



物質・材料研究データの探しやすさ

図書館総合展2020フォーラム「CiNii, 研究データに出会う。」

2020年11月6日



国立研究開発法人 物質・材料研究機構



統合型材料開発・情報基盤部門 材料データプラットフォームセンター

松田 朝彦 Asahiko MATSUDA

 <https://orcid.org/0000-0001-5989-027X>



物質・材料研究機構 (NIMS) について



国立研究開発法人物質・材料研究機構

National Institute for Materials Science

物質・材料科学 (Materials Science) の研究を専門とする公的研究機関

茨城県つくば市内3地区、兵庫県西播磨地区 (SPring-8内)

役職員数：1,564名 (2020年4月1日現在)

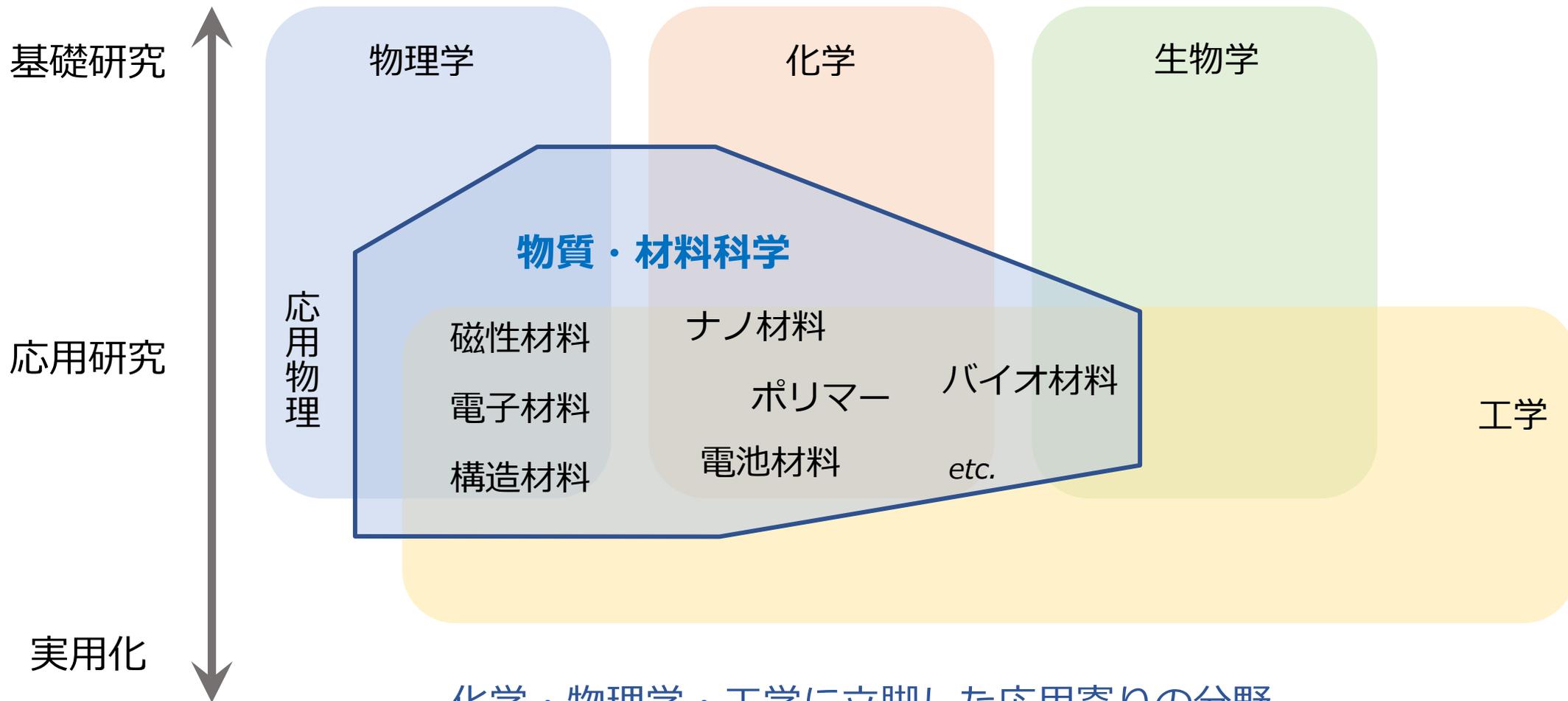
研究組織：機能性材料研究拠点
エネルギー・環境材料研究拠点
磁性・スピントロニクス材料研究拠点
構造材料研究拠点 (RCSM)
国際ナノアーキテクトニクス研究拠点 (MANA)
先端材料解析研究拠点
統合型材料開発・情報基盤部門 (MaDIS)
技術開発・共用部門 (RNFS)



