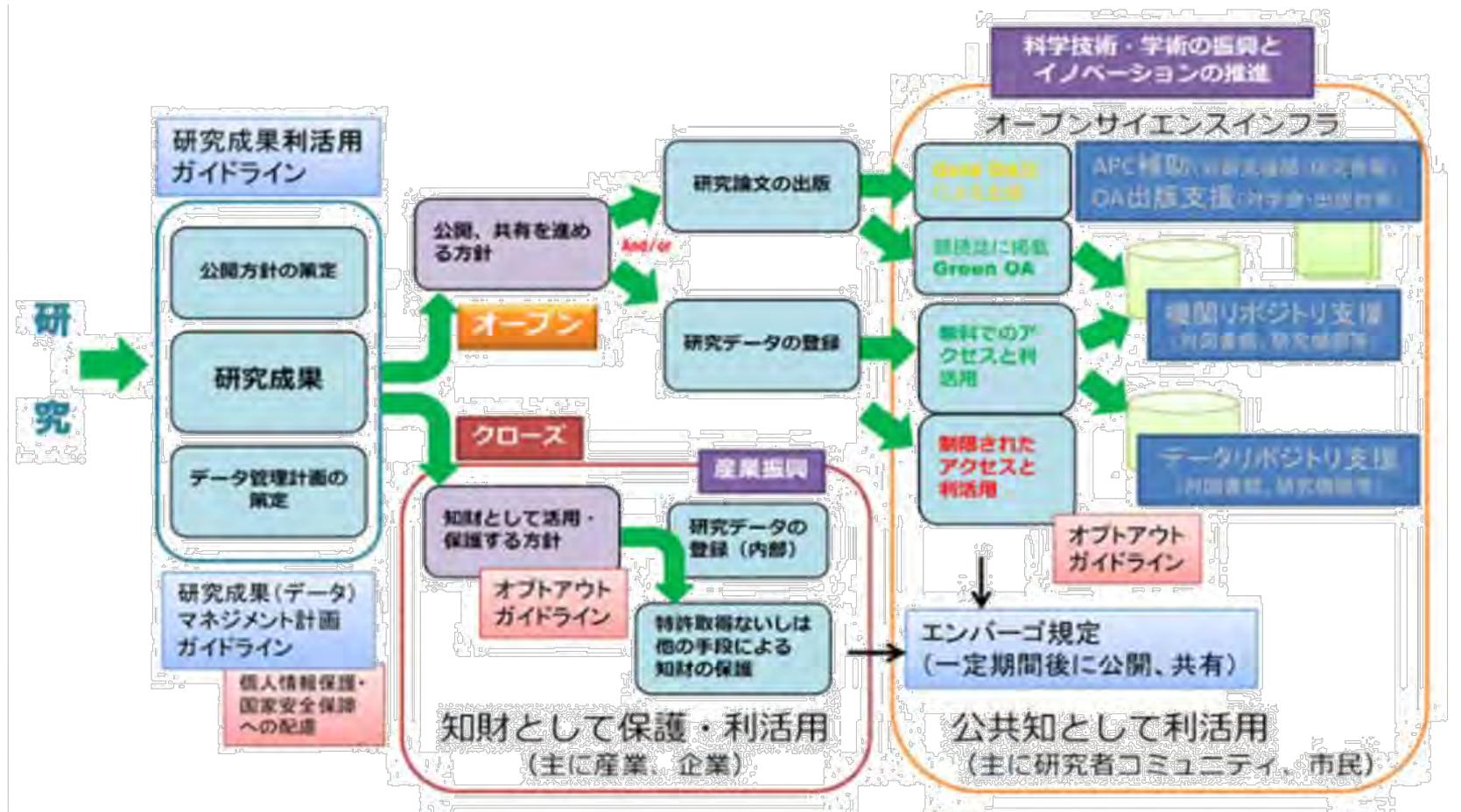
国としての基本姿勢・基本方針

公的研究資金による研究成果（論文、研究データ等）の利活用促進を拡大することを我が国のオープンサイエンス推進の基本姿勢とする。

その上で、各省庁、資金配分機関、大学・研究機関等のステークホルダーがオープンサイエンスの実施の責任を果たし、オープンサイエンスの実施方針及び実施計画を策定するものとし、それに当たって、参照すべき共通事項や留意点等をオープンサイエンスに関する基本方針として明示する。

オープンサイエンスの推進に当たっては、内閣府及び総合科学技術・イノベーション会議が政府全体を通じた旗振り役として中核的な役割を担い、各ステークホルダーにおける進捗状況をフォローする。

国際的動向を踏まえたオープンサイエンスに関する検討会（2015年） 我が国におけるオープンサイエンス推進のあり方について（抜粋）



公的資金：競争的資金及び公募型資金。運営費交付金を100%活用した研究も含むべき。

公的資金による研究成果については、その利活用を可能な限り拡大することを、我が国のオープンサイエンス推進の基本姿勢とする。その他の研究成果としての研究二次データについても、分野により研究データの保存と共有方法が異なることを念頭に置いた上で可能な範囲で公開する。

ただし、研究成果のうち、国家安全保障等に係るデータ、商業目的で収集されたデータなどは公開適用対象外とする。また、データへのアクセスやデータの利用には、個人のプライバシー保護、財産的価値のある成果物の保護の観点から制限事項を設ける。なお、研究分野によって研究データの保存と共有の方法に違いがあることを認識するとともに、国益等を意識したオープン・アンド・クローズ戦略及び知的財産の実施等に留意することが重要である。

