第9回FCワークショップ
MIX platform説明会

NIMS 袖山慶太郎, 冨中悟史

2024年10月2日 13:00-16:30
東京国際交流館 平成プラザ ３F メディアホール（オンライン配信および講習会場）

Contents

前半：講演
1. MI研究の動向紹介（MIでできること、今後の動向）
2. MIX platform紹介
    2.1. 開発背景と現状、今後の利用について
    2.2. 電気化学測定結果の解析（LSV, IV, ORR）
    2.3. 簡単なMI解析
    2.4. 自動自律実験による実験データ蓄積
    2.5. 自然言語処理による論文データ蓄積
3. データの連携、検索

後半：実習(オンサイトのみ)
① MIX Platformへの接続方法、使い方の説明
② 演習課題の説明

MI解析の目的

データ科学、計算科学、理論及び実験を融合させ、物質・材料研究開発のあり方を革新し、**研究開発のスピードを大幅に加速させる**ことを目指す。

材料研究
（マテリアルリサーチ）

＋

データ駆動科学、
機械学習

・超伝導材料（無機固体、有機固体）
・磁性材料（無機固体）
・高分子ポリマー材料（有機液体、固体）
・触媒材料(固体+気体、液体)
・リチウムイオン電池(無機固体+有機液体)
など、材料研究の適用先は多岐にわたる

