

- データ創出ー蓄積ー利用ー連携の基盤

2022.2.25

3. データ連携へ向けて「データ基盤共通7課題」

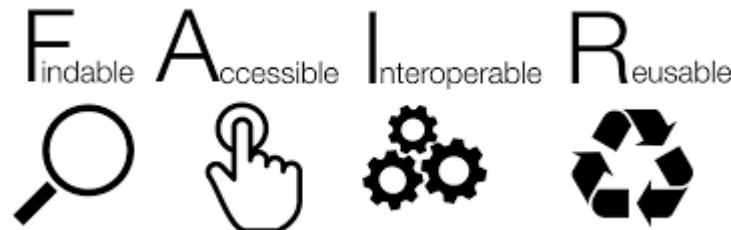




材料データプラットフォームDICE1.0 – 2020 FAIRに始動。



データ利用者の認証



研究データ
収集～高付加価値化

研究データ
蓄積～管理

研究データ
共用～公開

材料データベース
公知データベース

IoT
IoT DATA SYSTEM

M-DaC
Materials Data Conversion Tools

Labnotes

M
MATERIALS
INTEGRATION

データ

RDM

RDE
RESEARCH DATA EXPRESS

MDR Closed

データ

MDR
MATERIALS
DATA
REPOSITORY

リンク

装置等、共通マスターの管理システム

リンク

材料辞書MatVoc

リンク

リンク

MatNavi

PolyInfo
API

TDM

FigResourceMiner
powered by TDM and Data Mining

高速解析クラスター

ハードウェア基盤

VM基盤

アクセスログの統合管理・監視・利用モニタリングシステム



DPF2.0構想 (2019年に考えたこと)

全国利用へ向けた材料データプラットフォーム「データ中核拠点」

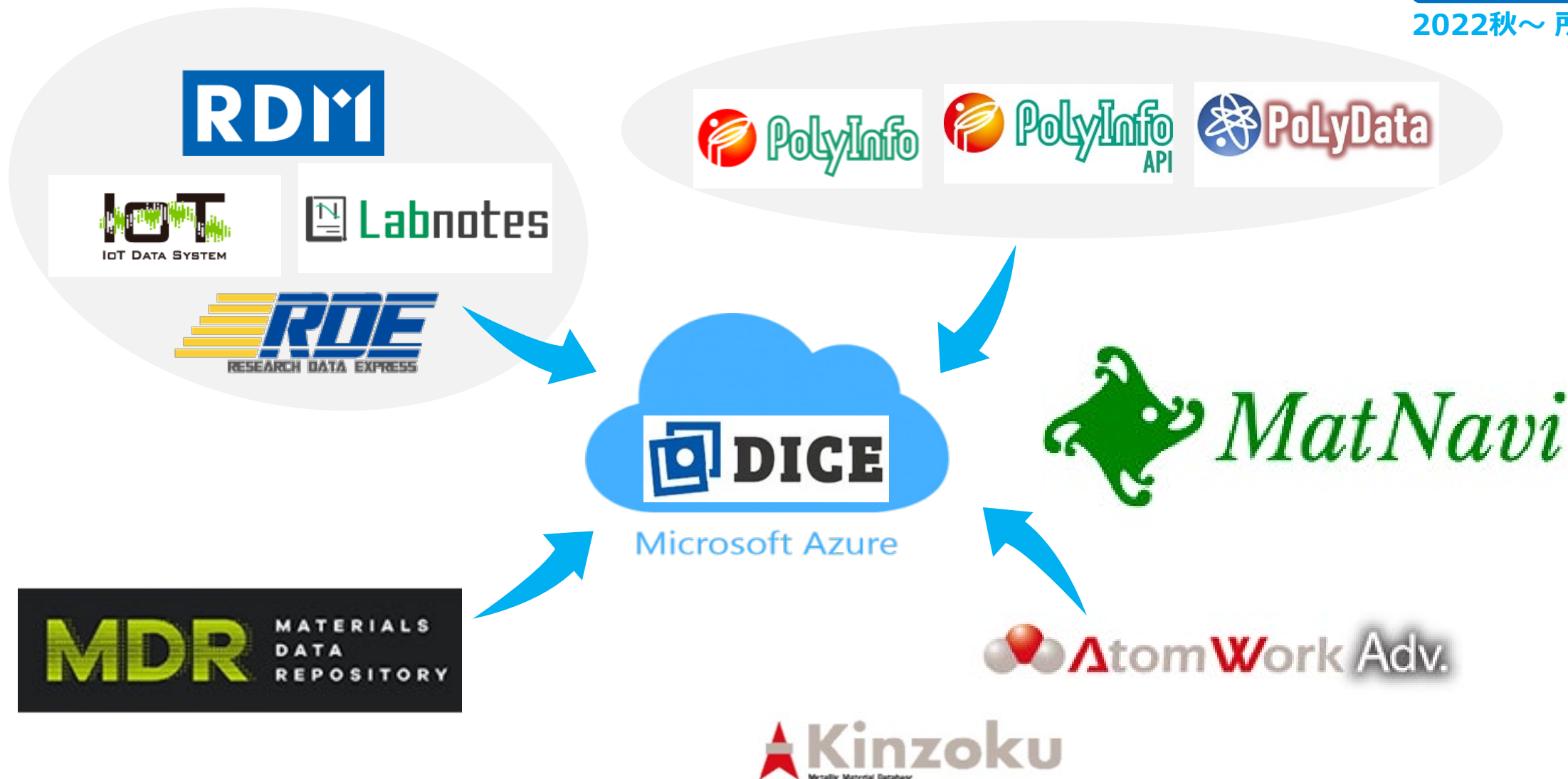
1. NIMS外のDPF利用者を認証(例: 国の研究プロジェクトや事業、企業連携など)
2. DPFデータと自分データを使い、MI解析や可視化など、その場理解を可能とする
3. (2が回る)データサイクルを機能強化し、プラットフォームの価値を高める
4. データのインターネット流通の安全を強化する
5. データ利活用の場はパブリッククラウドに外だしする



材料データプラットフォームDICE2.0 – SINET6・Azure認証



データ中核拠点
2022秋～ 所内試験





まとめ：データ連携へ向けて「データ基盤共通7課題」

DPF1.0実践から2.0取り組みを通して分かったこと。

1. **利用者の認証**: 身元と実在の確認
2. **データの運用**: 管理責任、アクセス権限(認可)、有効期間、期間終了後の運用
3. **データの可読性**: メタデータの共通スキーマと独自記述、機械学習向けデータ形式、データパッケージング方式(メタデータと実データ)、機械学習ツールへの転送
4. **データの転送**: リアルタイムなウイルスチェック、IoT技術、巨大ファイルの圧縮、セキュアな転送
5. **データの共用**: クローズデータ、シェアデータ(広域シェアのための制御・運用)
6. **オープンデータ**: 異なるデータ基盤間のメタ検索と実データローケーション、横断検索のためのRDF
7. **データのオーナーシップ**: データ基盤共通のフレームワーク(著作者、利用者、運用者、伴う費用分担)

SPRING-8 データワークショップ 第2回 2022年2月25日

語彙でつなげる研究データ： 材料語彙MatVocの紹介

松田朝彦, 長尾浩子, ダニエル・ネイションズ, 内藤裕幸, 石井真史

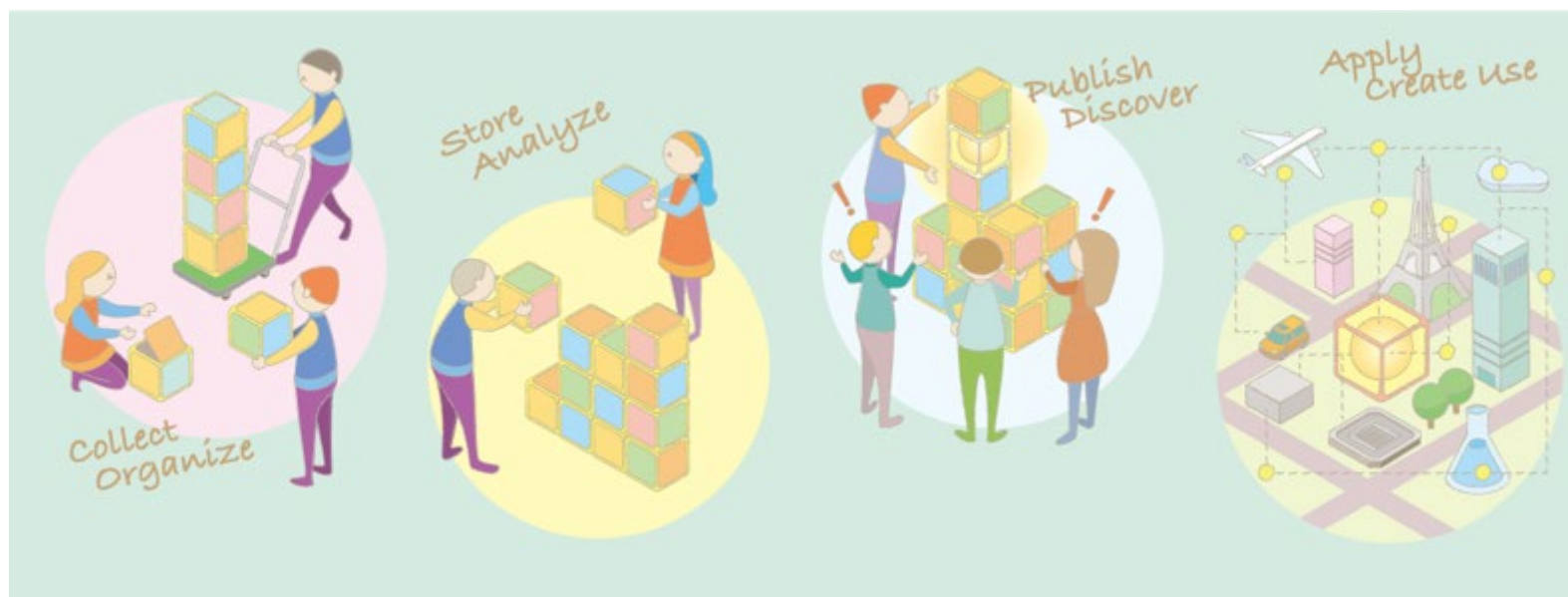
物質・材料研究機構 材料データプラットフォームセンター

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

材料データプラットフォーム “DICE”