エネルギー材料設計グループ
データ駆動高分子設計グループ
データ駆動構造材料グループ
デバイス材料設計グループ
データ駆動無機材料グループ
SIP-MIラボ
材料データプラットフォームセンター
└ DB Gr データ解析Gr システムGr
└ 図書T 出版T サービスT



物質・材料科学 (材料学) の立ち位置

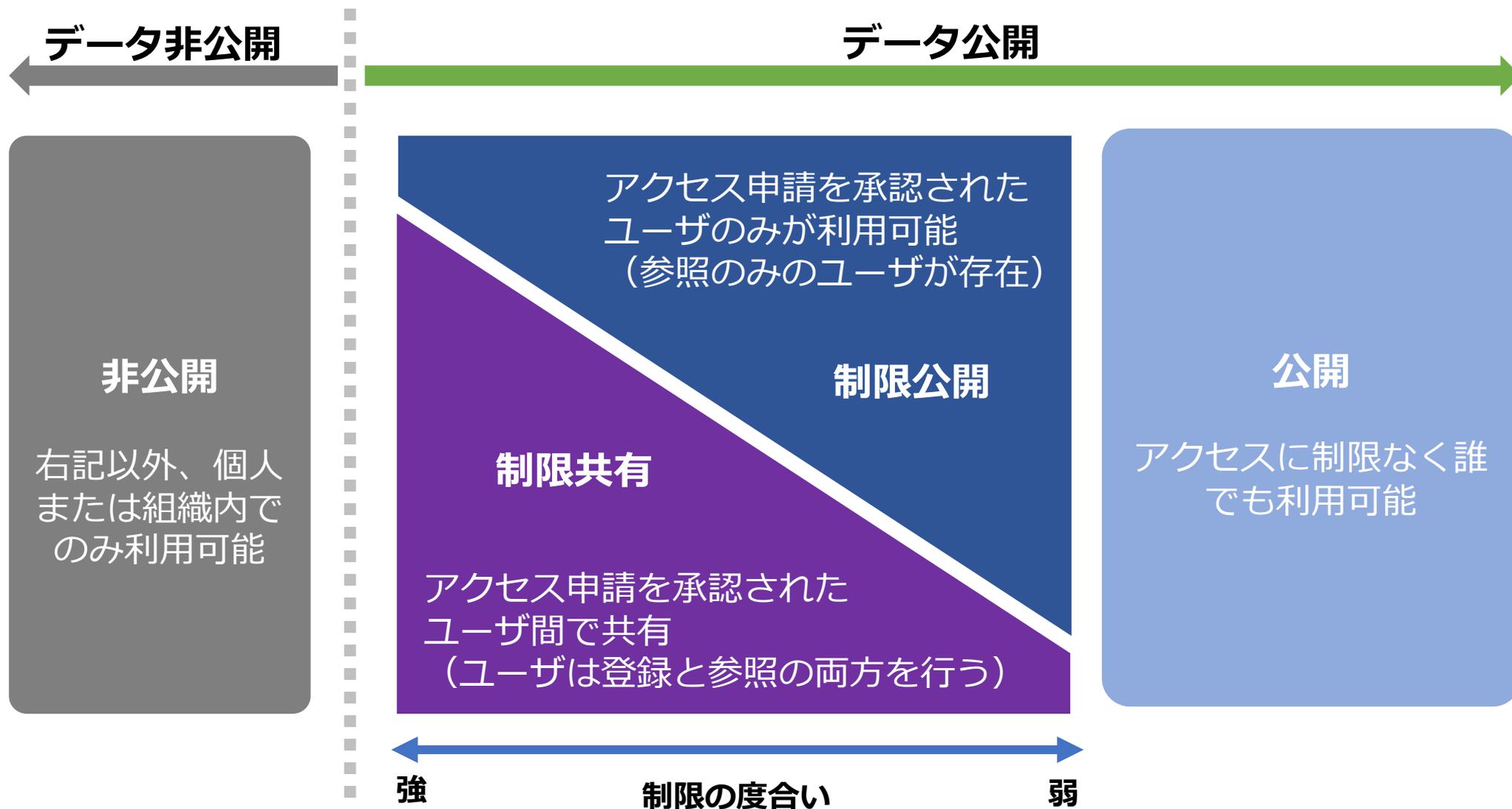


化学・物理学・工学に立脚した応用寄りの分野

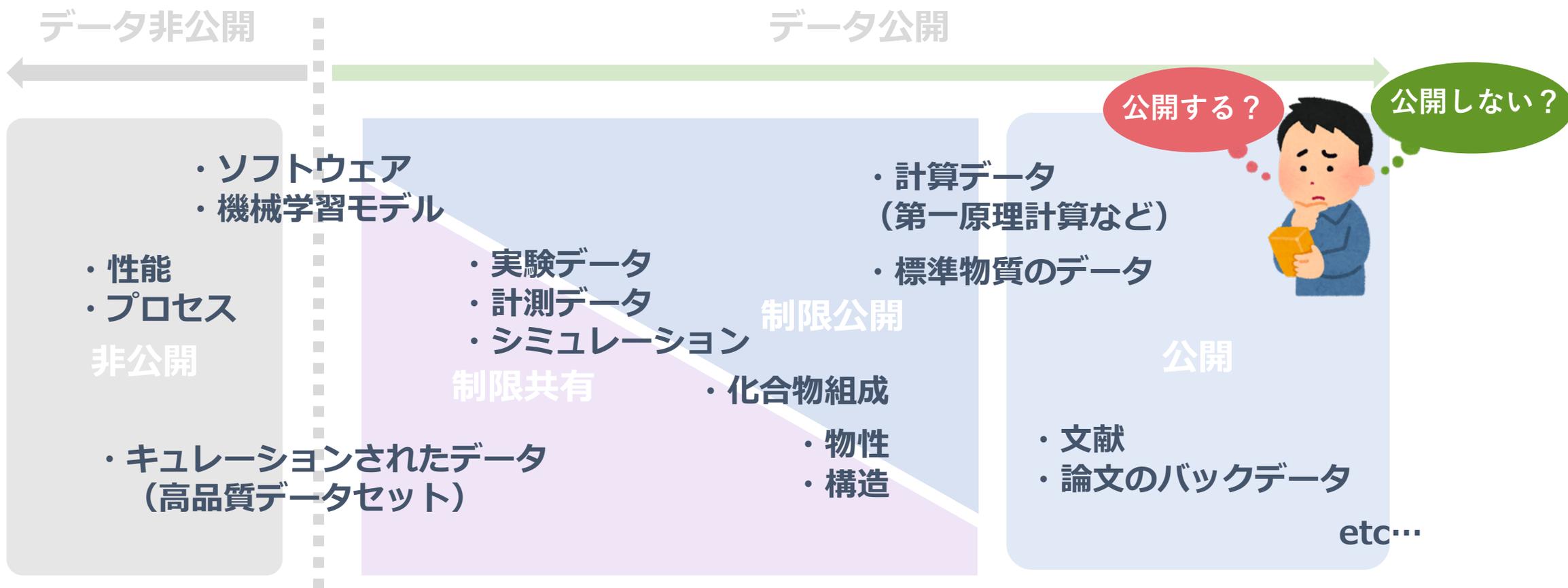
“分化ではなく融合で学術分野が生まれた最初の例” (R. Roy)



