

# マテリアルズ・インフォマティクスのための データ駆動型リポジトリ構築に向けて

国立研究開発法人 物質・材料研究機構  
石井 真史

革新的材料開発力強化（M<sup>3</sup>）プログラムを構成するMaterial Research Bank  
材料R&Dデータの入り口（収集）から出口（サービス）まで  
データ収集 → 処理 → 解析 → 蓄積 → 管理 → 公開

→ 駆動 → 協調的利用 → 競合的活用





つくる

## データ創出

- 実験
- 文献マイニング
- シミュレーション
- ラボノート

つかう

## データ活用

- DPFで
- インテグレーション
  - インフォマティクス
  - 研究者による分析

みる・引用

## データ公開

- ビュー
- パブリッシュ
- 他機関・分野リポジトリとの相互連携

ユーザー

## データの来歴・リソース・品質保証・利用者管理

ためる

## データ蓄積・保持

運用者

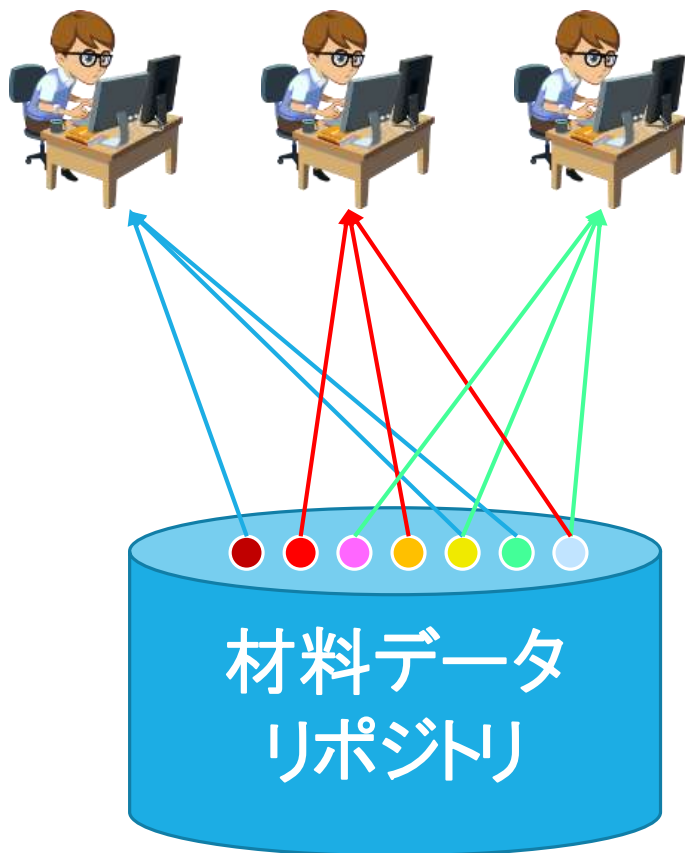
- 利活用データとして
  - MatNavi(機構内外の利用)
  - 計測データ
  - 計算データ
  - 標準データ
  - MIで生み出したデータ
- 機関アセットデータとして
  - 論文・補足・エビデンスデータ
  - 報告書データ
  - 受託事業で得られたデータ
  - 検証データ
  - 事故調査記録データ
  - 人プロフィールデータ