- 実験・計測装置、文献等からのデータ取り込み、データ管理・解析環境、データ共有・発見までを扱うシステム
- 材料開発における課題解決の糸口を与えるマテリアルズ・インフォマティクス・サービスを提供

<https://dice.nims.go.jp/>



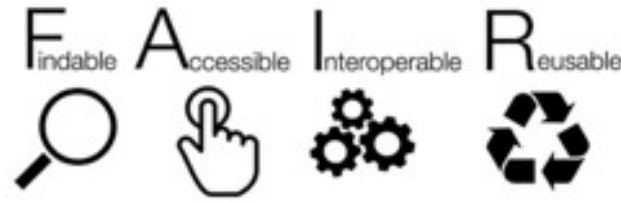
収集
●
整理

管理
●
解析

共有
●
発見

応用
●
活用

データ共有の基準： “FAIR 原則”



データの適切な共有・公開方法を表現した国際的な原則

• Findable: 見つけやすさ

- 一意な ID
- 十分なメタデータ記述
- 検索可能
- データの ID 明記

• Accessible: アクセス性

- 標準プロトコルによるデータ入手
- データが利用不可能となったとしてもメタデータにはアクセスできること

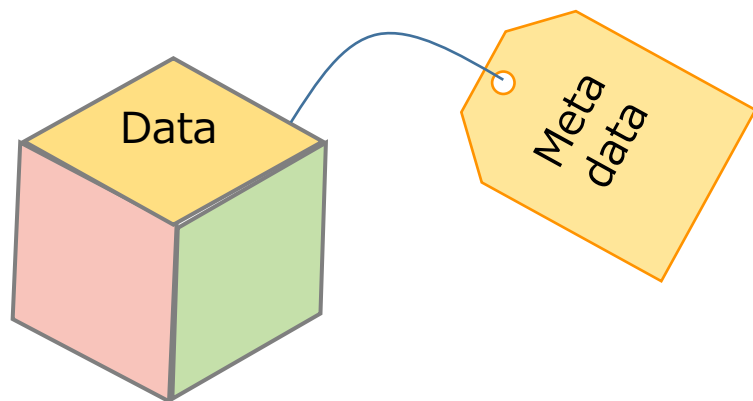
• Interoperable: 相互運用性

- 定形的で、共有されている記述
- 使われる語彙も FAIR であること
- 他のデータへの参照

• Re-usable: 再利用性

- データ利用ライセンス明示
- データ来歴
- 分野ごとのコミュニティの標準を満たすこと

メタデータ: データの“荷札”



書誌情報的メタデータ

- 表題
- 作成者
- 作成日

管理・利用メタデータ

- 管理者・責任者
- ライセンス

➡ **共通性が高い**

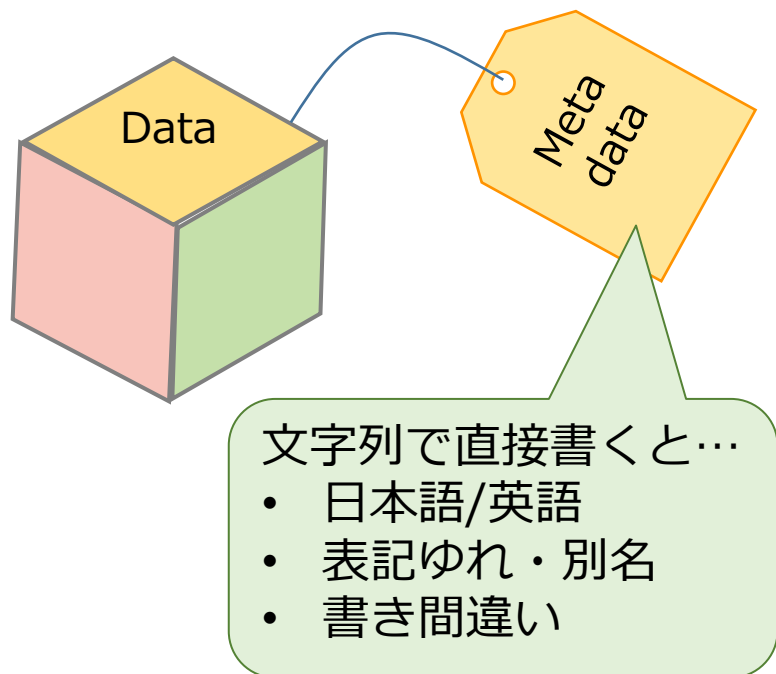
科学的メタデータ

- 対象の**試料・物質**
- 使った**測定法・実験条件**
- 使った**計算手法**
- 調べた**特性**
- 試料の**合成・プロセス**
など

➡ **分野依存性が高い**

Suzuki et al., JVST A **38** 023204 (2020)
“Raw-to-repository characterization data conversion for
repeatable, replicable, and reproducible measurements”
<https://doi.org/10.1116/1.5128408> など

メタデータにおける語彙管理



一ヶ所に概念・用語を登録し、
そこで表記ゆれ・各種情報を
管理することが望ましい

そこで…

- 人間可読ラベルはあくまでも属性のひとつとする
- データを記録するにはこのIDで行うようにする



語彙ID	Q123
日本語主ラベル	ポリエチレン
英語主ラベル	polyethylene
英語副ラベル	polyethene poly(methylene) polythene
化学式	(C₂H₄)_n

IDで語彙を特定し、ラベルや関係性を管理する

MatVoc ID: Q594

Item
Discussion
Read
View history
☆
Search MatVoc

two-dimensional electron/hole gas (Q594)

No description defined [edit](#)

2DHG | 2DEG

▼ In more languages [Configure](#)

Language	Label	Description	Also known as
English	two-dimensional electron/hole gas	No description defined	2DHG 2DEG
Japanese	電子/ホールガス層	工学的構造/電子ガス（ホール）層	電子ガス（ホール）層 二次元電子/ホールガス

Statements

has broader

engineered structures [edit](#)

▼ 1 reference

source item
DPFC expansion of the Materials Data Vocabulary

+ add reference

+ add value

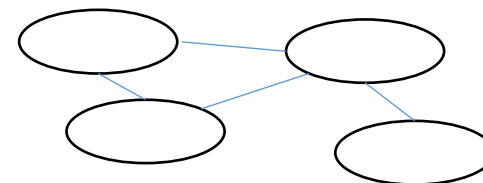
+ add statement

自然言語での記述、機械可読な語彙項目としての記述

百科事典記事



姉妹
プロジェクト



概念・語彙項目



使用ソフトウェア
(オープンソース)



星月夜

[編集 | ソースを編集] 3 users

出典: フリー百科事典『ウィキペディア (Wikipedia)』

星月夜（ほしづきよ、**フランス語**: *La nuit étoilée*、**オランダ語**: *De sterrennacht*、**英語**: *The starry night*）は、**オランダ**の画家・**ゴッホ**の代表作のひとつ。1889年6月、フランス**サン＝レミ＝ド＝プロヴァンス**のサン＝ポール・ド・モゾル修道院の精神病院で療養中に描かれた。1941年、**リリー・P・ブリス**遺贈。**ニューヨーク近代美術館**の永久コレクションである。

この絵は、フランス人作曲家**アンリ・デュティユー**の管弦楽曲「**音色、空間、運動** (Timbres, Espace, Mouvement)」にインスピレーションを与えた。また、この絵は、2011年公開の**ウディ・アレン**監督の映画『**ミッドナイト・イン・パリ**』のポスターに部分的に使用されている^[1]。

Labels

en	Starry Night	painting by van Gogh
ja	星月夜	ゴッホの絵画作品

Statements

P31 (instance of)	絵画作品 (Q3305213)
P170 (creator)	ゴッホ (Q5582)
P571 (date of creation)	June 1889
P195 (collection)	ニューヨーク近代美術館 (Q188740)

Q45585
星月夜

P31:
"instance of"

Q3305213
絵画作品

P170:
"creator"

Q5582
ゴッホ

P571:
"date of creation"

June 1889

P18:
"image"

(URL)



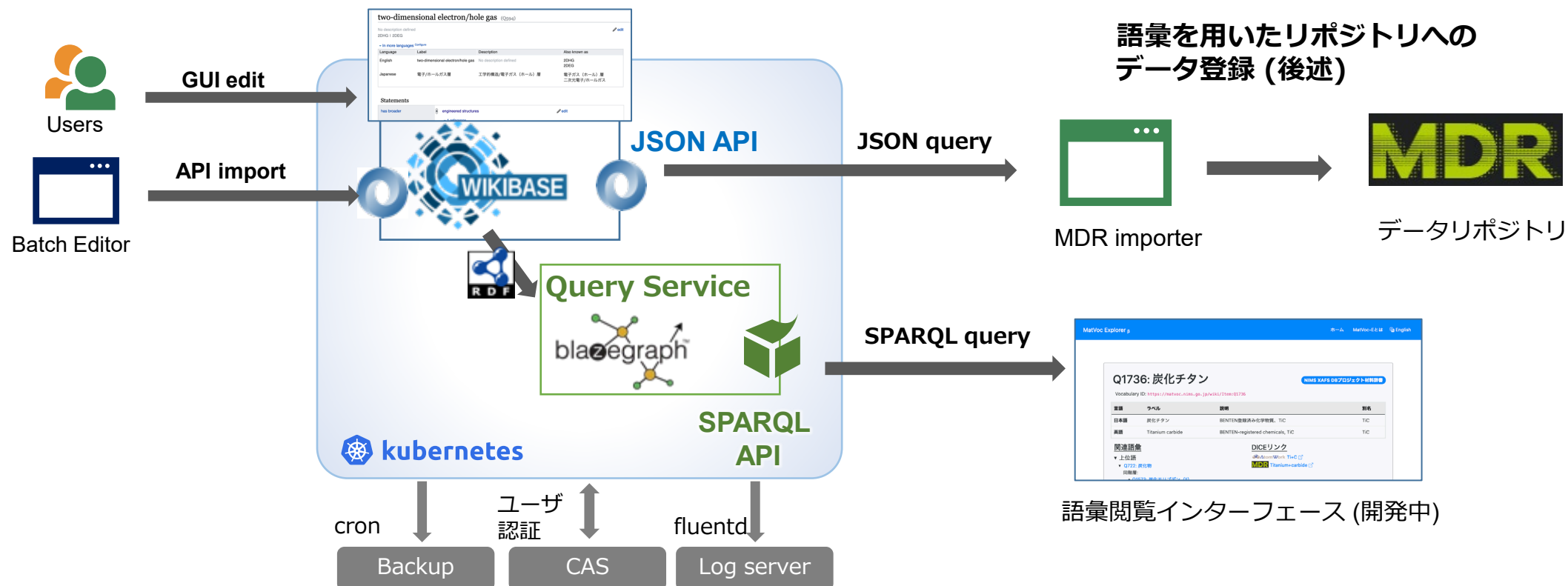
NIMSプラットフォーム上に展開した語彙システム

現時点では所内公開

語彙管理システム “MatVoc”