→

・新規材料開発の高速化
・新規メカニズムの解明

MI解析に必要な知識・ノウハウ

- 材料研究を加速するためにMI解析を利用する。
- **データのため方＋解析の仕方　両方の理解必要**

自分のデータだけ使う

自分のグループのデータだけ使う

外部データと合わせて使う

- 語彙辞書の作成
- 計測ルールの制定

どんなアウトプットが欲しいか

どんなインプットがあるか

どんなMI解析手法が使えるか

- 要因解析
- 条件最適化
- 分類
- 新規材料提案

MI研究事例紹介

1．メタノール酸化触媒　活性の高い触媒組成を高速探索
　初期17点　ベイズ最適化

2．Pt-HEA(111)モデル触媒の合成条件検討
　活性の高い触媒組成と合成温度を高速探索
　初期31点　ベイズ最適化

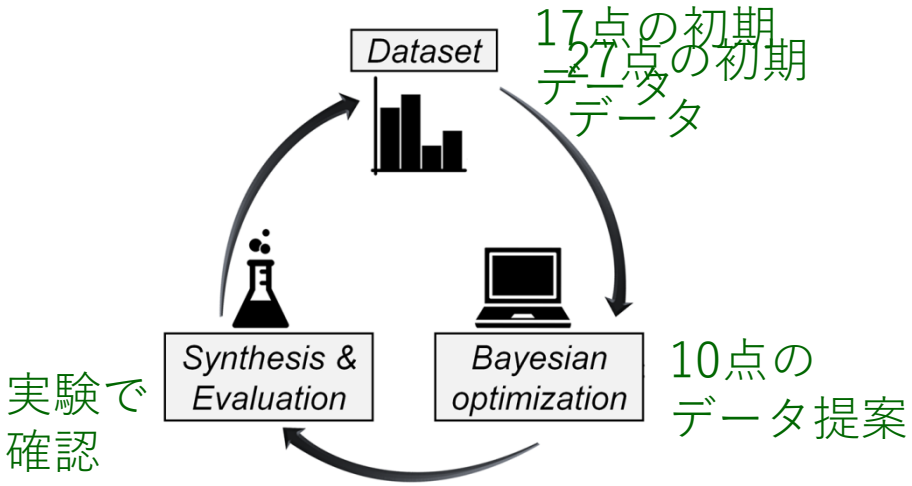
3．MI手法開発 分子式の方法のみから物性予測　SMILES-X

最近のMI研究の動向（成功事例と課題１）

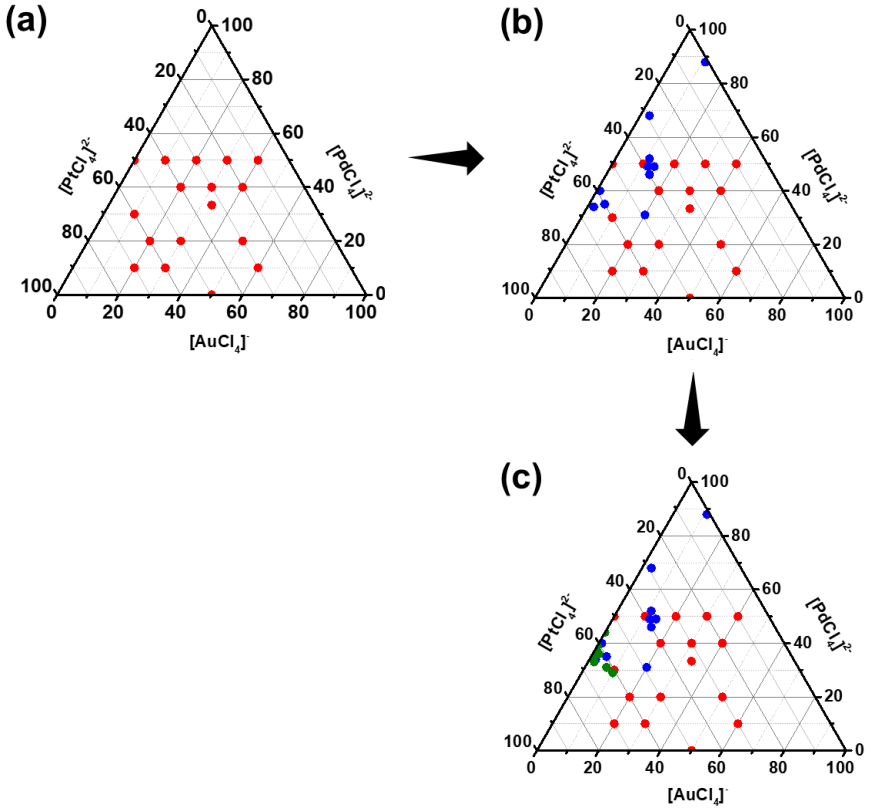
Output: 活性の高い触媒の組成を早く見つけたい
Input: 初期17点
MI手法: ベイズ最適化