材料データは相対的にオープンにされにくい傾向



材料データは相対的にオープンにされにくい傾向

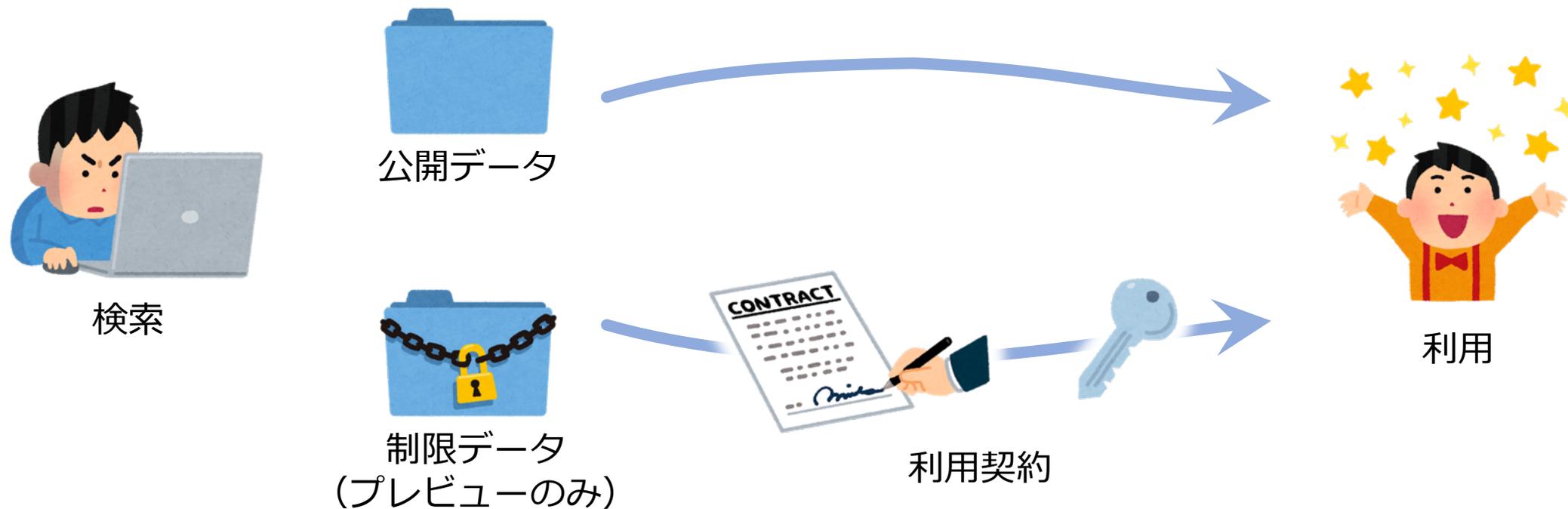


産業界との関連 知的財産としての性質
研究戦略やノウハウに直結しやすい

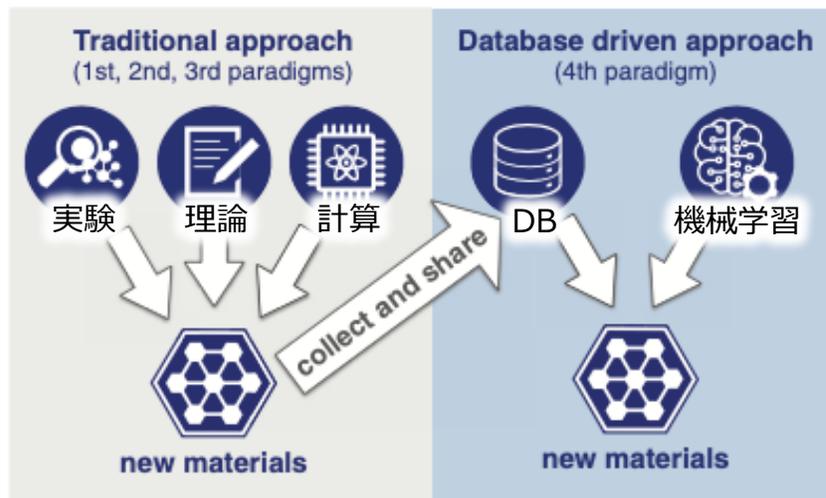


オープンサイエンスの流れ
データ駆動研究にはデータが多く必要

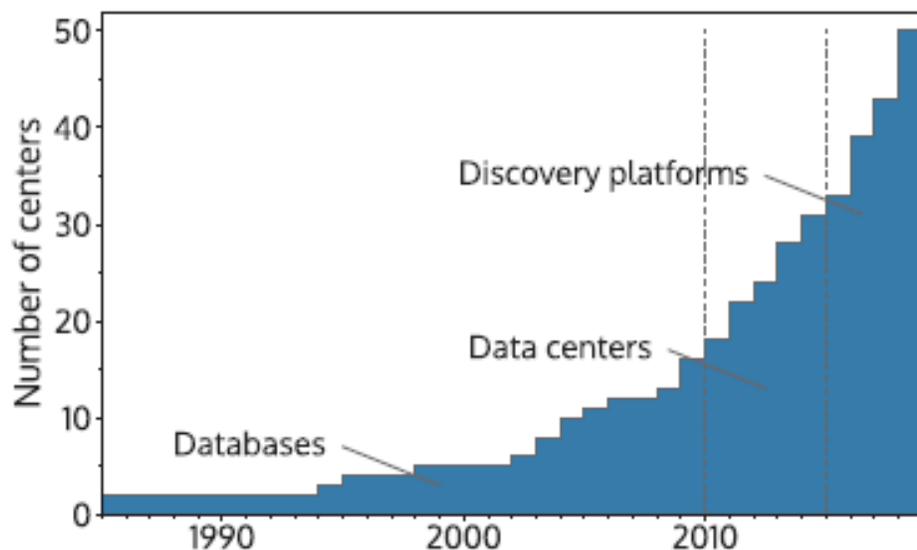
理想：制限データまでも含めたワンストップサービス



現実： システムの仕様として複雑
ステークホルダーが多くルール面でも調整が難しい

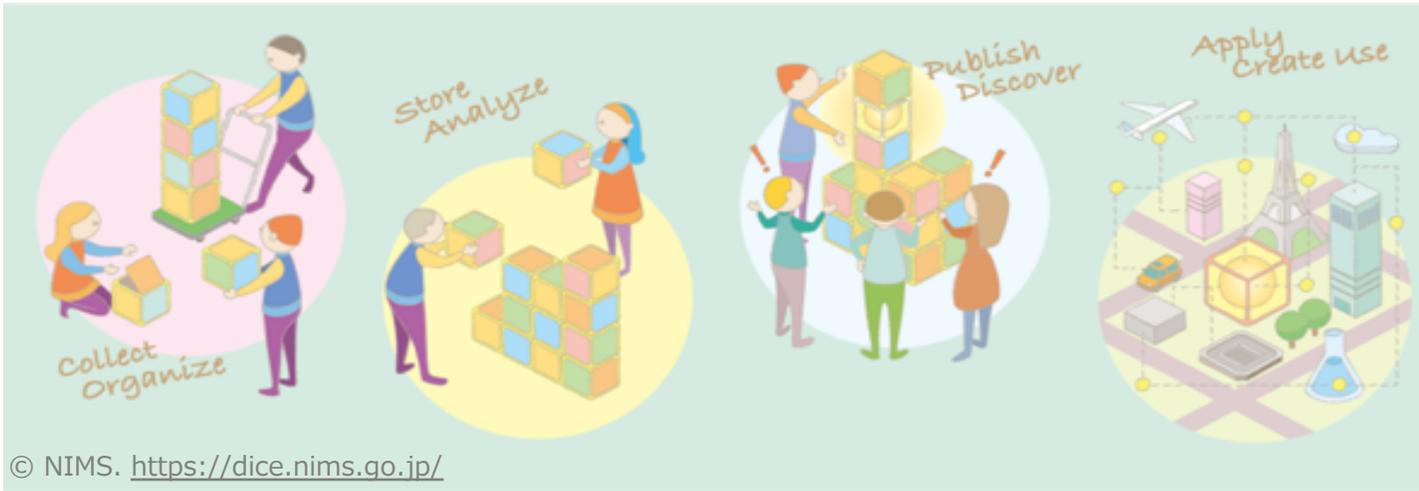


- 第1・2パラダイム = 実験・理論
 - 第3パラダイム = 計算（シミュレーション）
 - 第4パラダイム = データと機械学習
-
- 計算科学的なデータベースが拡充する形で材料データインフラの整備が進んだ
 - データを探し、そこでそのままそれを使い解析する機能など
 - 元素・測定手法・材料特性などのメタデータで整理されることが多い