他システムへの語彙配信

語彙を用いたリポジトリへの
データ登録 (後述)



松田, 内藤, 門平「材料データプラットフォームのための協働型材料語彙システム展開」
電子情報通信学会 SC研究会 (2020)

NIST材料語彙

<https://doi.org/10.18434/T4/1435037>

- Materials Resource Registry

<https://materials.registry.nist.gov/>

を検索するための分類

- 最上位概念:

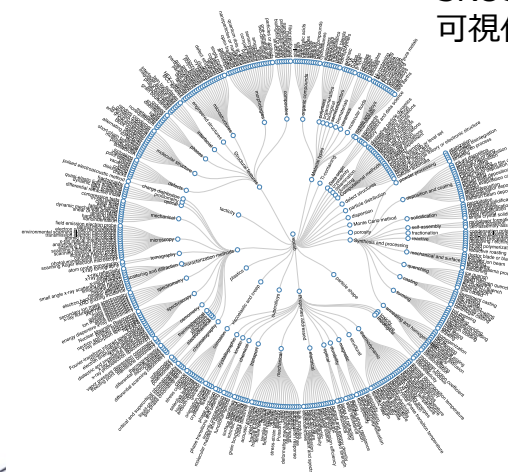
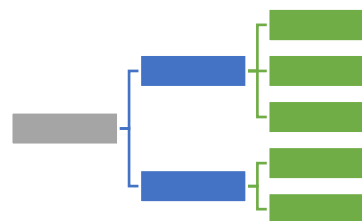
- Data origin
- Characterization methods
- Computational methods
- Material types
- Properties addressed
- Structural features
- Synthesis and processing

- そこから 2 レベル、下位に分岐する

配布されているExcelファイル

	A	B	C
1	This vocabulary describes the applicability to material science of records in the NIST Materials Resource Registry (NMRR). The NMRR allows for the registration of materials resources, bridging the gap between existing resources and the end users. The NMRR functions as a node in a federated system, making the registered information available for research to the materials		
2	This is being developed at the National Institute of Standards and Technology and is made available to solicit comments from the Material Science community. Please do not enter any proprietary data into this system.		
3	version 16, 7/13/17		
4	Feedback to Chandler Becker, cbecker@nist.gov		
5	See below for additional contributors and people/resources consulted		
6			
7			
8	Data origin	experiments	
9	Data origin	informatics and data science	
10	Data origin	simulations	
11	Data origin	theory	
12			
13	Material types	biological	.
14	Material types	biomaterials	.
15	Material types	ceramics	.
16	Material types	ceramics	carbides
17	Material types	ceramics	cements
18	Material types	ceramics	nitrides
19	Material types	ceramics	oxides
20	Material types	ceramics	perovskites
21	Material types	ceramics	silicates
22	Material types	metals and alloys	.
23	Material types	metals and alloys	Al-containing
24	Material types	metals and alloys	commercially pure metals
25	Material types	metals and alloys	Cu-containing
26	Material types	metals and alloys	Fe-containing
27	Material types	metals and alloys	intermetallics
28	Material types	metals and alloys	Mg-containing
29	Material types	metals and alloys	Ni-containing
30	Material types	metals and alloys	rare earths
31	Material types	metals and alloys	refractories

配布されている
SKOSファイルを
可視化した図



MatVoc 上での語彙の整理とつながり

RDF (Resource Description Framework)

主語 - 述語 - 目的語

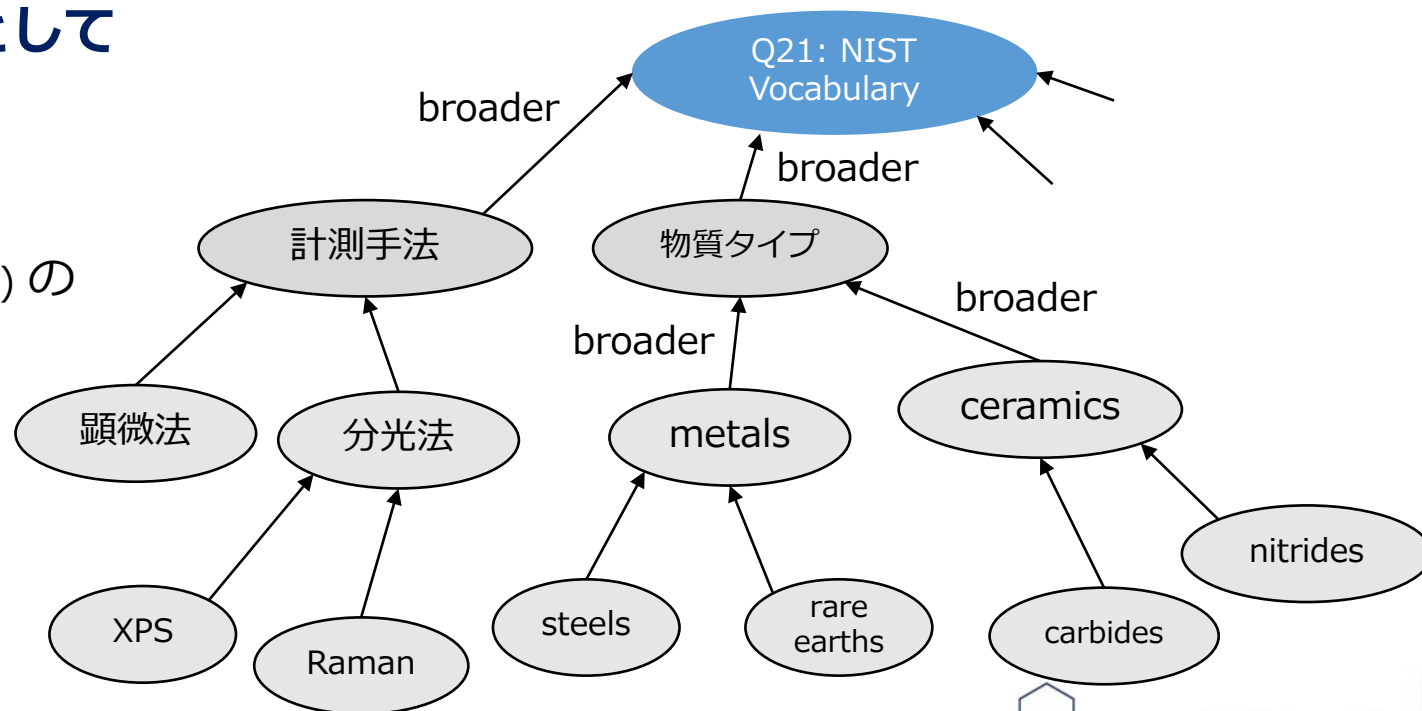
主語 - 述語 - 目的語



NIST語彙をRDF互換の有向グラフとして MatVocに格納

述語は、NIST原資料に沿い
SKOS (Simple Knowledge Organization System) の
“broader” (広義語) を主に用いた

SPARQLという言語でクエリ可能に



語彙利用事例: “MDR XAFS DB”

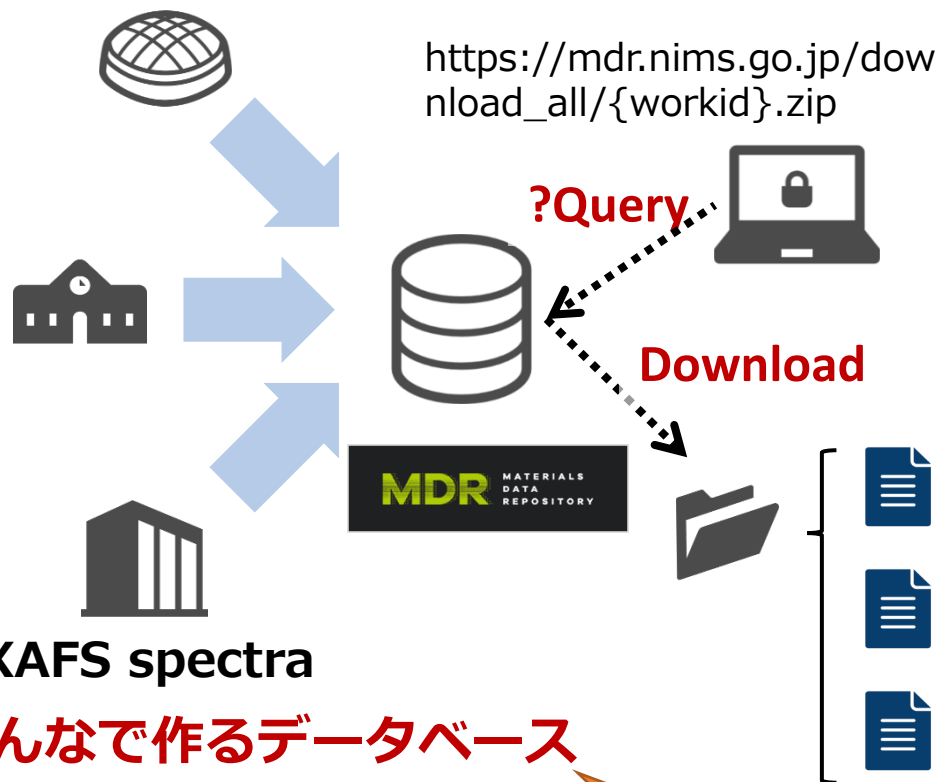
MDR XAFS DB:

<https://doi.org/10.48505/nims.1447>

公開中!