メタノール酸化触媒の設計:
触媒活性の高いPtPdAu合金の組成探索

Pt: Pd: Au = x: y: z



アクティブラーニングにより、
実験と予測を繰り返す。

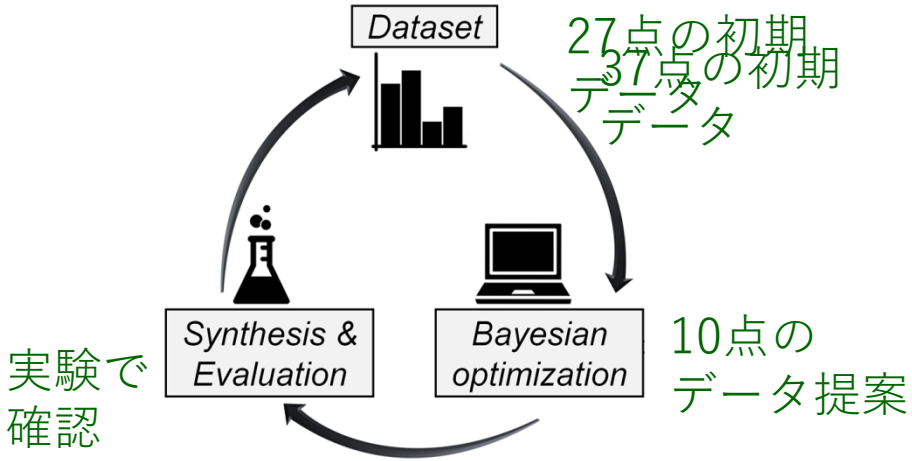


J. Mater. Chem. A **2020**, *8*, 13532–13540

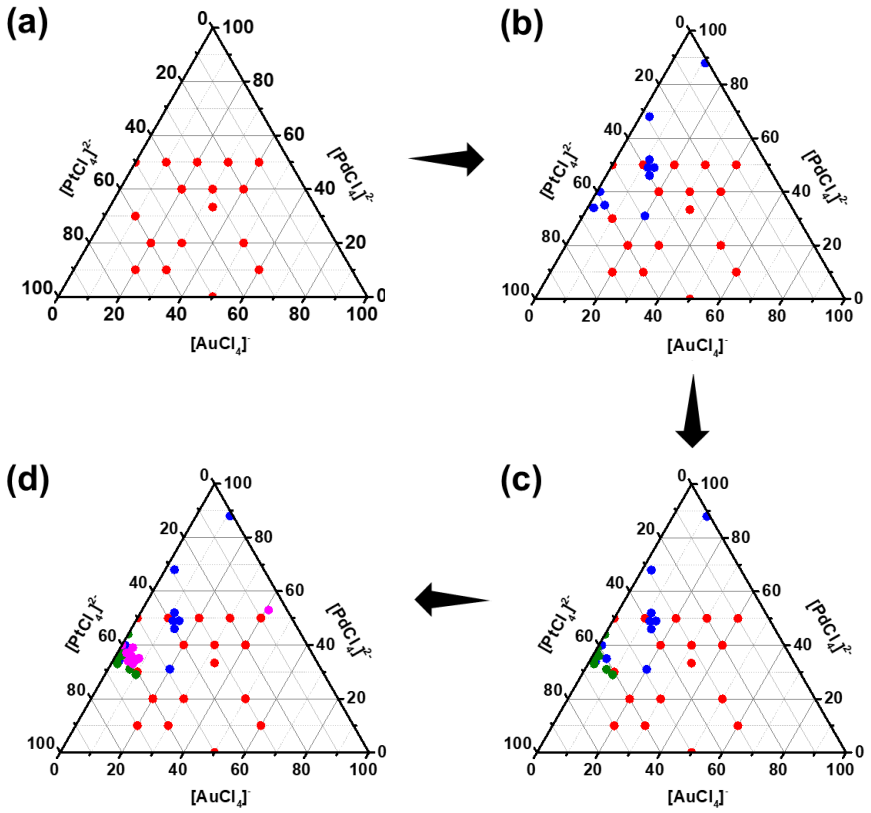
最近のMI研究の動向（成功事例と課題１）

メタノール酸化触媒の設計:
触媒活性の高いPtPdAu合金の組成探索

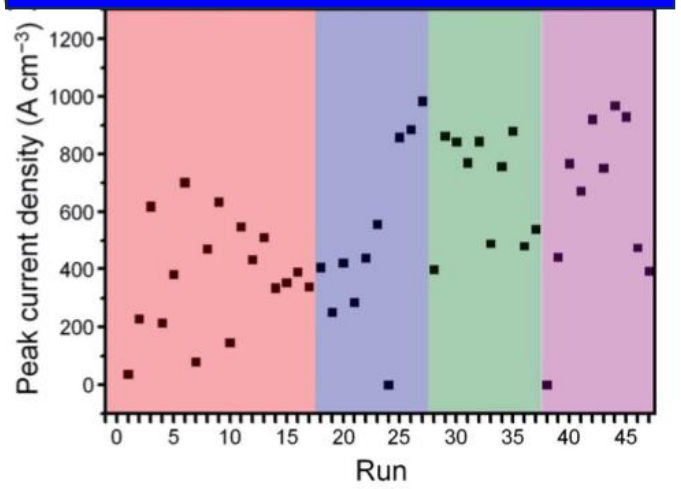
Pt: Pd: Au = x: y: z



アクティブラーニングにより、
実験と予測を繰り返す。



Output: 活性の高い触媒の組成を早く見つけたい
Input: 初期17点
MI手法: ベイズ最適化



・実験者にとって予想外の組成が高速に見つかった。