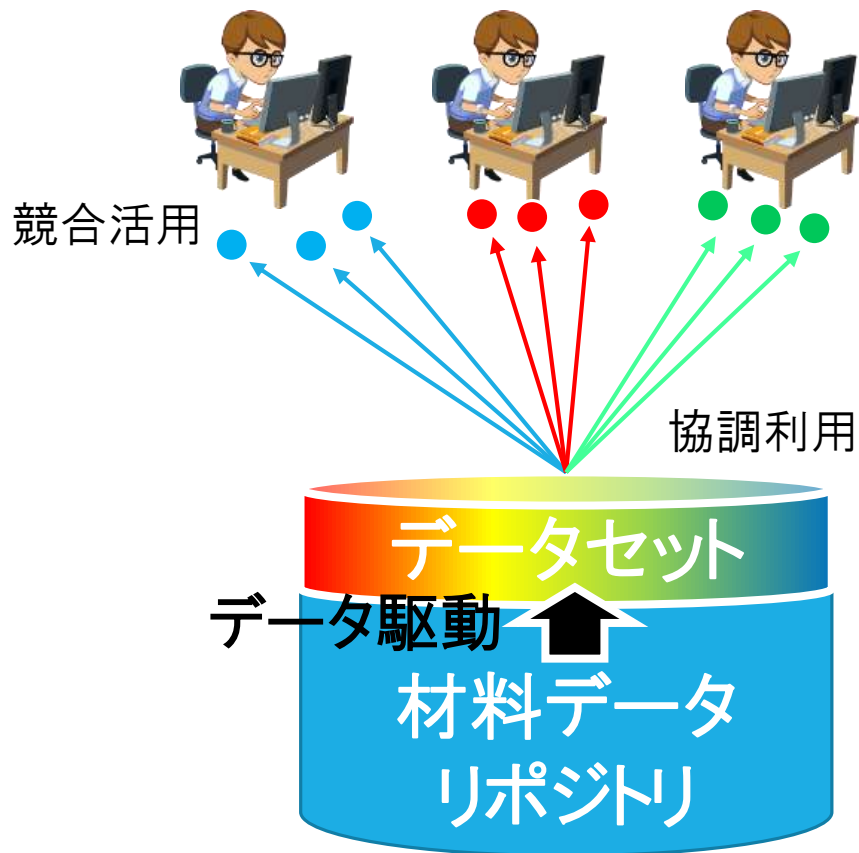
# MaDIS データ公開の先にあるもの： 次世代リポジトリとしての「データ駆動」の使い方