XAFS spectra



XAFS spectra

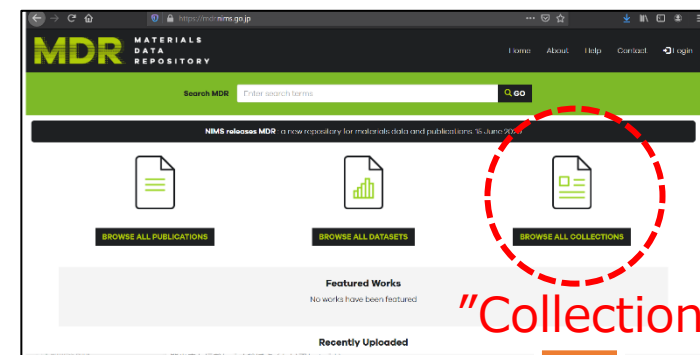
みんなで作るデータベース

- SPring-8 (HX XAFS)
- PF (HX XAFS)
- 立命館大学 (SX XAFS)
- 北海道大学
- あいちSR 他

Machine readable
and accessible



<https://mdr.nims.go.jp/>



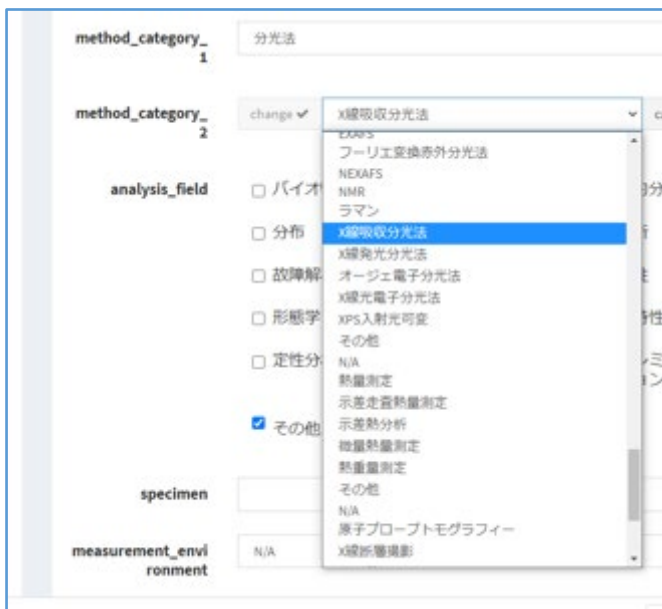
"Collections"



MDR XAFS DB
1,757 samples

石井「データサイエンスにおけるSPring-8の役割」
SPring-8 Symp. 2021

XAFSデータ登録時の MatVoc ID によるメタデータ移送



共通メッセージ形式

```
1,
"methods-category": [
  {
    "analysis-field-description": "",
    "analysis-field-items": [
      "0688",
    ],
    "energy-level-transition-structure": [
      "Cu K-edge",
    ],
    "measurement-environment": "0542",
    "measurement-environment-description": "",
    "method-category": [
      {
        "method-maincategory-code": "030",
        "method-subcategory-code": "1386",
      },
    ],
    "reference-source": [
```

材料データリポジトリ



Method	
Characterization methods	
spectroscopy --> x-ray absorption spectroscopy	
Instruments	
Instrument	
Title	BL14B2_XAFS
Description	SPring-8 産業利用ビームライン XAFSセットアップ
Instrument function	
Category	spectroscopy
Sub category	x-ray absorption spectroscopy
Manufacturer	
Organization	Japan Synchrotron Radiation Institute
Managing organization	
Organization	Japan Synchrotron Radiation Institute
Specimen details	
Specimen type	
Title	Copper

API問い合わせ

MatVoc ID: Q386

x-ray absorption spectroscopy (Q386)

参照

同期

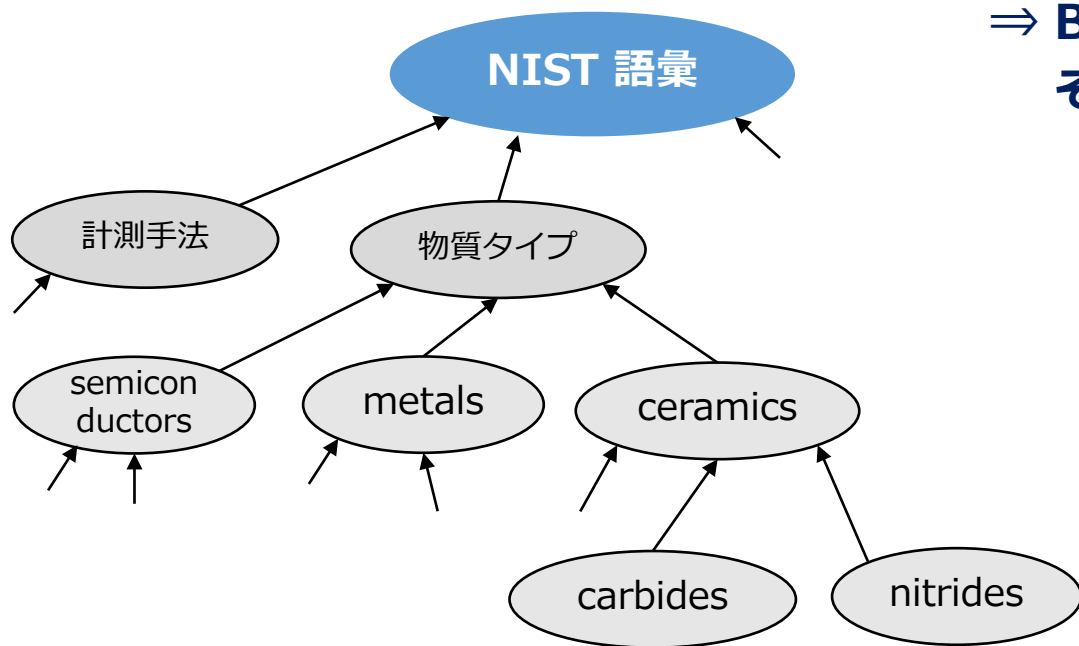
MatVoc



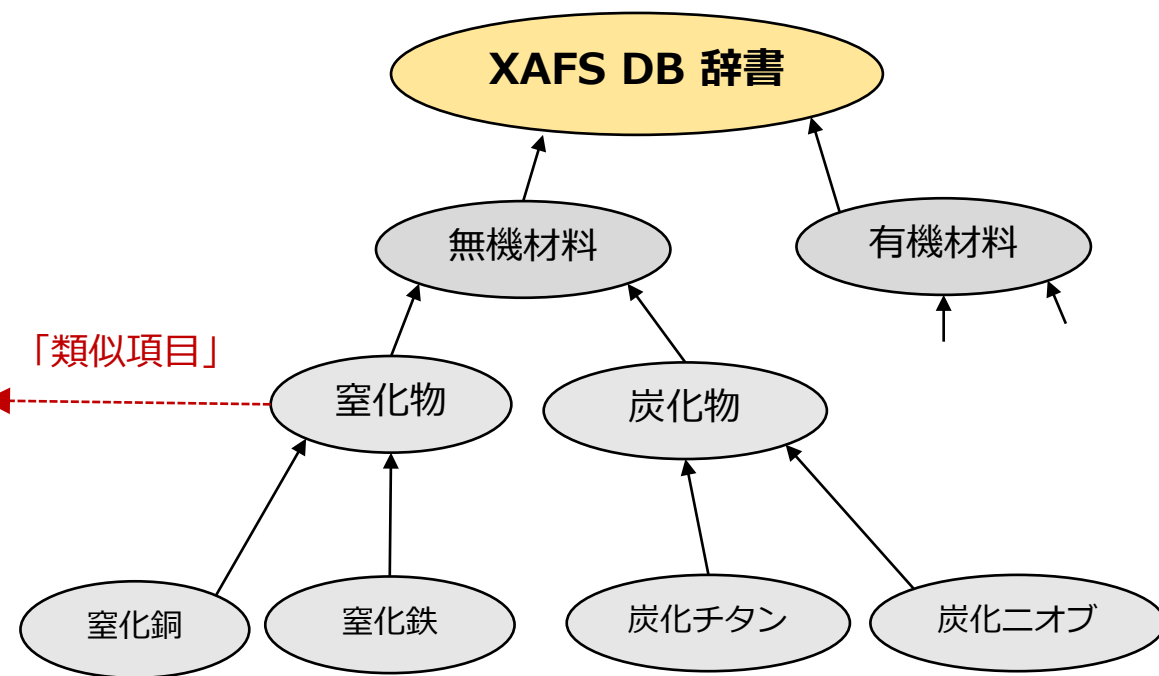
システム辞書

MatVoc 上に作った、XAFS DB のための語彙

NIST語彙の物質タイプ分類はBENTENのものとは合わない
⇒ **BENTENの分類に寄せた「XAFS DB 辞書」を整備、そのIDでメタデータを記録**



分類の下層側で一致する概念があれば、
「類似項目」として辞書横断リンク



MatVoc を核にした NIMS リソース連携

石井「データサイエンスにおけるSPring-8の役割」
SPring-8 Symp. 2021

 **Q1286: 塩化銀 (I) 無水**  **NIMS XAFS DBプロジェクト材料辞書**




語彙ID: <https://matvoc.nims.go.jp/wiki/Item:Q1286>


言語	ラベル	説明	別名
日本語	塩化銀 (I) 無水	BENTEN登録済み化学物質、AgCl	AgCl
英語	Silver(I) chloride, anhydrous	BENTEN-registered chemicals, AgCl	AgCl

関連語彙

- ▼ 上位語
 - ▶ Q725: 塩化物

DICEリンク

 **AtomWork** [Ag+Cl](#) 
MDR [Silver\(I\)+chloride%2C+anhydrous](#) 

 **AtomWork** [Search phase diagrams](#) [Search materials](#)

[Home](#) > [Search materials](#) > List of substances


Search materials - List of found substances

Results 1-2 of 2 for Chemical system: Ag Cl


Find materials that have ...