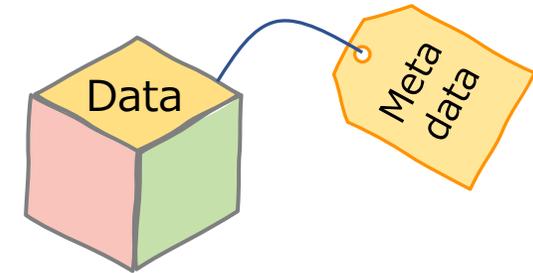


図：いずれも L. Himanen et al. *Adv. Sci.* **6**, 1900808 (2019)
<https://doi.org/10.1002/advs.201900808> より (CC BY)

データプラットフォームDICEと材料メタデータの分類



DICE



材料学的メタデータ

- 対象の試料・物質
- 使った測定法
- 使った計算手法
- 調べた特性
- 試料の合成・プロセス

➡ データの検索・再現性確保に重要だが、分野依存性が高く標準化困難

収集・整理

管理・解析

共有・発見

応用・活用

M-DCaC
Materials Data Conversion Tools

IoT
IoT file transfer system

Research Data Management
Research Data Express

DataVisualizer

MDR
MATERIALS
DATA
REPOSITORY



LabCard
LabNote

書誌情報的メタデータ

- データに関わった人
- 作成日
- 表題
- キーワード

管理・利用メタデータ

- 管理組織
- ライセンス

➡ 共通性が高い
標準化されている



NIMS材料データプラットフォームDICEの共通メタデータ



■ 2列n行の key-value CSV

Software	Gaussian09
Calculation	B3LYP
Basis set	cc-pVDZ

■ Schema.org の拡張, RO-Crate などの JSON-LD



```
{ "@graph": [
  { "@id": "./",
    /* Bibliographic metadata */
    "name": ...,
    "author": ...,
    /* Scientific metadata */
    "variableMeasured": ...,
    "hasPart": { "@id": "data.csv" }
  }, .....
]}
```

※ イメージ

- 高い機械可読性
- 大手検索サービスのサポート
- 分野ごとの標準を定める必要がある (材料分野では未確立)

MDR

表示例

Preview: data.csv

Show 10 entries Search:

Key	Value
Software	Gaussian09
Calculation	B3LYP
Basis set	cc-pVDZ

Showing 1 to 3 of 3 entries Previous 1 Next

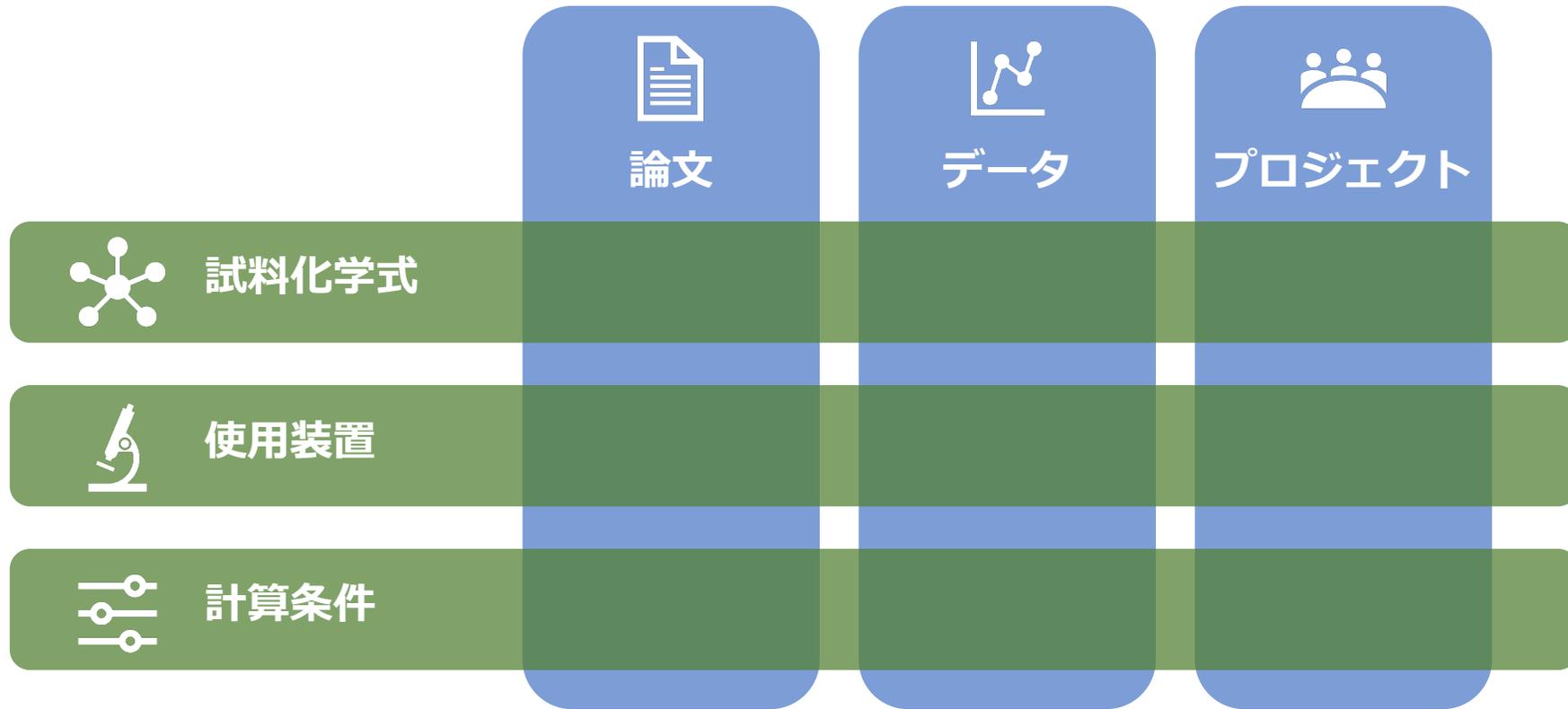
Items

Thumbnail	Title	Date Uploaded	Size	Visibility	Actions
	data.csv	18/10/2020	56 Bytes	NIMS	SELECT AN ACTION

- ユーザによる作成は容易
- Key の語彙統一は期待できない



資料種類による探し方と分野特有の探し方



- 研究者にとってデータの探しやすさは、分野特有の探し方が可能かどうかで決まる。
 - 資料種類による探し方とは直行するもの。
- しかしその定義は困難。特に CiNii のような学際的なサービスではどうすればいいのか。
 - まずとにかくメタデータを記録し、全文検索に頼りたい。定型的な取り扱いは課題。

ぜひ皆様と議論しながら考えていきたい