目次

1 . 本ガイドラインの位置付け

2 . データポリシー策定のポイント及び並行して取り組む事項

(1) ポリシー策定の目的

(2) ポリシー策定の主体

(3) 管理対象とするデータが具備すべき要件

(4) データ利活用のための要件

(5) ポリシー策定とともに取り組むべき事項

(6) その他

3 . データポリシーで定めるべき項目

(1) 機関におけるポリシー策定の目的について

(2) 管理する研究データの定義、制限事項について

(3) 研究データの保存・管理・運用・セキュリティについて

(4) 研究データに対するメタデータ、識別子の付与、フォーマットについて

(5) 研究データの帰属、知的財産の取り扱いについて

(6) 研究データの公開、非公開および猶予期間ならびに引用について

目次

- 1 . 研究データの保存・公開とデータリポジトリの整備・運用
- 2 . 研究データリポジトリの定義及び役割
- 3 . FAIR 原則
- 4 . 信頼できる研究データリポジトリの整備・運用に関する要件
 - 4.1 . 運営体制
 - 4.1.1 . 研究データリポジトリのミッション
 - 4.1.2 . 運営組織
 - 4.1.3 . 運営計画
 - 4.1.4 . データポリシー
 - 4.2 . 情報基盤
 - 4.2.1 . 研究データリポジトリのICT インフラ（狭義の情報基盤）
 - 4.2.2 . データ管理システム
 - 4.2.3 . データバックアップシステム
 - 4.2.4 . データ検索システム
 - 4.3 . 人的基盤
 - 4.3.1 . 研究データリポジトリにおける業務
 - 4.3.2 . 運用スタッフに求められる技能・経験
 - 4.3.3 . 運用スタッフの配置・育成

国際的動向を踏まえたオープンサイエンスの推進に関する検討会（2017.12～）

国際動向を踏まえたオープンサイエンス推進のための方策等について検討し、本政策分野における国際プレゼンスの向上、及び国内施策の充実を図るため、「国際的動向を踏まえたオープンサイエンスの推進に関する検討会」を開催する。

研究データ基盤整備と国際展開ワーキング・グループ（2019.2～）

国際的な学術情報流通の潮流の中で、我が国独自の研究データ基盤を構築し、利活用を促すとともに、国際連携等により我が国のプレゼンスの向上を図る必要がある。このための基本的な考え方及び方策について議論するため、「国際的動向を踏まえたオープンサイエンス推進のに関する検討会」の下に、「研究データ基盤整備と国際展開ワーキング・グループ」を設置する。

研究データ基盤整備と国際展開WG委員

(2019年6月21日時点)

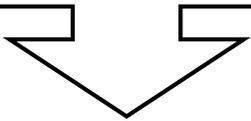
氏名 (敬称略・50音順)	所属・役職	備考
家 泰弘	独立行政法人 日本学術振興会 理事	
小賀坂 康志	国立研究開発法人 科学技術振興機構 情報基盤事業部長	
加藤 治	国立研究開発法人 日本医療研究開発機構 基盤研究事業部長	
喜連川 優	大学共同利用機関法人 情報・システム研究機構 国立情報学研究所 所長 国立大学法人 東京大学 生産技術研究所 教授	主査
後藤 文郷	住友化学株式会社 先端材料開発研究所 グループマネージャー	
末吉 互	潮見坂綜合法律事務所 弁護士	
谷 幹也	日本電気株式会社 中央研究所 セキュリティ研究所 所長	
林 和弘	文部科学省 科学技術・学術政策研究所 科学技術予測センター 上席研究官	主査代理
松井 啓之	国立大学法人 京都大学 経営管理大学院 教授	
眞野 浩	一般社団法人 データ流通推進協議会 代表理事	
村山 泰啓	国立研究開発法人 情報通信研究機構 戦略的プログラムオフィス研究統括	
谷治 和文	国立研究開発法人 新エネルギー・産業技術総合開発機構 技術戦略研究センター 標準化・知財ユニット長	
山地 一禎	大学共同利用機関法人 情報・システム研究機構 国立情報学研究所 教授	

引原隆 京都大学教授/図書館長 (国際的動向を踏まえたオープンサイエンスの推進に関する検討会主査)

上山隆大 総合科学技術会議議員

が随時出席

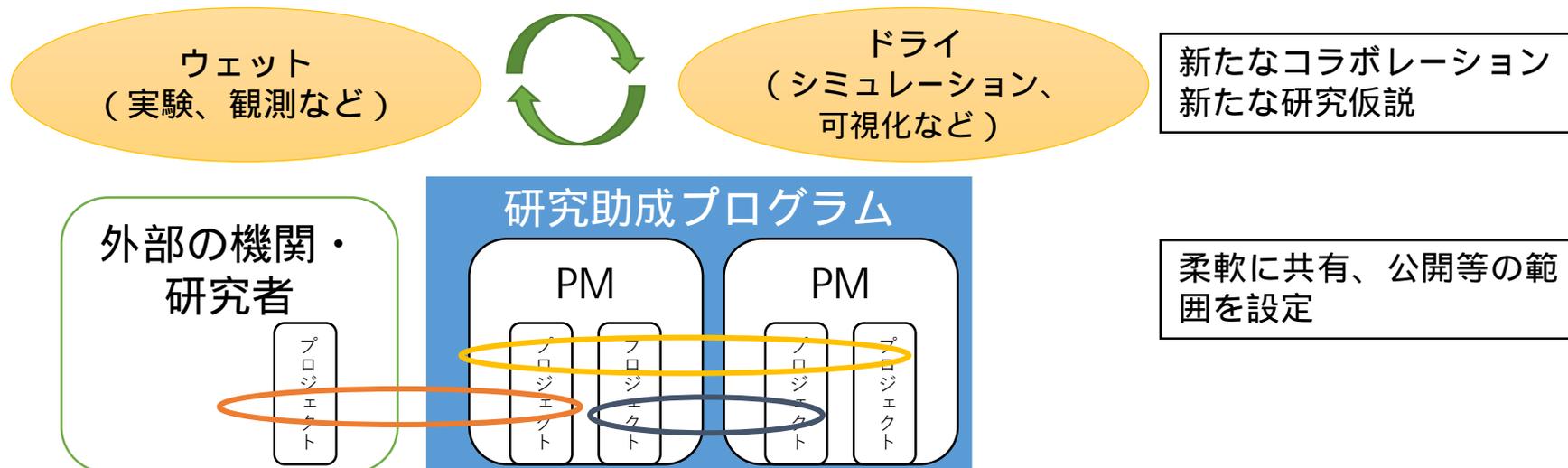
- 我が国の研究データを適切に管理・共有・公開・検索するための研究データ基盤システムの構築や政府全体の体制に関する検討
- 国レベルでの研究データ戦略やデータポリシー等に関する検討
- 国際的な対応のあり方の検討
- 地球環境測定系、生命科学系、材料科学系、工学系等の各研究分野における最先端研究者の研究データマネジメントに関するケーススタディ