## 公開データの利用



## 駆動されたデータの協調利用 →競合活用



## 先進データベース

- 文献からのデータマイニング
- 機械学習手法を取り入れた先進的データベース

## データ収集と高付加価値化

- 実験装置、計測機器、IoT
- 日本の大型研究施設から
- メタ情報とトレーサビリティを担保
- 信頼性のあるデータ、標準データ



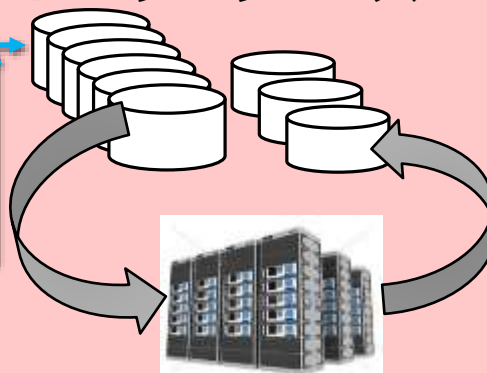
NanotechJapan  
Nanotechnology Platform

## SIP-MI統合システムによる利用技術の活用

## データ駆動型 データパブリッシュ

## 材料情報統合 データプラットフォーム

### 物質・材料 データベース



### 解析サーバ

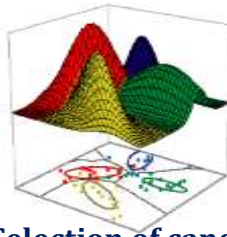
ツールや  
アプリケーション

連携・共同研究

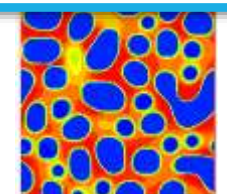
## 協調利用→ 競争活用

Visualization of materials

## 公開+利用



Selection of candidate materials



Physical model of micro-structure



民間利用

学術利用



## 各種機能性

高運航効率・低騒音



CO<sub>2</sub> ゼロ排出



超軽量



極限環境素材

超スピード開発  
経験も勘も不要



自動で  
未知情報の宝庫へ



マテリアルズ・  
インフォマティクス

1																	18																		
1	H																	2	He																
3	Li	4	Be													7	N	8	O	9	F	10	Ne												
11	Na	12	Mg									13	Al	14	Si	15	P	16	S	17	Cl	18	Ar												
19	K	20	Ca	21	Sc	22	Ti	23	V	24	Cr	25	Mn	26	Fe	27	Co	28	Ni	29	Cu	30	Zn	31	Ga	32	Ge	33	As	34	Se	35	Br	36	Kr
37	Rb	38	Sr	39	Y	40	Zr	41	Nb	42	Mo	43	Tc	44	Ru	45	Rh	46	Pd	47	Ag	48	Cd	49	In	50	Sn	51	Sb	52	Te	53	I	54	Xe
55	Ba	56	La	57	Ce	58	Pr	59	Nd	60	Pm	61	Sm	62	Eu	63	Gd	64	Tb	65	Dy	66	Ho	67	Er	68	Tm	69	Yb	70	Lu				
87	Fr	88	Ra	89	Ac	90	Th	91	Pa	92	U	93	Np	94	Pu	95	Am	96	Cm	97	Bk	98	Cf	99	Es	100	Fm	101	Md	102	No	103	Lr		



従来



ひたすら頑張る





## データセット

## 米の収穫量 平均気温 降水量 日照時間

	A	B	C	D	E	F
1	平成13年度水稲単収					
2	単収(g/m <sup>2</sup> ) 年平均気温(°C) 降水量(mm) 日照時間(時)					
3	北海道	526	8.2	1130	1805	
4	青森	575	9.7	1360	1695	
5	岩手	528	9.8	1265	1815	
6	宮城	536	11.9	1205	1843	
7	秋田	574	11.1	1745	1642	
8	山形	601	11.2	1126	1667	
9	福島	577	10.7	1252	1802	
10	茨城	510	13.2	1308	1820	
11	栃木	528	13	1282	1842	
12	群馬	476	13.9	1130	1980	
13	埼玉	486	14.2	1168	1989	
14	千葉	504	15	1058	1911	
15	東京	400	15.6	1405	1811	
16	神奈川	471	15.2	1089	1880	
17	新潟	557	13.2	1778	1697	
18	富山	545	13.5	2296	1608	
19	石川	518	14.1	2060	1608	
20	福井	522	14.1	2368	1609	
21	山梨	548	13.9	1055	2075	
22	長野	520	11.5	938	1952	
23	岐阜	495	15.1	1094	2067	
24	静岡	528	16.3	2217	2073	
25	愛知	510	15.1	1094	2015	
26	三重	496	15.1	1055	1986	
27	滋賀	517	14.1	1054	1825	
28	京都	516	15.2	1081	1708	
29	大阪	486	16.2	1218	1944	
30	兵庫	514	15.6	1216	1918	
31	奈良	506	14.4	1055	1949	
32	和歌山	488	16.1	1262	2021	
33	鳥取	540	14.5	1050	1695	
34	島根	517				
35	岡山	522				
36	広島	523				
37	山口	521				
38	徳島	471				
39	香川	502				
40	愛媛	508				
41	高知	459				
42	福岡	515				
43	佐賀	544				
44	長崎	485				
45	熊本	522				
46	大分	512				
47	宮崎	487				
48	鹿児島	486				
49	沖縄	285				



	A	B	C	D	E	F
1	平成13年度水稲単収					
2	単収(g/m <sup>2</sup> ) 年平均気温(°C) 降水量(mm) 日照時間(時)					
3	北海道	526	8.2	1130	1805	
4	青森	575	9.7	1360	1695	
5	岩手	528	9.8	1265	1815	
6	宮城	536	11.9	1205	1843	
7	秋田	574	11.1	1745	1642	
8	山形	601	11.2	1126	1667	

## 都道府県



説明変数(年平均気温、降水量、日照時間)で収穫量を予測

$$\text{水稲の収穫量} = 714 - 17 \times \text{年平均気温} + 0.01 \times \text{降水量} + 0.02 \times \text{日照時間}$$