No.	Chemical formula	Substance name	Number of elements	Phase identifier			
				Structure type (Prototype)	Pearson symbol	Space group	Space group number
1	AgCl		2	CsCl	cP2	Pm-3m	221
2	AgCl		2	NaCl	cF8	Fm-3m	225

MDR MATERIALS DATA REPOSITORY [Home](#)

Search MDR 

Limit your search

Filtering by: **KEYWORD > SILVER CHLORIDE, ANHYDROUS** 

1 - 2 of 2 [SORT BY RELEVANCE](#) [10 PER](#)

Keyword

Ag K-edge 2
 AgCl 2
 BL14B2 2
 Chloride 2
 SPring-8 2
 more >

Description/Abstract: This dataset consists of X-ray absorption fine structure (XAFS) spectra at Ag K-edge of Silver(I) chloride, anhydrous measured at SPring-...

Keyword: Ag K-edge, AgCl, BL14B2, Chloride, SPring-8, Si(31), Silver(I) chloride, anhydrous, XAFS,

Resource type >

MatVoc を核にした NIMS 外 リソース連携

石井「データサイエンスにおけるSPRING-8の役割」
SPRING-8 Symp. 2021

 **Q1276: クロム酸銀 (I)**  NIMS XAFS DBプロジェクト材料辞書

語彙ID: <https://matvoc.nims.go.jp/wiki/Item:Q1276>

言語	ラベル	説明	別名
日本語	クロム酸銀 (I)	BENTEN登録済み化学物質、Ag ₂ CrO ₄	Ag ₂ CrO ₄
英語	Silver(I) chromate	BENTEN-registered chemicals, Ag ₂ CrO ₄	Ag ₂ CrO ₄


関連語彙
▶ 上位語

DICEリンク

 AtomWork [Ag+Cr+O](#) 
 MDR [Silver\(I\)+chromate](#) 

外部リンク

 [Silver_chromate](#) 
 [Q416277](#) 

 **WIKIPEDIA**
The Free Encyclopedia

Main page
Contents
Current events
Random article
About Wikipedia
Contact us
Donate

Contribute
Help
Learn to edit
Community portal
Recent changes
Upload file

Article Talk

Silver chromate

From Wikipedia, the free encyclopedia

Silver chromate is an inorganic compound with formula Ag₂CrO₄ which appears as distinctively coloured brown-red crystals. The compound is insoluble and its [precipitation](#) is indicative of the reaction between silver chromate and silver precursor salts (commonly [potassium/sodium chromate](#) with [silver nitrate](#)).^{[2][4][9]} This reaction is important for two uses in the laboratory: in [analytical chemistry](#) it constitutes the basis for the [Mendeleev method of argentometry](#),^[6] whereas in [neuroscience](#) it is used in the [Golgi method](#) of staining neurons for microscopy.^[7]

In addition to the above, the compound has been tested as a photocatalyst for [wastewater treatment](#).^[4] The most important practical and commercial application for silver chromate, however, is its use in Li-Ag₂CrO₄ batteries, a type of [lithium battery](#) mainly found in [artificial pacemaker](#) devices.^[8]

As for all [chromates](#), which are [chromium\(VI\)](#) species, the compound poses a hazard of toxicity, [carcinogenicity](#), and [genotoxicity](#), as well as great environmental harm.



 **WIKIDATA**

Main page
Community portal
Project chat
Create a new item
Recent changes
Random item
Query Service
Nearby

Item Discussion

silver(I) chromate (Q416277)


chemical compound
silver chromate

▼ In more languages
Configure

Language	Label	Description
English	silver(I) chromate	chemical compound
Japanese	クロム酸銀(I)	No description defined


その先、概念による連携 (Ontology)


石井「データサイエンスにおけるSPring-8の役割」
SPring-8 Symp. 2021





[Home](#)
[Ontologies](#)
[Documentation](#)
[About](#)

[XAFS/Chemical Method Ontology:](#)
<https://github.com/rsc-ontologies/rsc-cmo>

[OLS / Chemical Methods Ontology](#)
CHMO / CHMO:0000299
 Copy

XAFS (OLS ID: 299)
 JSON

X-ray absorption fine structure spectroscopy

 http://purl.obolibrary.org/obo/CHMO_0000299
 Copy

Tree view Term mappings Term history

CHMO

Chemical Method Ontology

- process
 - planned process
 - assay
 - spectroscopy**
 - X-ray spectroscopy
 - X-ray absorption spectroscopy
 - X-ray absorption fine structure spectroscopy
 - extended X-ray absorption fine structure spectroscopy
 - quick extended X-ray absorption fine structure spectroscopy
 - reflection extended X-ray absorption fine structure spectroscopy
 - surface extended X-ray absorption fine structure spectroscopy
 - near-edge X-ray absorption fine structure spectroscopy
 - total-reflection X-ray absorption fine structure spectroscopy

Graph view Reset tree Show all siblings

上位概念經由で sibling連携

上位概念經由で sibling 連携

最後に

- 研究データを取り扱うにあたって、
データの見つけやすさ・アクセス性・相互運用性・再利用性が重要
- **メタデータ記述の際の語彙管理**はその重要な要素
 - 「この表現/分類が正しい」(他は誤りか?) という話にならないよう、
柔と剛を併せ持つ「管理」が必要
- DICEではソフトウェアにWikibaseを採用して共同編集、
ポリシーとして**様々な記述** (異なる表記・異なる分類) を受入、**リンク**
 - 辞書を横断し、項目単位でのリンクを張り結びつけ
- **データセット/データベース単位**での語彙管理
 - ↔ **プラットフォームとして**の語彙管理
 - ↔ **内外リソースとの連携** 標準規格、外部オントロジーなど