- 
- 国レベルの研究データ戦略やポリシー
 - 統合イノベーション戦略 2019
 - 次期科学技術基本計画 など

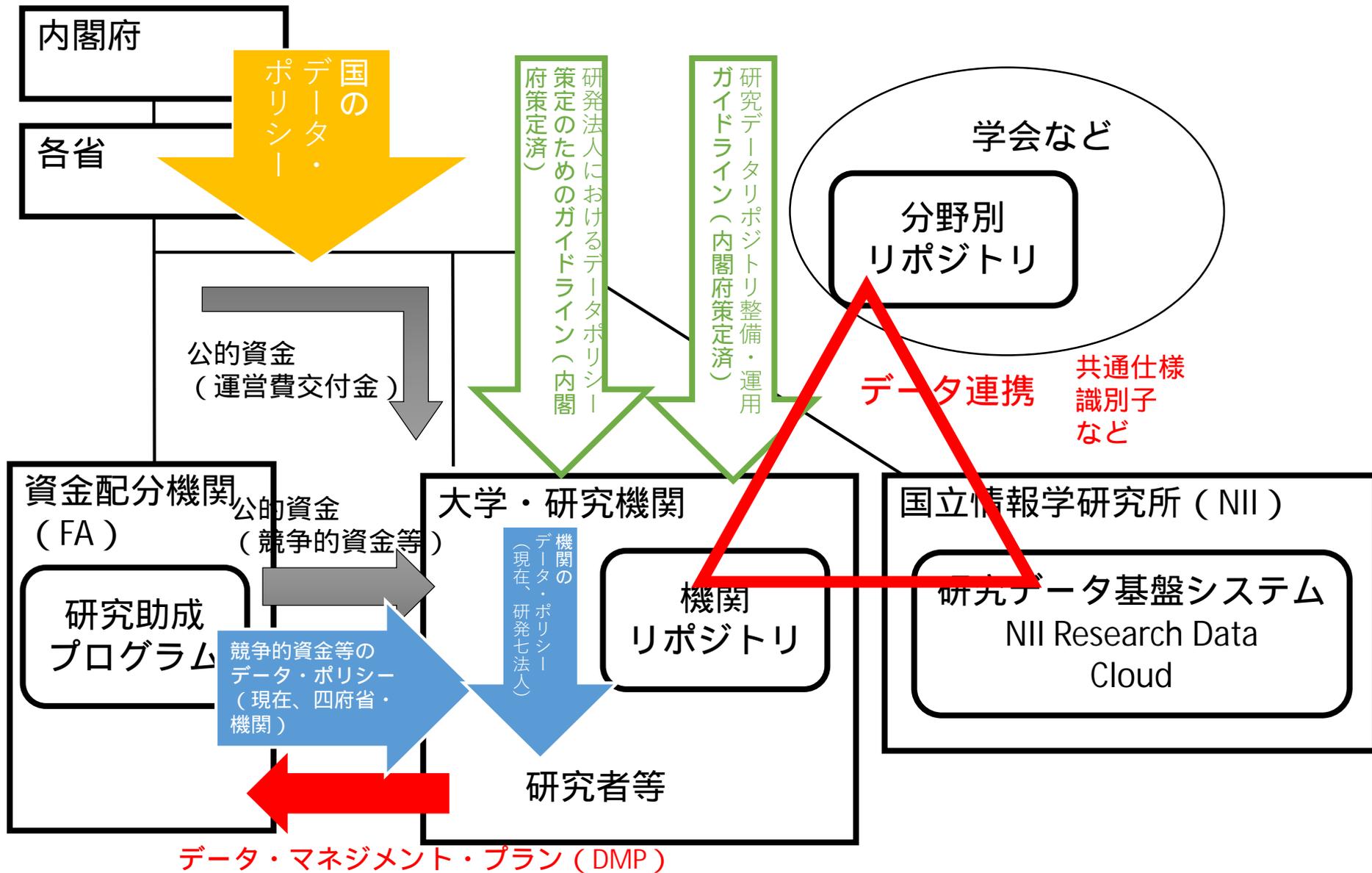
研究プロセスにおける研究データのマネジメントの意義

研究データを体系化、共有、公開等を行うことにより、新たなコラボレーションによる独創的な研究成果・イノベーションの創出や、研究プロセス・マネジメントの高度化・効率化につながる。

- プログラムマネージャー（PM）やプロジェクトの研究代表者（PI）の方針の下で、プロジェクト、分野、領域等をまたいで、柔軟に共有、公開等の範囲を設定することにより、新たな研究や事業化の芽を生み出すことが期待できる。
- PMやPIが、プロジェクトにおける研究データの体系化、共有化を促すことにより、研究プロセスが関係者間で可視化、共有化され、研究プロセスやマネジメントの高度化・効率化につながることを期待できる。
- また、研究者等の評価においても、論文等の成果だけでなく、データの体系化、共有、公開等を通じ、研究プロセスに関するきめ細かい配慮の在り方を検討していくことが求められる。