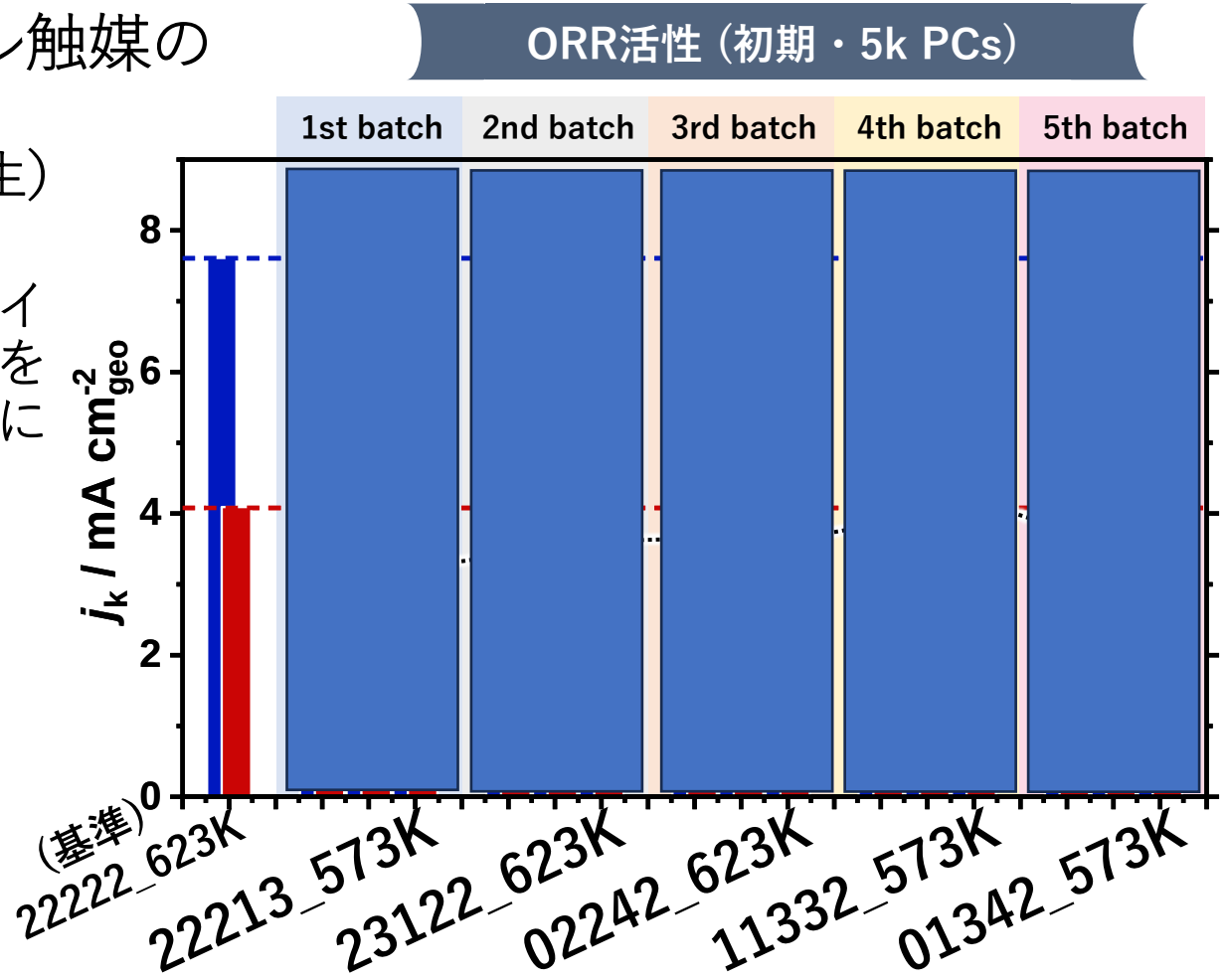
Pt: Pd: Au=59:40:1
の組成で従来より40%性能向上（予想外の組成が得られた）

J. Mater. Chem. A **2020**, *8*, 13532–13540

最近のMI研究の動向（成功事例と課題２）

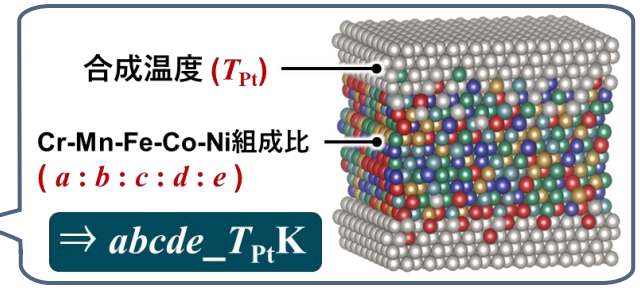
Pt-HEA(111)モデル触媒の合成条件検討
（東北大学 和田山先生）

ハイエントロピーアロイ触媒の組成と実験条件をアクティブラーニングにより最適化。



Output: 高性能新規触媒組成と合成温度を提案したい
Input: 初期
MI手法: ベイズ最適化

・触媒組成情報と合成温度のみから物性予測に成功



✓ 1st ~ 4th batchまでは、batch内で最高の特性を示した試料の”5kPCs後活性”がbatch毎に向上 ⇒ 4th batchで Pt / 等組成Cr-Mn-Fe-Co-Ni と同等のORR特性