例えるならばお米作りでのデータ活用

## マテリアルズ・インフォマティクス

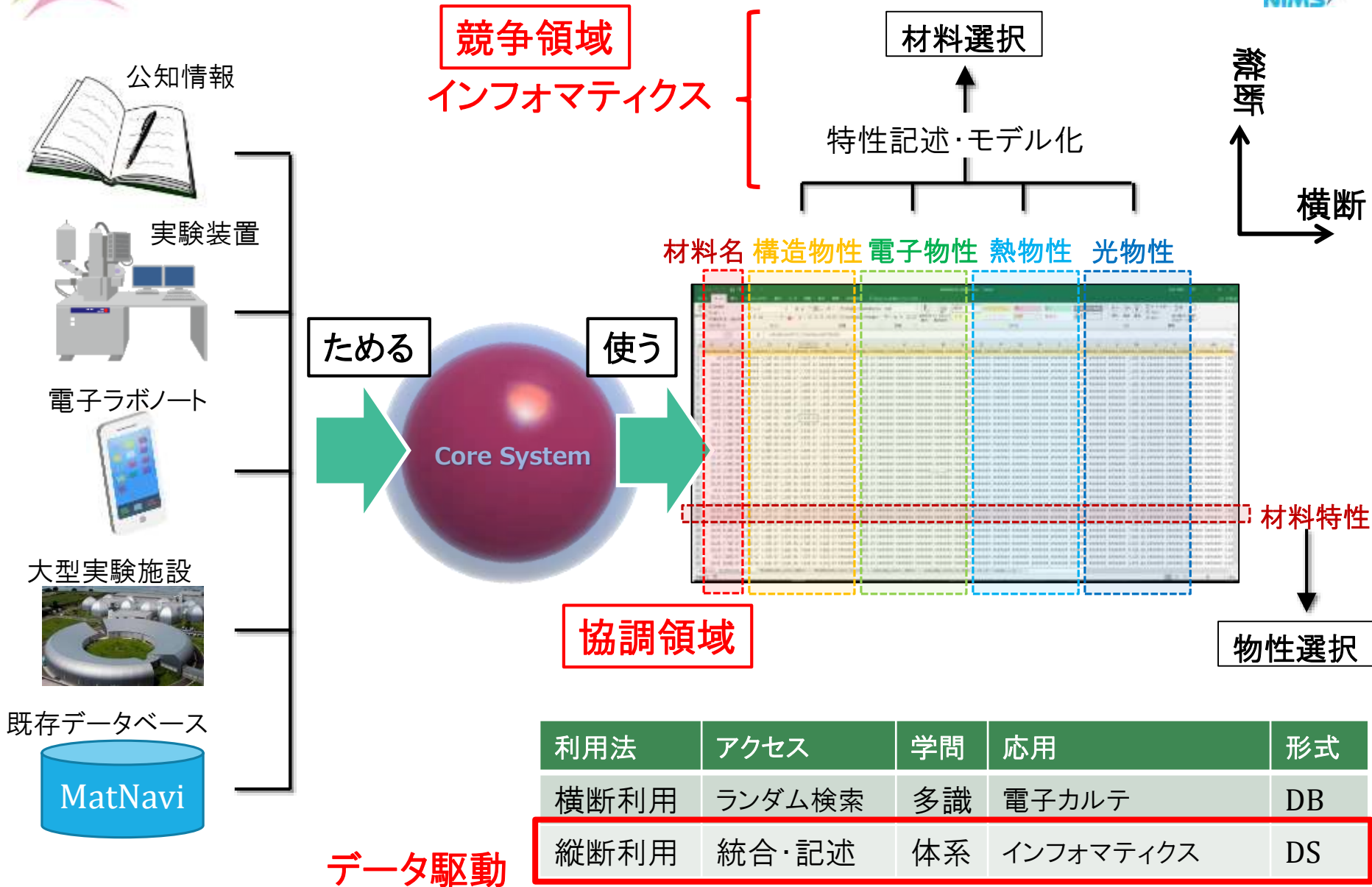
- ・都道府県→材料名
- ・収穫量→機能性(例:磁石の強さ)
- ・説明変数→物性(結晶構造、原子の大きさ、電子の数・・・)



得られた式から新材料を予測

参考:島根大学ホームページ

[http://www.ipc.shimane-u.ac.jp/food/kobayasi/multipleregression\\_excel.htm](http://www.ipc.shimane-u.ac.jp/food/kobayasi/multipleregression_excel.htm)



## NIMS材料データプラットフォームセンター

マテリアルズ・インフォマティクスを背景に  
「データ公開」は、「データ駆動→協調利用→競争利用」に変化しつつある

利用者の観点では、  
「次世代リポジトリ＝データ駆動型リポジトリ」  
(NGRの必要要件はとにかく)