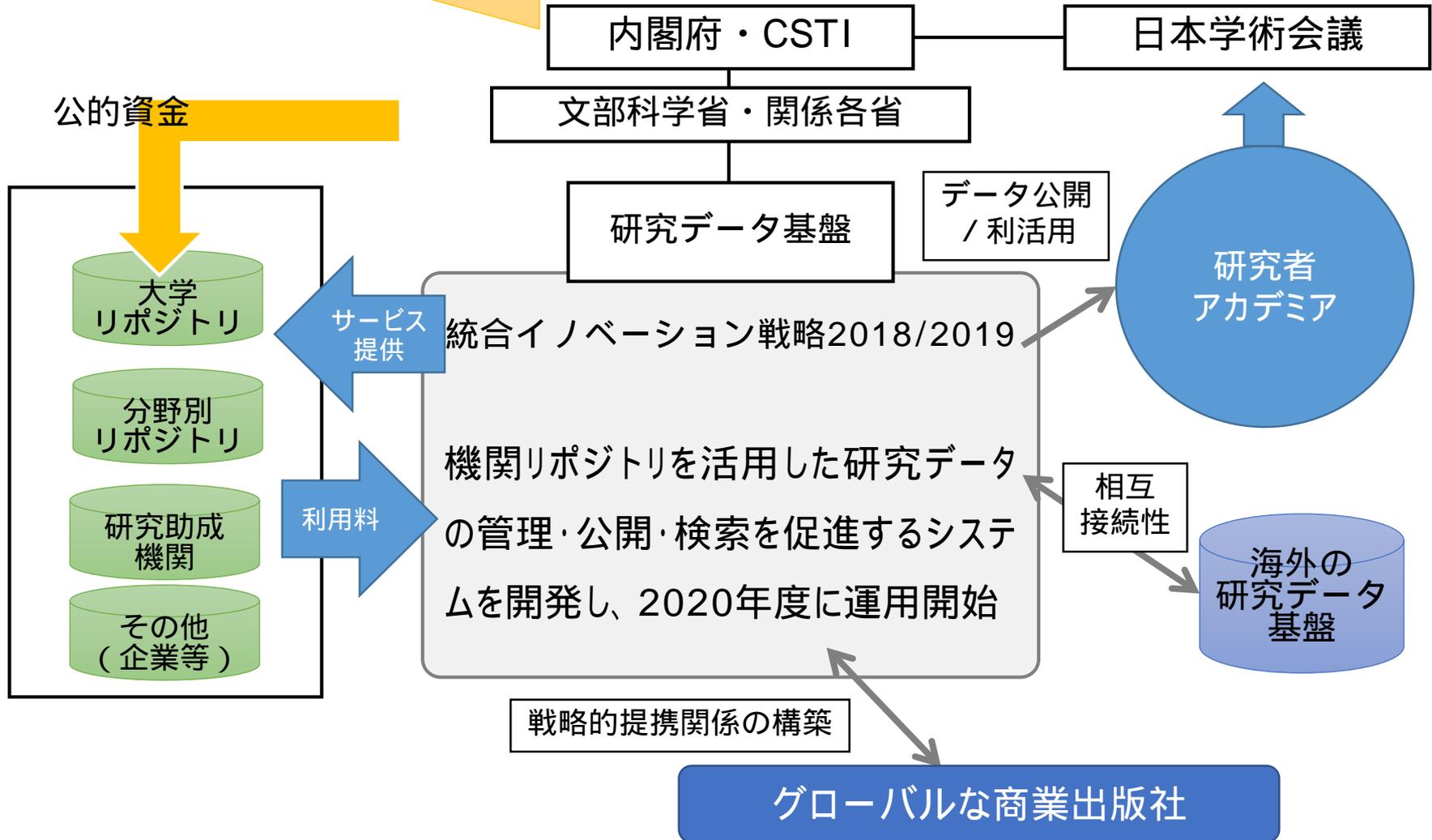
研究データ基盤システムと関連体制



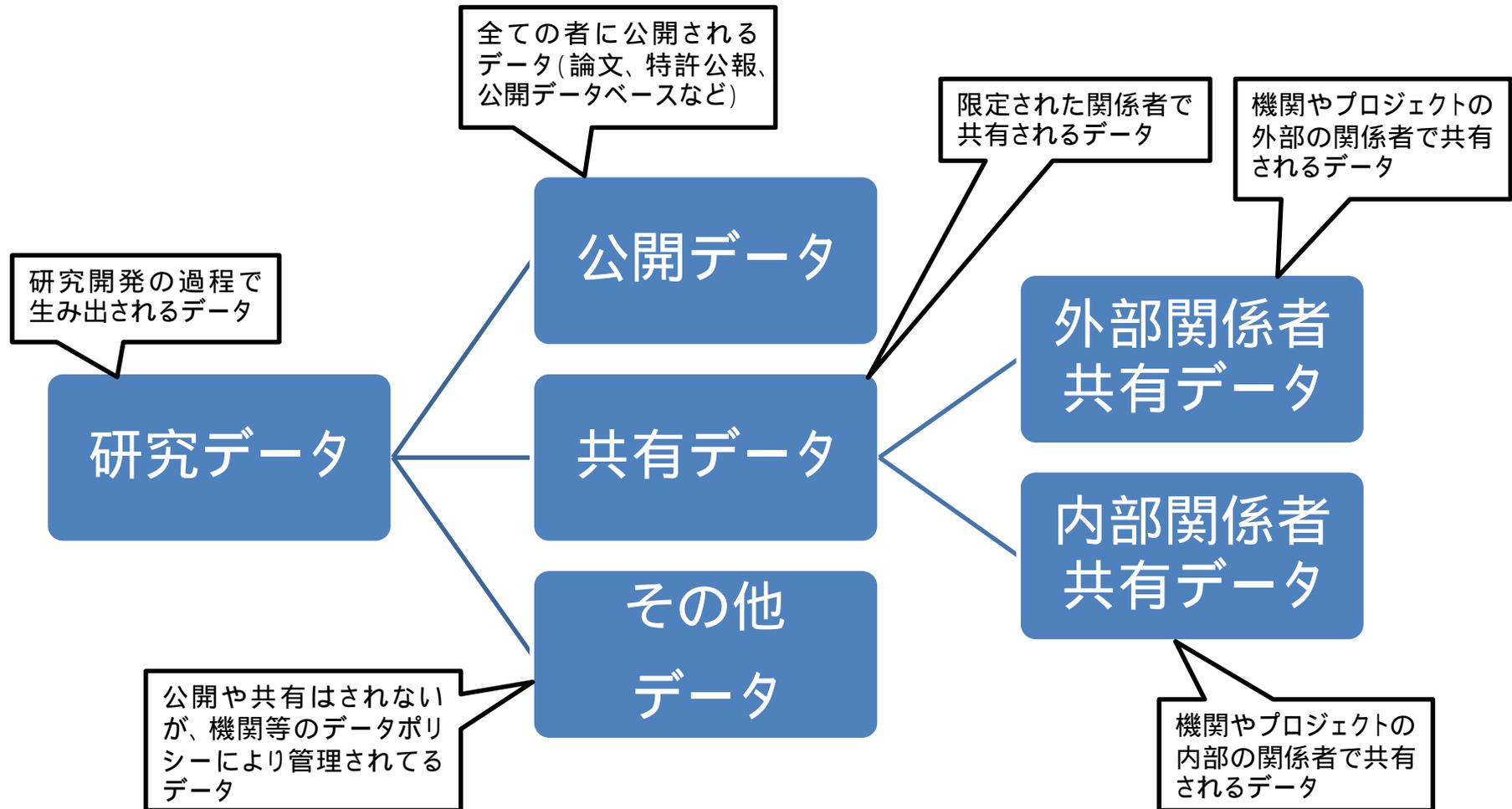
研究データ基盤と国際展開の全体像

- ・ ナショナル・データ・ポリシーの策定
- ・ G7等の国際展開

アカデミアの
立場からの検討

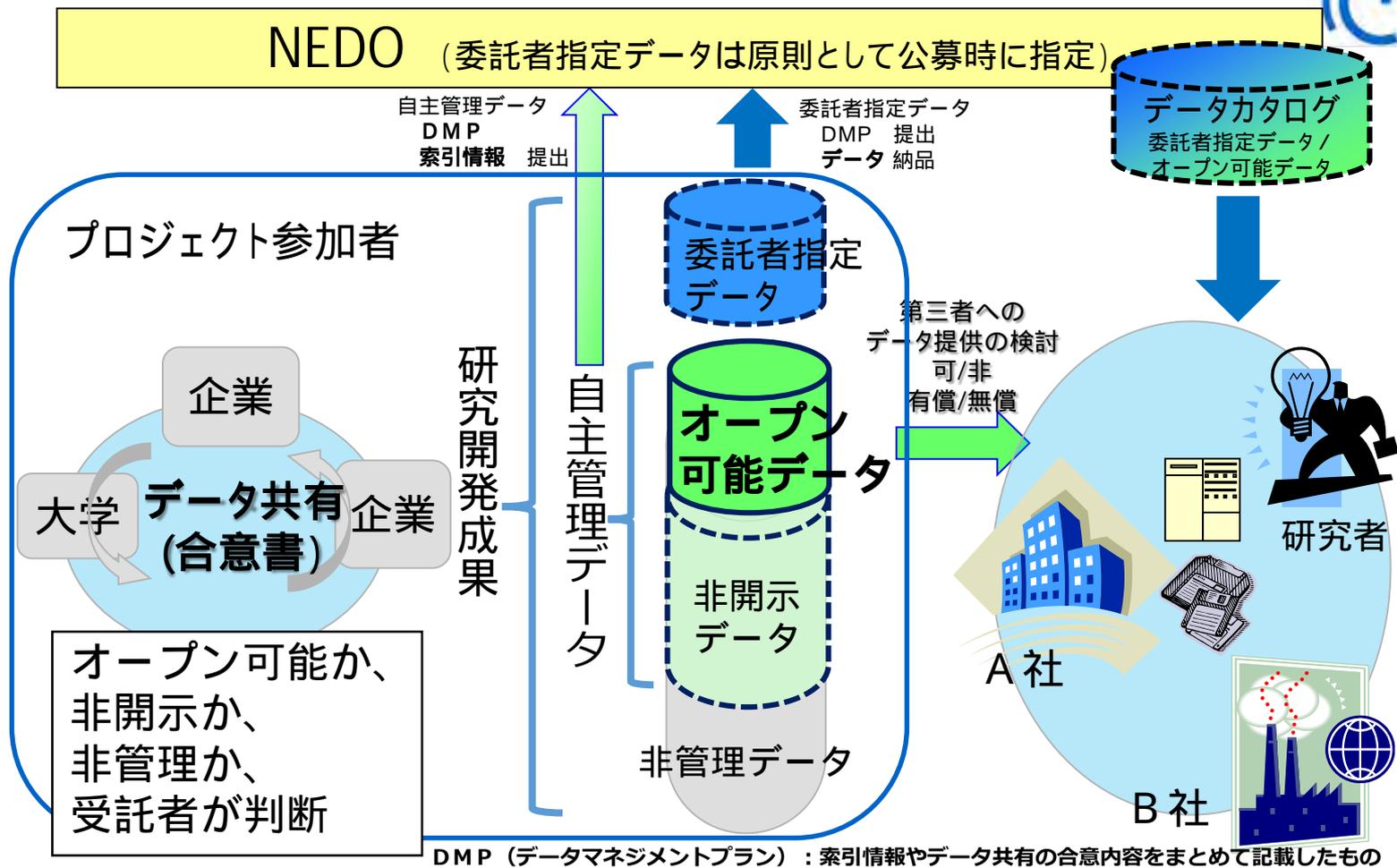