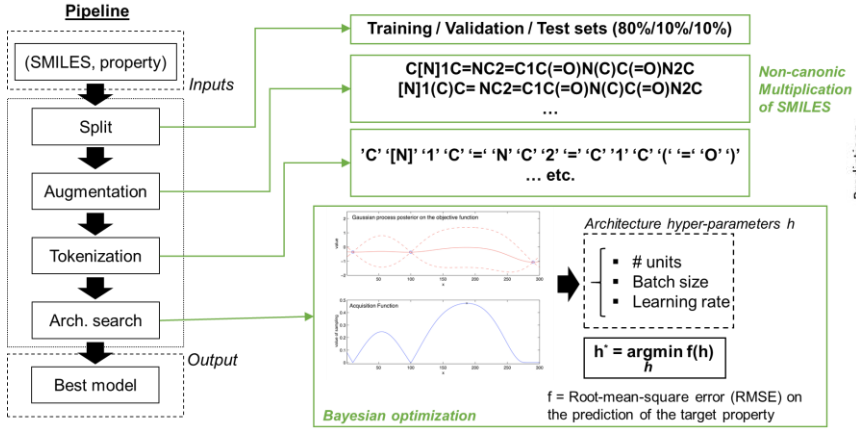
✓ 非等組成比5元系, 4元系の合金化元素でも高ORR特性を示す合成条件を発見

最近のMI研究の動向（成功事例と課題3）

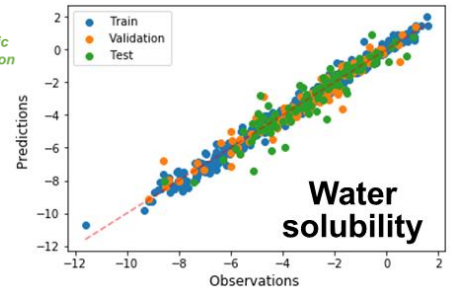
分子式の情報のみから物性予測・新分子生成（言語

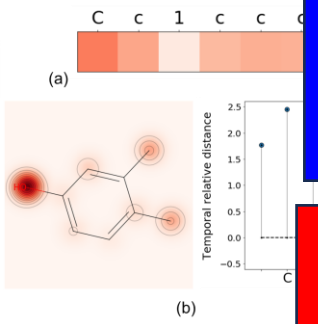
Output: 分子の物性を予測したい。高性能新規分子を提案したい
Input: 初期1783点,106点
MI手法: SMILES-X

・分子式の情報のみから物性予測に成功した。

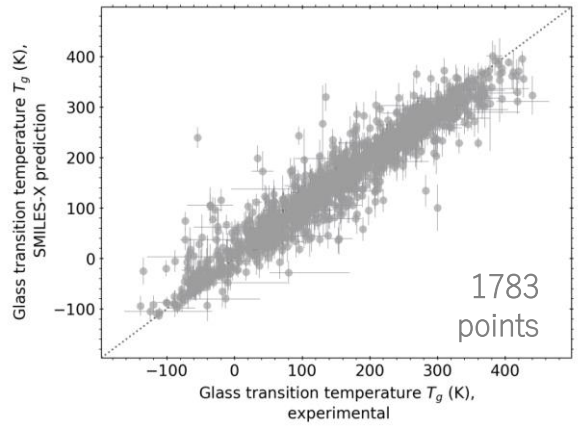


Input: SMILES strings (only)
Output: Physical properties prediction

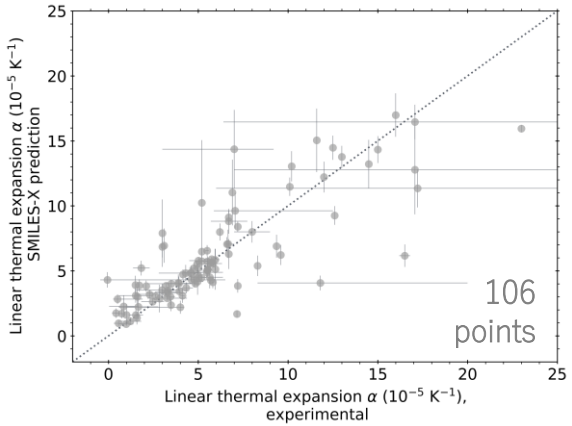




Water solubility prediction
Interpretation is also possible.
(G. Lambard et al., Mach. Learn.: Sci. Technol., 1(2), 025004, (2020) .)