データプラットフォームにおけるデータ活用のための 可視化・構造化・モデル化

物質・材料研究機構（NIMS）
統合型材料開発・情報基盤部門（MaDIS）
材料データプラットフォームセンター（DPFC）
吉川英樹

NIMS 材料データプラットフォーム(DPF)の目的と課題

目的

- ① 蓄積した文献／他者／自己の過去データ
- ② 自己の新データ
- ③ AIを志向して開発した文献／実験のデータ解析ツール



を三位一体とする**データ駆動型の材料研究**を実施する仕組み作り

上記目的の実現のためには

第3者が容易にデータを検索/抽出/分析できる形式でデータを蓄積する必要がある



データの構造化が必要

AI活用を目指した
データ構造化の階段

明快なライセンスでデータが流通することでビッグデータ化し、それよりデータセットを抽出／分析／活用するツールやアプリが豊富にある



特徴量（物質名, 組織構造, 特性・計測値, 合成・加工・計測条件等）が必要に応じて可視化やモデル化をされ共通の形式でデータに特記される



数値行列を含むデータ全体が分野毎に共通のフォーマットで記述される



データ全体が分野毎に共通の用語（語彙）で記述される

個体識別のIDのメタデータがついて同一系統のデータが検索できる

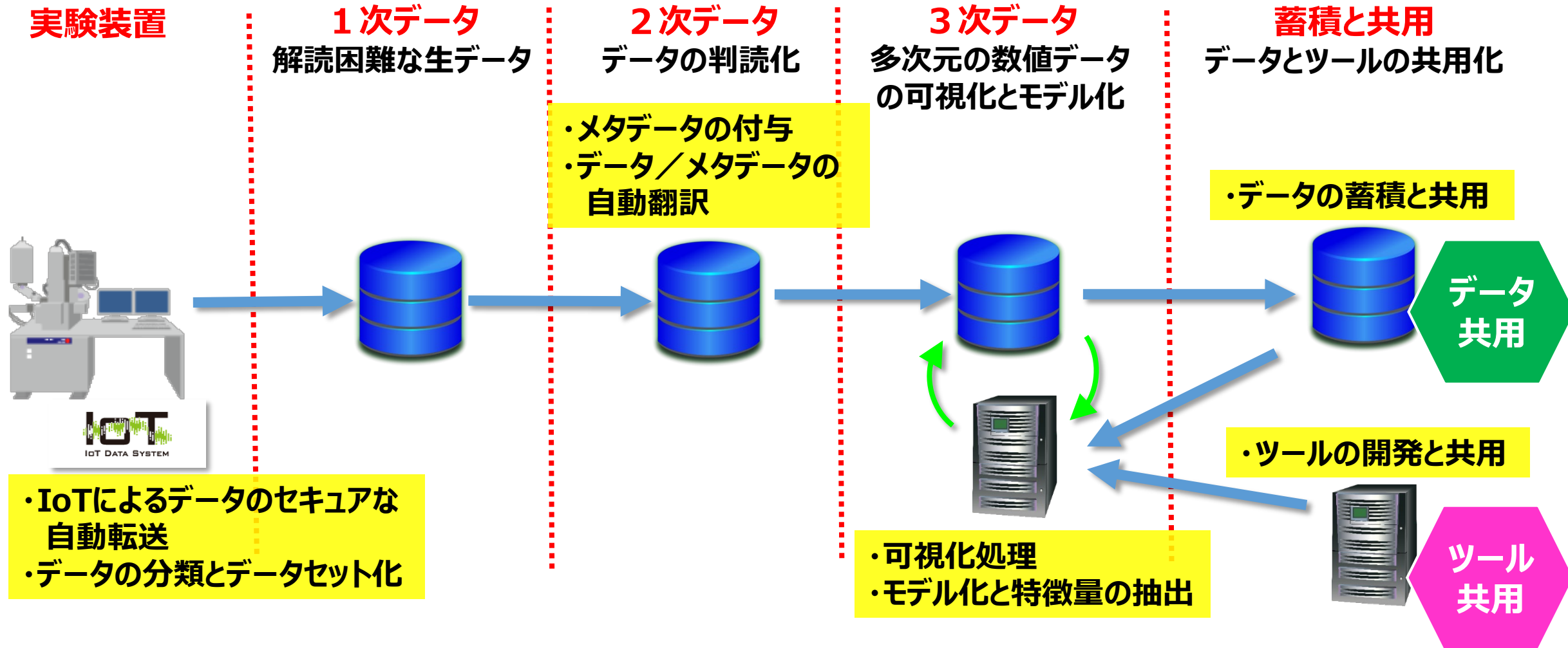


キーワード等の語彙のメタデータがついて類似のデータが検索できる

デジタルデータとしてどんな機械でも読める



実験データの構造化のワークフローの概念図



データ構造化に際してのフォーマットの2つの考え方

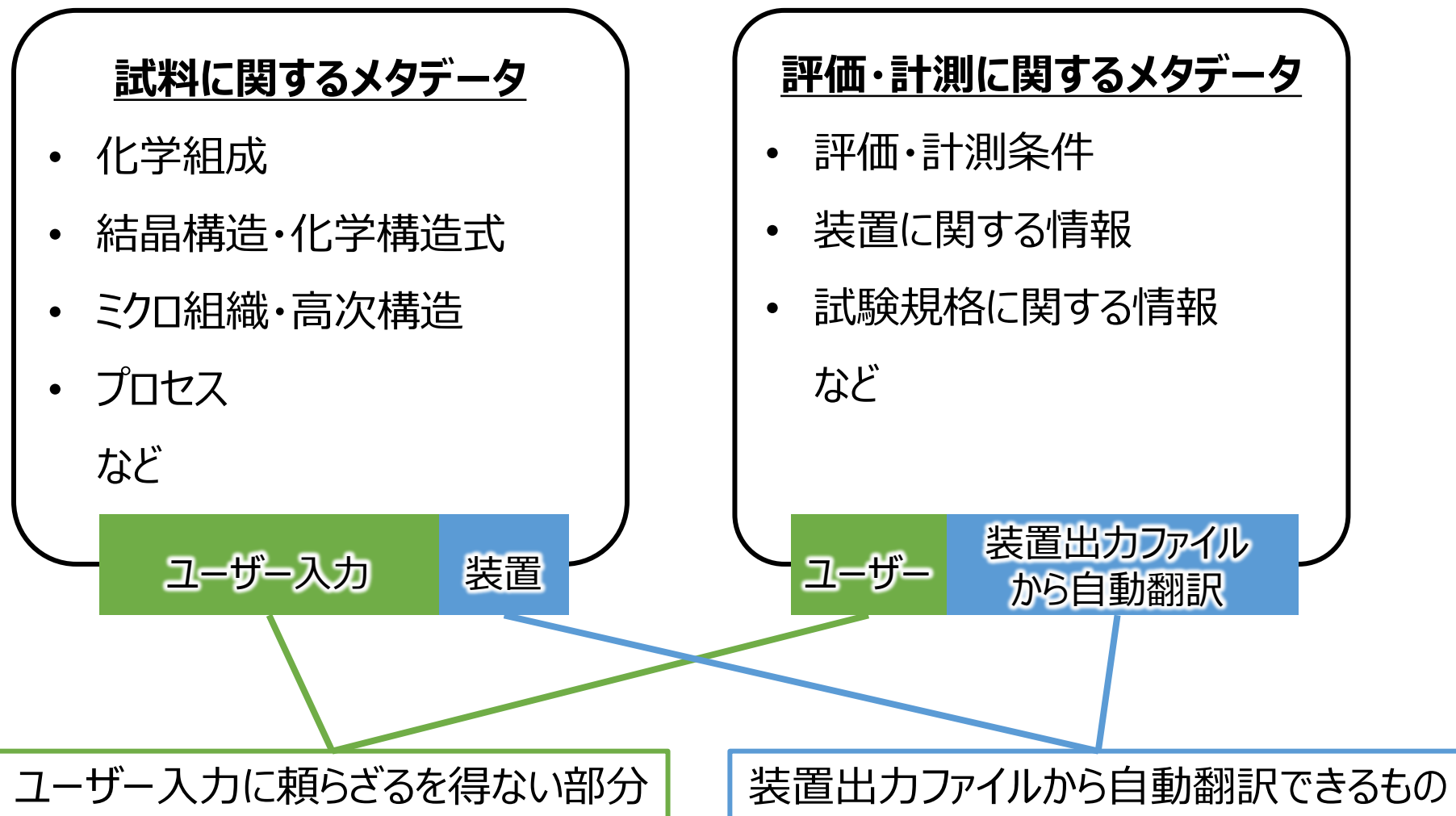
★ Schema-on-Read 型のデータ登録

- ✓ データフォーマットは、データのWrite 時には一つに限定せず多数個の存在を許容
- ✓ データの判読性の確保のため、データフォーマット情報とセットでデータを記録
- ✓ 後日データを Read する際に改めてデータフォーマットを整備することを想定

Schema-on-Write 型のデータ登録

- ✓ 標準化された一つのデータフォーマットをデータの Write 時に使用

データ構造化の基本方針 — データ登録者の負担をできるだけ減らすために



スマートラボトリ

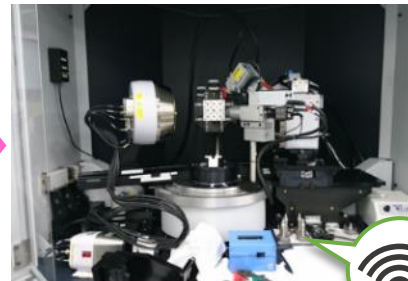
材料の合成から分析までの工程のデータをためて総合評価し材料研究を加速



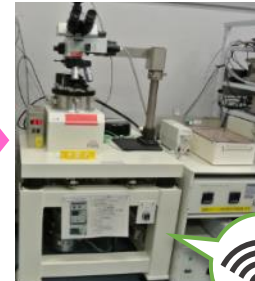
成膜装置:
新材料の合成



XRF:
元素組成の分析



XRD:
結晶構造の分析



AFM:
表面形状の分析



XPS:
化学種の分析



(古いOSで動作しているPCを含む)
実験装置制御PCをネットワークに接続
してデータを安全に収集するための
IoT情報セキュリティデバイス
現在の所内の設置装置台数 123台

無線タイプ



有線タイプ

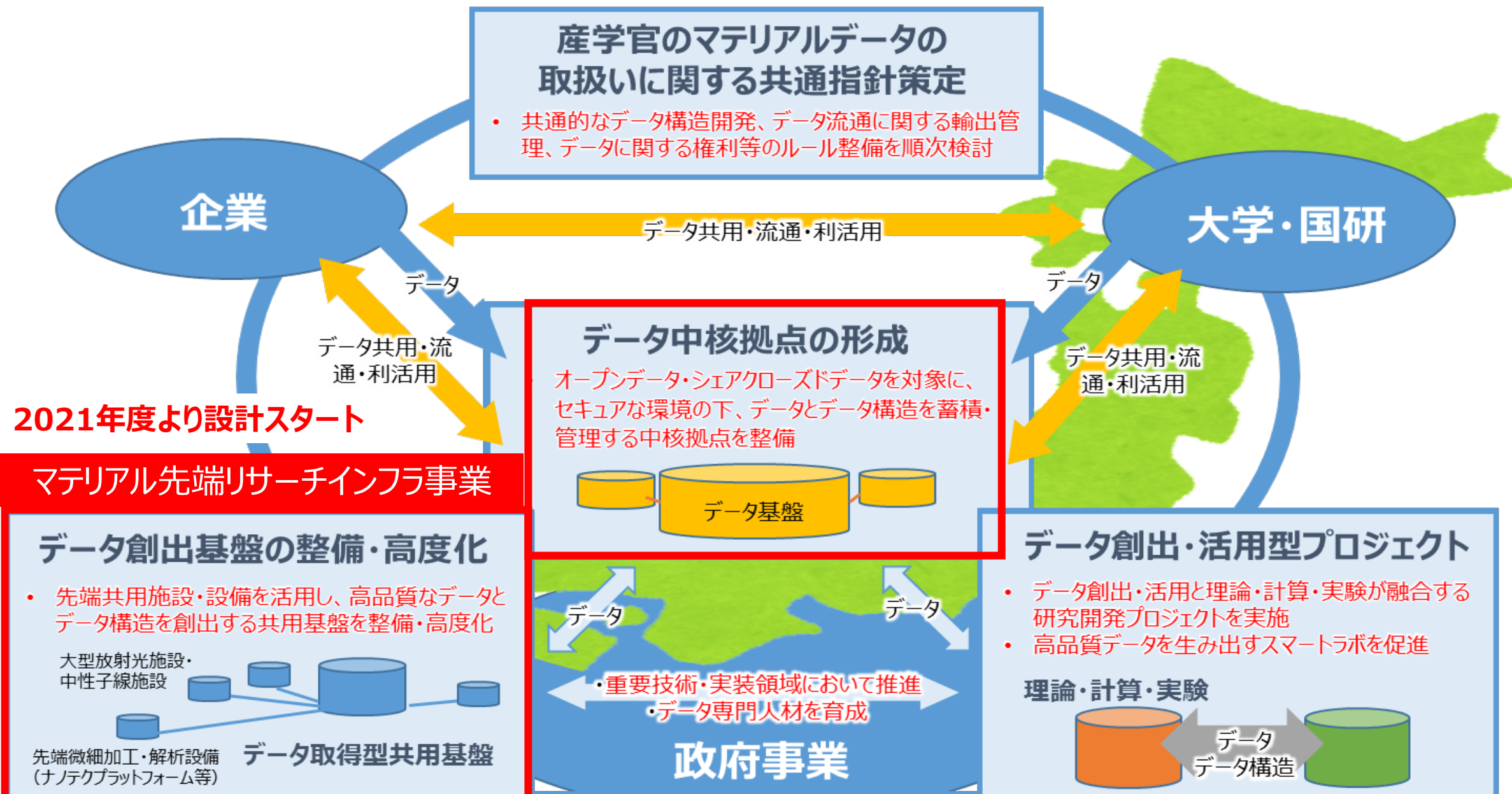


データの収集と構造化のシステム

- 材料の合成から分析までの一貫した各種実験装置のデータを自動で収集し、自動翻訳およびデータ形式を整えてためる
- ためたデータを総合評価し、材料合成プロセスを最適化