研究データの範囲と区分



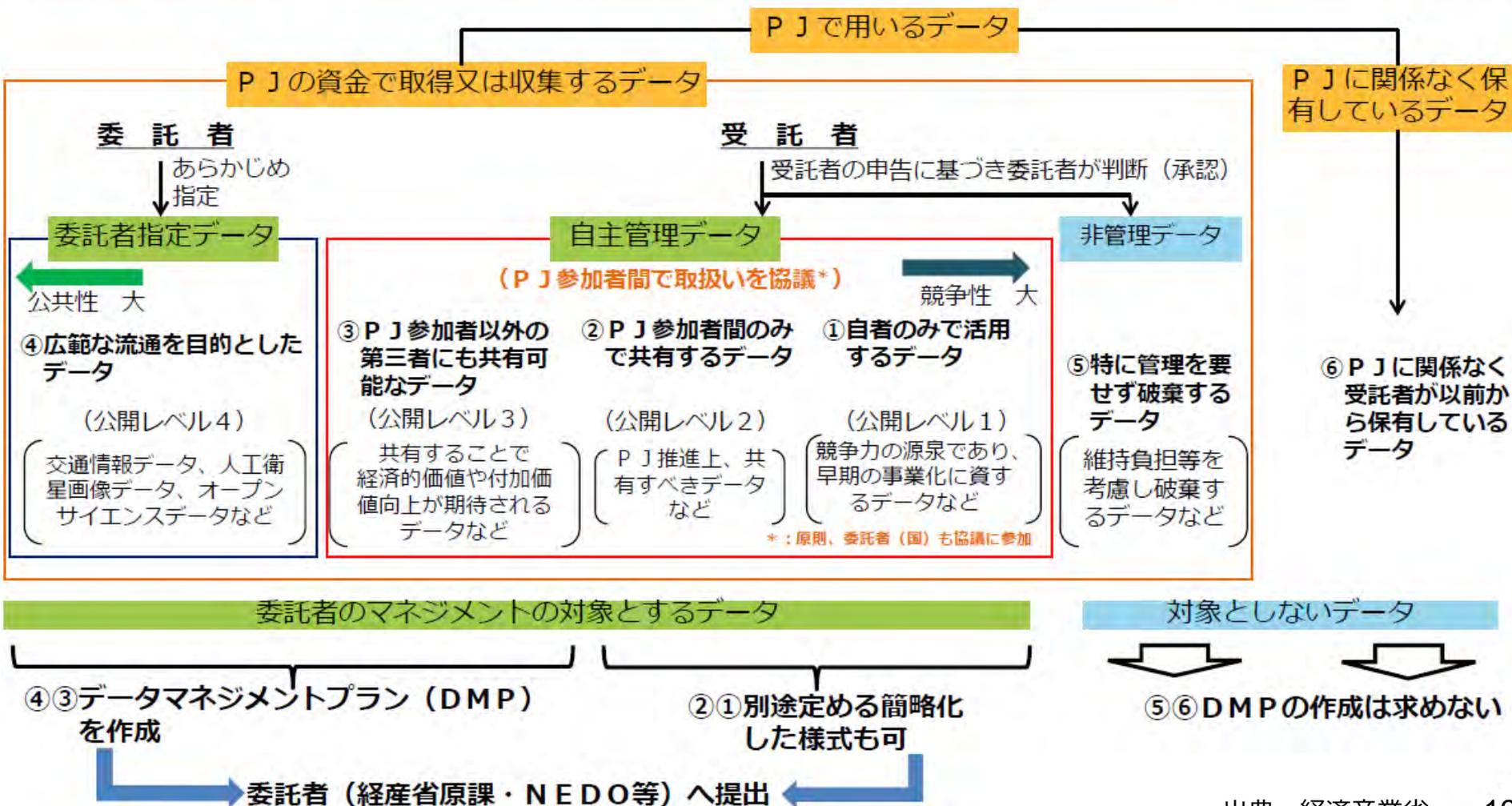
* 全ての研究データにおいて、セキュリティの確保、関係諸法令(個人情報保護法、不正競争防止法)の遵守が求められる。また、データ取り扱いに関する条約等の国際約束等(ロンドンガイドライン、EU・GDPRなど)や、データ管理の原則であるFAIR原則等の国際的な規則や慣行等との整合性に十分留意することが求められる。

経産省・NEDOにおける研究データマネジメントの考え方（1）



研究開発データのマネジメントの基本的な考え方

- データの利活用の促進
- プロジェクト参加者間でのデータの共有、ルール化による事業成果の最大化
- 知財権同様に、委託事業により発生するデータについても委託者のマネジメントの対象として取扱いを定める