grass transition temperature (Tg)



Linear thermal expansion (β)

SMILES-X: using chemical structure information, not chemical fingerprint

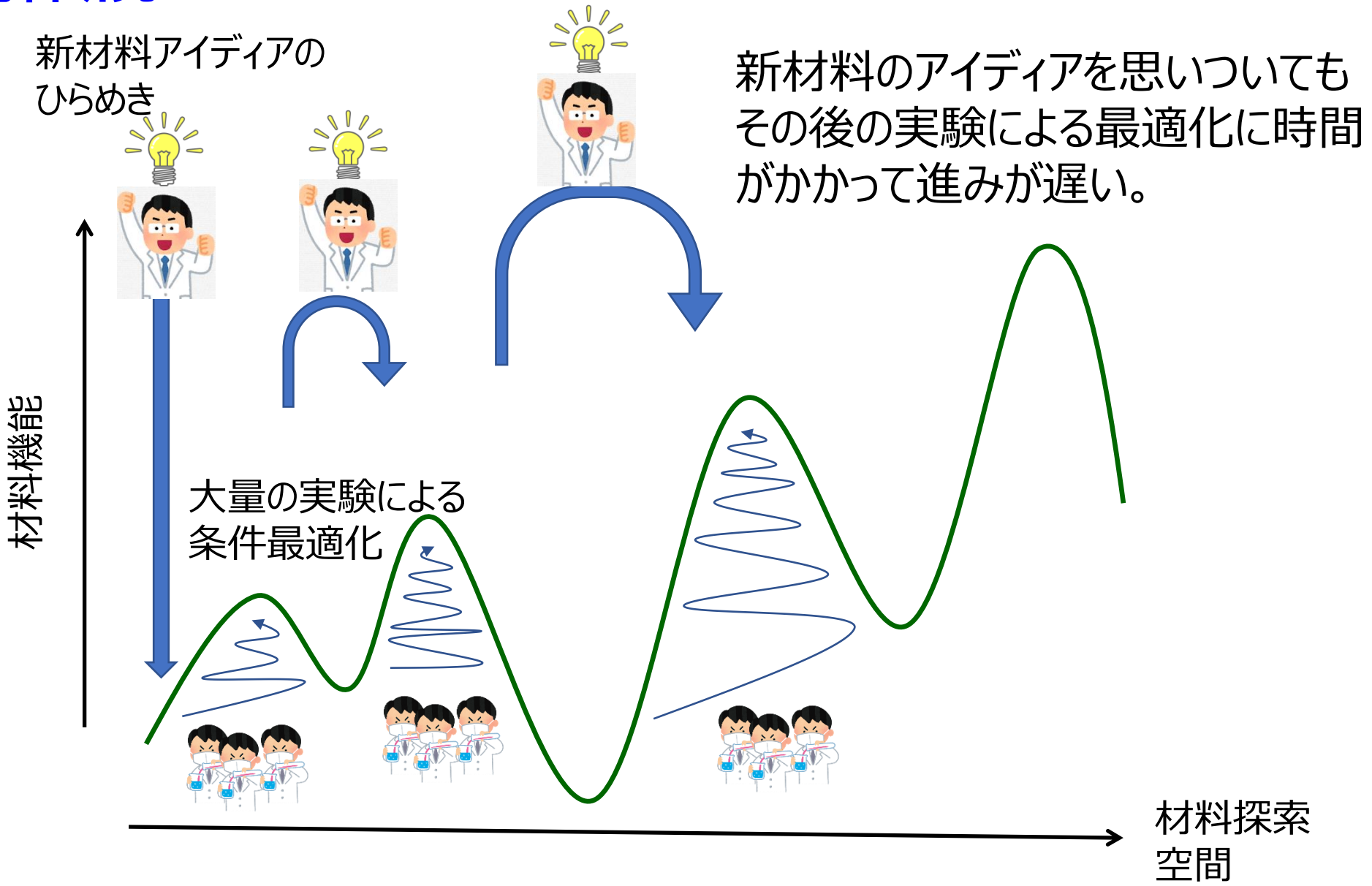
Property	Model	R^2-score	RMSE
α (10^{-5} K^{-1})	SMILES-X	0.68 ± 0.06	2.44 ± 0.17
	Random forest	0.65 ± 0.02	2.54 ± 0.07
T_g (K)	SMILES-X	0.92 ± 0.00	33.21 ± 0.68
	Random forest	0.84 ± 0.00	45.70 ± 0.62

Only monomer information, β value is predicted.