文科省の新事業「マテリアル先端リサーチインフラ事業」 の概要説明

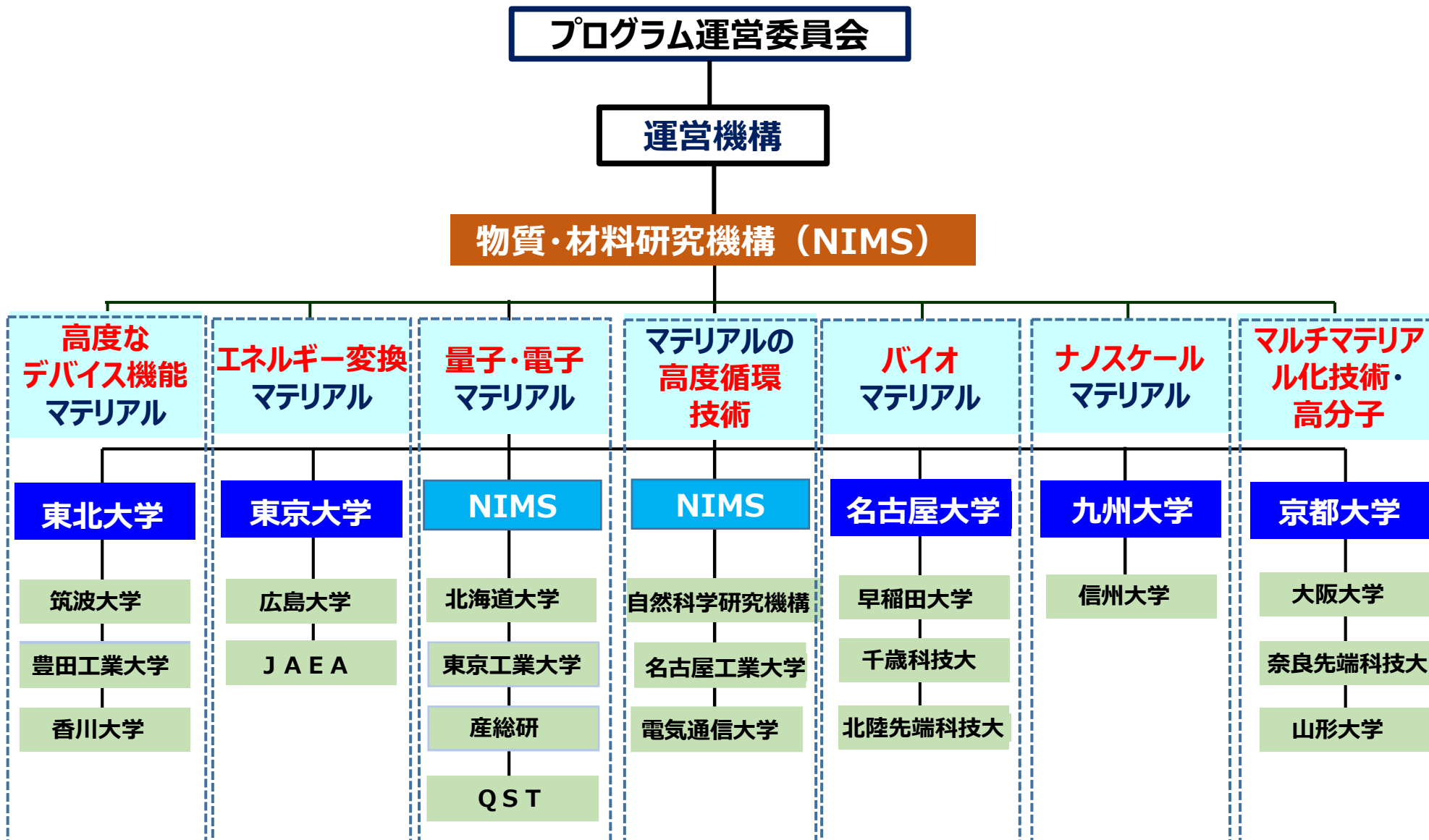
マテリアルDXプラットフォーム構想のイメージ



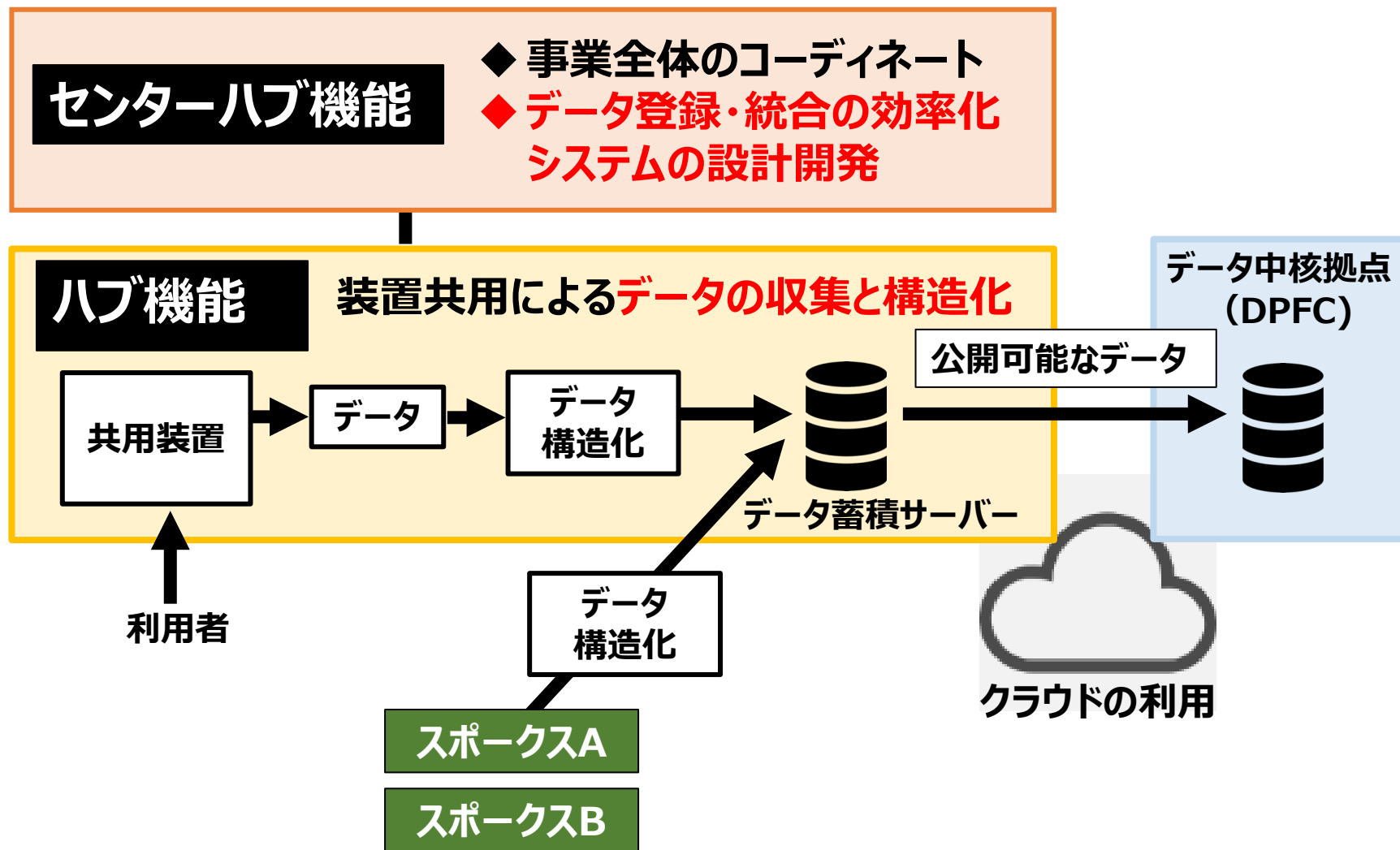
センターハブ機関

ハブ機関

スポーク機関

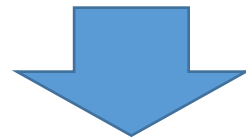


データの収集・構造化・共用のフロー



- データ構造化は、異なる材料研究分野の多様かつ多数の実験装置に対応する必要がある
- データ管理は、時間軸で変化する closed → shared → open の流通段階ごとに区別して実施

溜まったスペクトルや画像などの多次元の実験データを
第3者がデータ駆動型研究として利用するために



実験データの数理モデル化の目的と
X線光電子分光における事例紹介

数理モデル化：特徴量の抽出、スパースモデリング

手法の説明

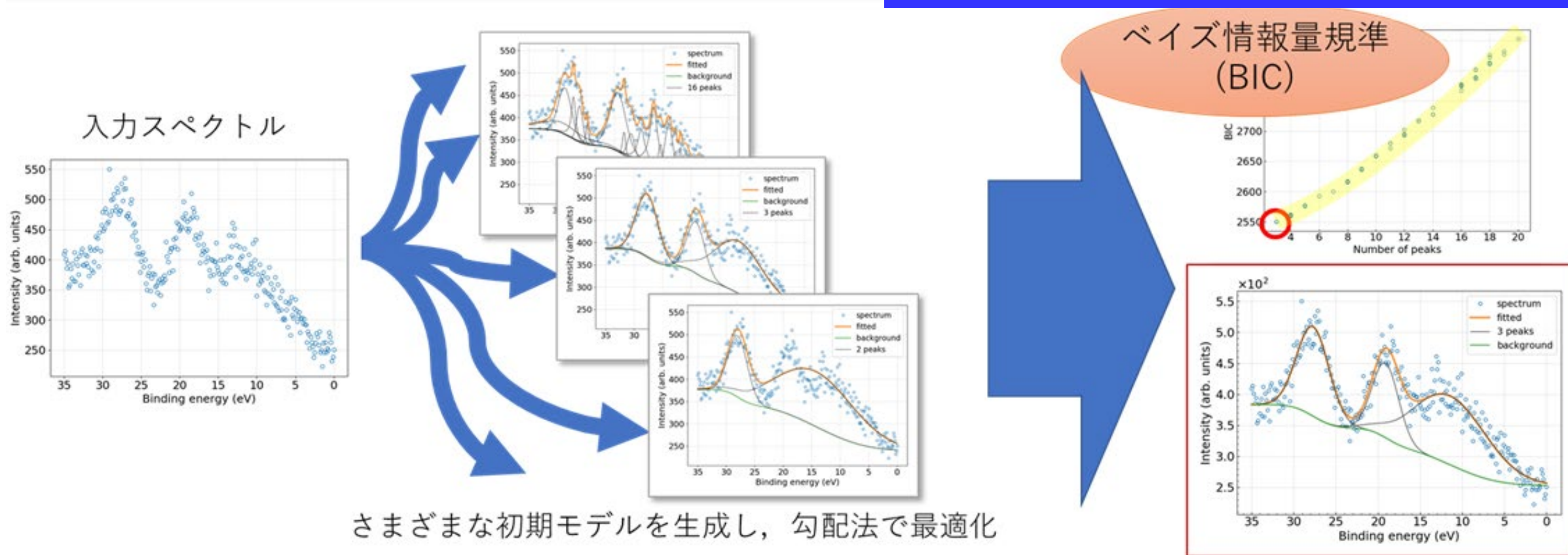
- 実験で得られた多次元の数値データ^(注)を出来るだけ少数個の数学パラメータで過不足なく表現する
- 数学パラメータの解が（局所解ではなく）大域解であること及び解の値の揺らぎを情報工学的に評価