経産省・NEDOのナショプロデータカタログ

ナショプロデータカタログ

掲載URL : http://www.meti.go.jp/policy/innovation_policy/data_management.html

公表可能データ(2019年3月版)										
データ No.	掲載日	データ名称	データの説明	管理者	プロジェクト終了後のリポジトリ	概略データ量	研究データの利活用・提供方針	連絡先	契約件名(プロジェクト名)	備考
例1	2019年3月10日	〇〇実験においてセンサより撮像したデータ及び関連データ	〇〇実験においてセンサより撮像したデータであり、道路の画像データ	〇〇研究所	自社に保存	10GB以上100GB未満	一定期間後に事業の実施上有益なものに対して有償又は無償で提供を開始。但しデータのクレジット表記を条件とする。なお、サンプルデータを公開している。	〇〇研究所 〇〇部門〇〇課 TEL:00-0000-0000(内線0000) Mail: 〇〇〇〇@〇〇〇〇	〇●に関する研究開発	サンプルデータを下記URLにて提供。 http://www. 〇〇.co.jp
例2	2019年3月10日	〇〇のシミュレーションデータ	〇〇のシミュレーションデータ	〇〇大学	自社に保存	10GB	一定期間後に事業の実施上有益なものに対して有償又は無償で提供を開始。但しデータのクレジット表記を条件とする。なお、サンプルデータを公開している。	〇〇大学 〇〇研究室 TEL:00-0000-0000(内線0000) Mail: 〇〇〇〇@〇〇〇〇	●〇に関する実験研究	サンプルデータを下記URLにて提供。 http://www.〇〇〇.html
1	2019年3月12日	公開された有害性情報および毒性推定モデルに基づく化学物質分類データ	化学法や化学法の規制根拠情報として、既に公開されている試験結果、有害性情報、規制に関わる有害性クラス、等を入力データとして、統計解析や毒性推定モデルにより得られる化学物質分類データ。またこの過程で得られた中間データ。	国立大学法人大阪大学	本学に保存	1GB未満	一定期間後に事業の実施上有益なものに対しての提供を開始する。但しデータのクレジット表記を条件とする。	大阪大学大学院工学研究科(環境・エネルギー工学専攻)東海明宏 TEL: 06-6879-7676 Mail: tokai@see.eng.osaka-u.ac.jp	平成30年度化学物質安全対策(ベストミックスの視点からみたレスポンスシナリオのあり方)	関連する学術発表状況を下記URLにて提供。 http://www.see.eng.osaka-u.ac.jp/soem/soem/
2	2019年3月12日	非定常流出を想定したリスク評価結果のデータ。また、この過程で得られた中間データ。	化学法に関する公開情報およびそれと関連する補助調査の公開情報等を入力データとして、解析シミュレーションより得られた評価結果のデータ。また、この過程で得られた中間データ。	国立大学法人大阪大学	本学に保存	1GB以上10GB未満	一定期間後に事業の実施上有益なものに対しての提供を開始する。但しデータのクレジット表記を条件とする。	大阪大学大学院工学研究科(環境・エネルギー工学専攻)東海明宏 TEL: 06-6879-7676 Mail: tokai@see.eng.osaka-u.ac.jp	平成30年度化学物質安全対策(ベストミックスの視点からみたレスポンスシナリオのあり方)	関連する学術発表状況を下記URLにて提供。 http://www.see.eng.osaka-u.ac.jp/soem/soem/
3	2019年3月12日	参加した地域対話において取得したデータ	参加した地域対話において取得した、配布資料、当日の意見交換内容等を整理したデータ	横浜国立大学	自社に保存	1GB以上10GB未満	プロジェクト期間中: 同一プロジェクト参加者または管理者が認める関係者に無償で提供 プロジェクト終了後: 一定期間後に広く公表する。但しデータのクレジット表記を条件とする。	横浜国立大学 環境情報研究所/リスト共生社会想像センター(併任) 齋藤 真枝子 TEL:045-330-3994 Mail:kumasaki-mieko-pd@ynu.ca.jp	レスポンスシナリオ活動におけるリスク情報を活用した地域対話の在り方に関する研究	
4	2019年3月12日	PRTR排出量の推移分析データ	(1)公開情報であるPRTR届出データから、「大気の排出」について、事業所ごとの排出量とその内訳を抽出したもの(燃料小売量を除く45業種、業種別)。 (2)事業所・物質ごとの排出量の年次推移・変化量の算出結果(化学工業ほか一部の業種のみ)	横浜国立大学	自社に保存	1GB以上10GB未満	プロジェクト期間中: 同一プロジェクト参加者または管理者が認める関係者に無償で提供 プロジェクト終了後: 一定期間後に広く公表する。但しデータのクレジット表記を条件とする。	横浜国立大学 環境情報研究所/リスト共生社会想像センター(併任) 齋藤 真枝子 TEL:045-330-3994 Mail:kumasaki-mieko-pd@ynu.ca.jp	レスポンスシナリオ活動におけるリスク情報を活用した地域対話の在り方に関する研究	
5	2019年3月12日	PRTR制度の比較データ	諸外国において整備されているPRTR、または類似制度を調査し、比較整理したデータ	横浜国立大学	自社に保存	1GB以上10GB未満	プロジェクト期間中: 同一プロジェクト参加者または管理者が認める関係者に無償で提供 プロジェクト終了後: 一定期間後に広く公表する。但しデータのクレジット表記を条件とする。	横浜国立大学 環境情報研究所/リスト共生社会想像センター(併任) 齋藤 真枝子 TEL:045-330-3994 Mail:kumasaki-mieko-pd@ynu.ca.jp	レスポンスシナリオ活動におけるリスク情報を活用した地域対話の在り方に関する研究	
6										
7										
8										

出典：経済産業省

各資金配分機関が管理する研究データ区分の関係

データの内容	JST	NEDO*1	AMED*2
公開データ	一般公開	委託者指定データ	非制限公開データ
外部関係者共有データ	研究グループ等での 構成員限りでの共有 / 限定されたものへの公開	PJ参加者以外の 第三者にも 共有可能なデータ	制限公開データ
内部関係者共有データ		PJ参加者間のみで 共有するデータ	制限共有データ
その他データ	非公開	自者のみで 活用するデータ	その他の 研究データ

* 1) NEDOは、NEDOが管理しないデータ区分として、記載のデータ区分以外に「非管理データ」という区分を設けている

* 2) 「ゲノム医療実現のためのデータシェアリングポリシー」より。なお、ゲノム関連事業以外のデータポリシーにデータ区分を設けていない

資金配分機関におけるデータ・マネジメント・プラン（DMP）

大項目	詳細	AMED	JST	NEDO
プロジェクト情報	事業年度	○	○	○
	事業名	○	○	○
	研究開発課題名	○	○	○
データ情報	研究から産出されるデータ及びデータ群の総称	○	○	○
	委託者指定データ、自主管理データの分類			○
	研究データの種類の分類		○	
	研究開発データの説明		○	○
	想定データ量			○
	加工方針（ファイル形式、メタデータに関する事項を含む）			○
	研究計画書対応項目		○	
担当者	データサイエンティストの所属・氏名等	○		
	研究開発データを取得又は収集した者			○
	研究開発データの管理者			○
取得・収集	研究開発データの取得又は収集方法			○
保存・管理	リポジトリ（保存場所）	○	○	○
	管理対象となる研究データの保存・管理方針		○	○
公開・提供	研究データの公開・非公開に係る方針		○	○
	公開可能な研究データの提供方法・体制		○	○
	公開研究データの想定利用用途		○	○
	公開研究データの利活用促進に向けた取り組み		○	○
	研究開発データの利活用・提供方針 （他者に提供する場合）円滑な提供に向けた取り組み； （秘匿して自ら利活用する場合）秘匿期間、秘匿理由			○

なお、表中のAMEDのDMPは汎用的なDMPについてであり、ゲノム関連プロジェクトにおいては、データシェアリングポリシーを設け詳細なDMPを義務づけている

DCC (Data Curation Centre ; 英) のDMP

大項目	詳細	備考
管理上のデータ	助成機関や所属機関が定めた関連ID	
	助成機関・助成金番号	
	プロジェクト名称・内容・連絡先	
	プロジェクトの主任研究員 (PI) もしくは主要研究者の名前とその永続的識別子 (ORCID等)	
	データ管理計画書の作成日・最終更新日	
	関連するポリシー	
データ収集	どのようなデータを収集・生成するのか	データの種別、フォーマット、容量etc.
		フォーマットやソフトウェアのデータ共有・長期アクセスへの適性 再利用できる既存のデータの有無
	どのようにデータを収集・生成するのか	用いる標準や方法論
		フォルダやファイルの構造化の方法 バージョン管理の方法
文書化とメタデータの付与	文書化とメタデータの付与	長期にわたりデータが可読性を有し解釈され得るために必要な情報
		文書化及びメタデータ作成の方法
		用いるメタデータ標準とその理由
倫理・法律のコンプライアンス	倫理的問題	データの保存、共有についての同意取得の有無 (必要な場合の) 参加者の個人識別性の保護
		センシティブデータの安全な保管・移動のためのとりくみ
	著作権、知的財産権等に関する問題	データの所有者
		再利用のために付与するライセンス 第三者による再利用についての制限の有無
保管とバックアップ	研究期間中のデータの保管・バックアップ方法	十分なストレージを有しているか
		付加サービスの利用に必要な料金を使途に含める必要があるか
		データのバックアップ方法
	障害発生時のデータの復元方法	
	アクセス制限やセキュリティ確保の管理	

DCC (Data Curation Centre ; 英) のDMP

大項目	詳細	備考
選定と保存	どのデータを保持・共有・保存すべきか	契約、法律、規制等の理由で保持あるいは破棄すべきデータ その他の保存するデータの決定方法 データの保持・保存の期間
	データセットの長期保存計画の内容	リポジトリやアーカイブ等のデータの保存場所
データ共有	どのように共有するか	
	データ共有に何らかの制限が必要か	
責任と リソース	データ管理の責任者	
	計画実行のために、どのようなリソースが必要か	

DMPの一般的な項目

項目	備考
1 . プロジェクト情報	
・事業年度	
・事業名	
・研究開発課題名	
・体系的課題番号 等	
2 . データ情報	
・取得・収集される研究開発データの名称	データ種別、フォーマット、容量 等
3 . 担当者情報	
・研究代表者	名称、識別子（NRID、ORCID 等）、所属機関 等
・研究分担者	同上
・研究開発データの管理者	同上
・研究開発データの取得者又は収集者	同上
・（仮）データキュレーター（データ専門職）	同上、データキュレーターの定義が必要
4 . 取得・収集	
・研究開発データの取得又は収集方法	関連する標準や方法、品質保証、データの組織化（命名規則、バージョン管理、フォルダ構造 等

DMPの一般的な項目

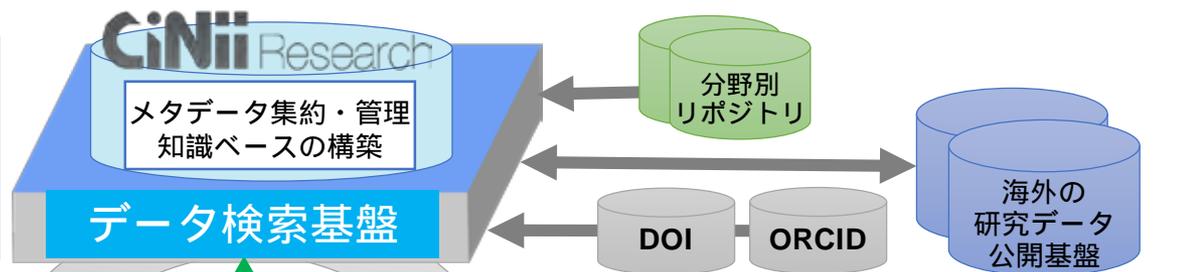
項目	備考
5 . 文書化とメタデータ	
・再利用を可能にするための文書	データの収集・分析・処理方法に関する情報、変数の定義、測定単位、設定した仮説、データのフォーマットとファイル種別、使用したソフトウェア 等
・使用するメタデータ標準（フォーマット）	汎用（Dublin Core、DataCite、JPCOARスキーマ 等） 分野（DDI（社会科学）、SPASE（観測データ） 等）
6 . 研究プロセス中の保管とバックアップ	
・研究中のデータおよびメタデータの保管とバックアップ方法	保管とバックアップの場所や、バックアップの頻度
・研究中のデータのセキュリティと機密データの保護方法	インシデント発生時のデータ復旧方法 アクセス可能なメンバー、アクセスの管理方法 機密データのリスクや管理方法 どの機関のデータ保護方針が適用されるか
7 . 倫理・法律のコンプライアンス	
・倫理的問題への対処	データの保存や共有に関する同意、匿名化処理、センシティブデータの扱い 等
・著作権や知的財産権に関する問題への対処	データの所有者は誰か、再利用のためのライセンス付与、第三者による再利用の制約 等

DMPの一般的な項目

項目	備考
8 . 研究プロセス後の保存・管理	
・研究開発データの保存・管理方針	保存すべきデータの選定 等
・保存期間	日本学術会議は研究公正の観点から10年程度
・保存場所	リポジトリ、ストレージ
9 . 公開・提供	
・研究開発データの公開・提供方針	公開・限定的公開・非公開の方針 非公開の場合の期間、理由 等
・研究開発データの公開・提供場所	リポジトリ（汎用、分野、機関）
・委託者指定データ、自主管理データの分類	NEDOでは、交通情報データ、人口衛星画像データなどを委託者指定データとして例示

NII Research Data Cloud

- 機関リポジトリ + 分野別リポジトリやデータリポジトリとも連携
- 研究者や所属機関、研究プロジェクトの情報とも関連付けた知識ベースを形成
- 研究者による発見のプロセスをサポート



- データ収集装置や解析用計算機とも連携
- 研究遂行中の研究データなどを共同研究者間やラボ内で共有・管理
- 組織が提供するストレージに接続した利用が可能



- データ管理基盤における簡便な操作で研究成果の公開が可能
- 図書館員やデータキュレータによる、メタデータや公開レベル統計情報などの管理機能の提供