（注）ここで言う材料研究における多次元の数値データとは、スペクトルデータや画像データなど及びそれらが複合化したハイパースペクトルデータを指す

目 的

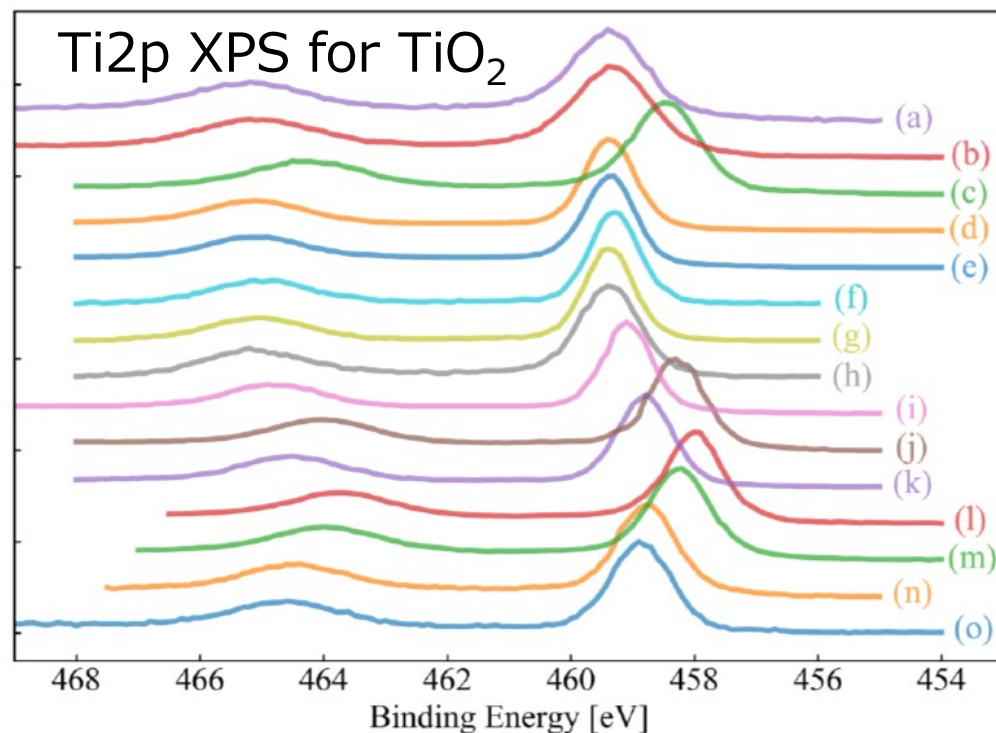
- その実験分野の専門外のデータ解析者に渡せるデータにする
 - ✓ 物理化学や材料学などの知見に基づいて実験データを数理モデル化する
 - ✓ 実験装置の機種や実験条件の違いが実験データに与える揺らぎも数理モデル化する
 - ✓ 数値データに付随する統計誤差も数理モデルに組み込む
- データの情報量（ファイルサイズ）の大幅な圧縮
- データのパラメトリック検索の高速化
- 複数のデータ間の類似度判定の高速化

モデル化：数学関数のパラメータで記述
ベイズ情報量規準を用いたモデルの自動探索



X線光電子分光(XPS)スペクトルを例として

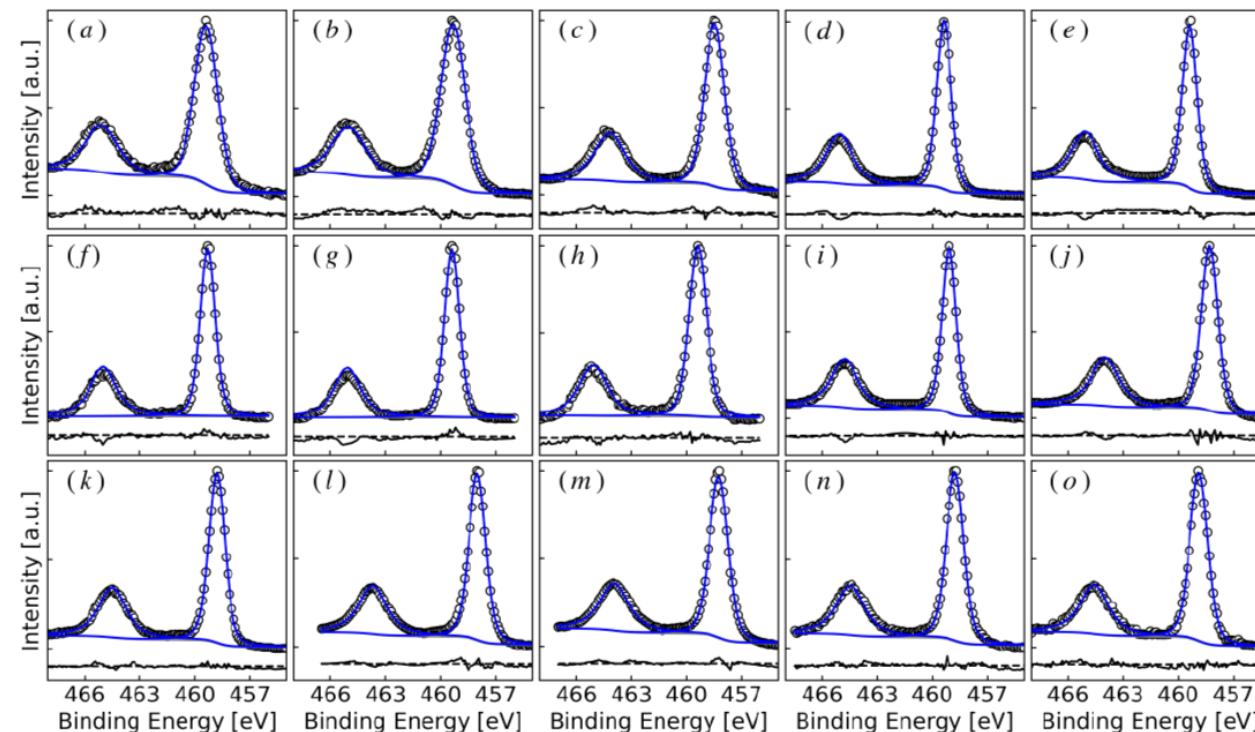
文献を含む他者のデータ群における実験装置由来のパラメータ（ピークの位置、幅、形状およびバックグラウンドの端点、形状など）の値のばらつきや統計誤差を考慮した自動数理モデル化のアルゴリズム開発



15種の文献にある TiO_2 のTi2pXPSスペクトルデータ

R.Murakami et al., STAM Methods 1, 182 (2021)

全データを最も良く説明するピーク本数 $K=2$ で統一



BICを使って判定をした自動スペクトル分解の結果
K値は最適ピーク本数を示す

Voigt関数で表現をした各ピークの位置と半値幅（Gauss成分）は、
実験装置に依存して変化することを数理モデルに取り込む

多機関で取得されたデータをデータ駆動型で扱う際に必要となる機能

データプラットフォーム構築にかかる主な発表文献

1. 出村 雅彦: [データ駆動材料研究の進展～NIMSにおける実践を例として](#), 情報の科学と技術, 71, 252 (2021).
2. 谷藤 幹子: [材料データプラットフォームシステムの設計と構築](#), 機能材料 40, 4 (2020).
3. X. Liu, L. Yang, Z. Hou, Bo Da, K. Nagata, H. Yoshikawa, S. Tanuma, Y. Sun, Z. Ding. [Machine learning approach for the prediction of electron inelastic mean free paths](#). Physical Review Materials. 5 [3] (2021).
4. H. Oka, A. Yoshizawa, H. Shindo, Y. Matsumoto, M. Ishii : [Machine extraction of polymer data from tables using XML versions of scientific articles](#), Sci. Technol. Adv. Mater. Methods 1, 12 (2021).
5. T. Takemura, M. Ishii, M. Tanifuji : [PoLyInfo RDF: A Semantically Reinforced Polymer Database for Materials Informatics](#). CEUR Workshop Proceedings. 2019, 69-72.
6. L. Foppiano, L. Romary, M. Ishii, M. Tanifuji : [Automatic Identification and Normalisation of Physical Measurements in Scientific Literature](#). DocEng '19 Proc. ACM Symp. on Document Eng 2019.
7. S. Dieb, K. Amano, K. Tanabe, D. Sato, M. Ishii, M. Tanifuji: [Creating Research Topic Map for NIMS SAMURAI Database Using Natural Language Processing Approach](#): Sci. Technol. Adv. Mater. Methods 1, 2 (2021).
8. 谷藤 幹子, 吉川 英樹: [材料データプラットフォームシステムDICEにおける研究データフローの構築—実践と課題](#). デジタルプラクティス. (2021) 57.
9. 松波 成行, 松田 朝彦, 知京 豊裕, 原田 善之, 吉川 英樹. [IoTデータ収集システムのデータアーキテクチャ](#). デジタルプラクティス. (2021) 80.
10. M. Suzuki, H. Nagao, Y. Harada, H. Shinotsuka, K. Watanabe, A. Sasaki, A. Matsuda, K. Kimoto, H. Yoshikawa: [Raw-to-Repository Characterization Data Conversion for Repeatable, Replicable, and Reproducible Measurements](#): J. Vac. Sci. Technol. A38 [2] (2020) 023204.
11. 石井 真史, 上杉 文彦, 小澤 哲也. [計測インフォマティクスとデータベースの統合による客観・高速結晶構造解析](#). 日本結晶学会誌. 62 [1] (2020) 35-42 10.5940/jcrsj.62.35.
12. M. Tanifuji, A. Matsuda, H. Yoshikawa: [Materials Data Platform - a FAIR System for Data-Driven Materials Science](#). Proc. 2019 8th Int. Congress on Advanced Applied Informatics (IIAI-AAI). 2019