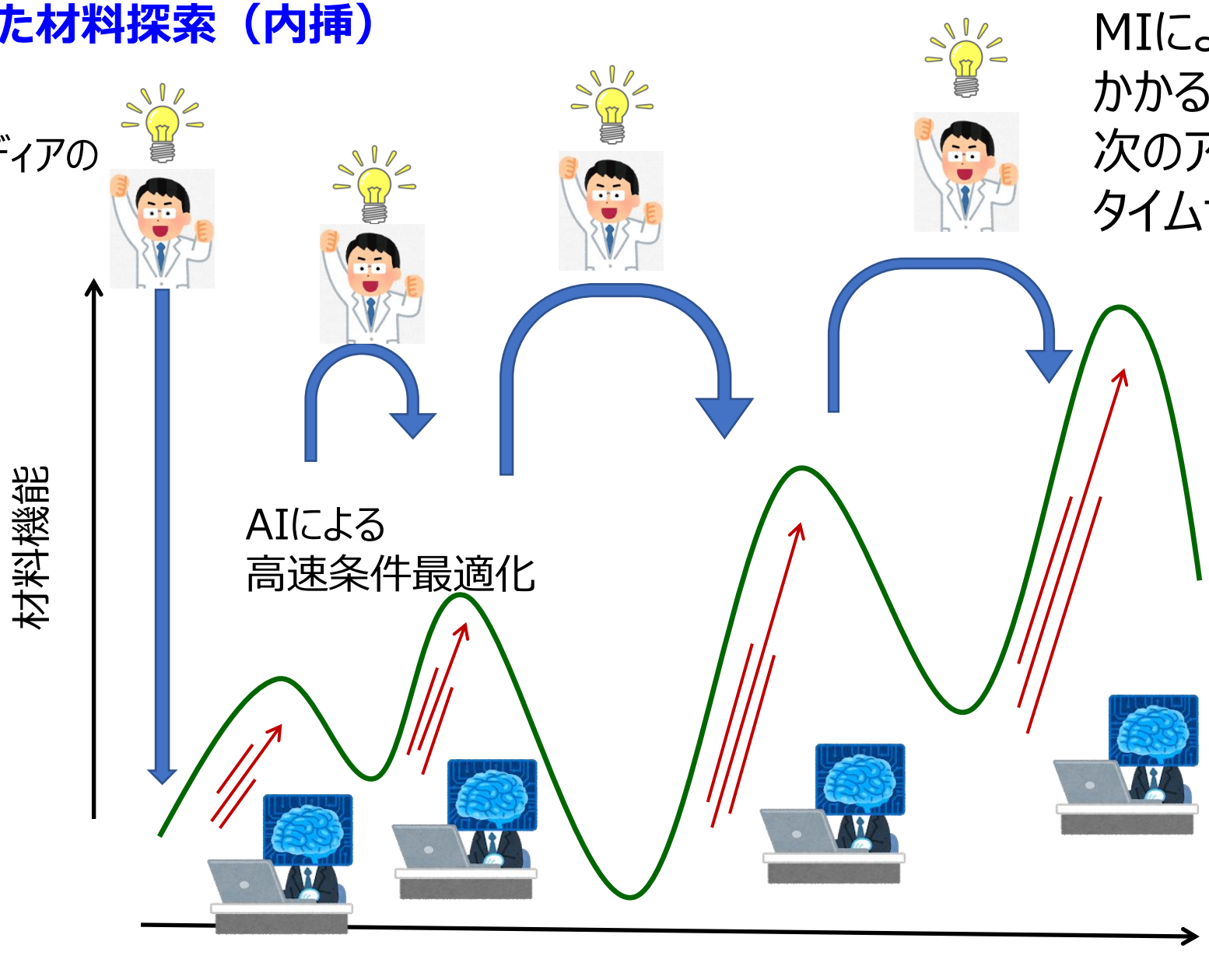
E. Gracheva, G. Lambard, S. Samitsu, K. Sodeyama, A. Nakata, *STAM-methods, 1:1*, 213 (2021).

最近のMI研究の動向（今後の方針）

一般的な材料研究



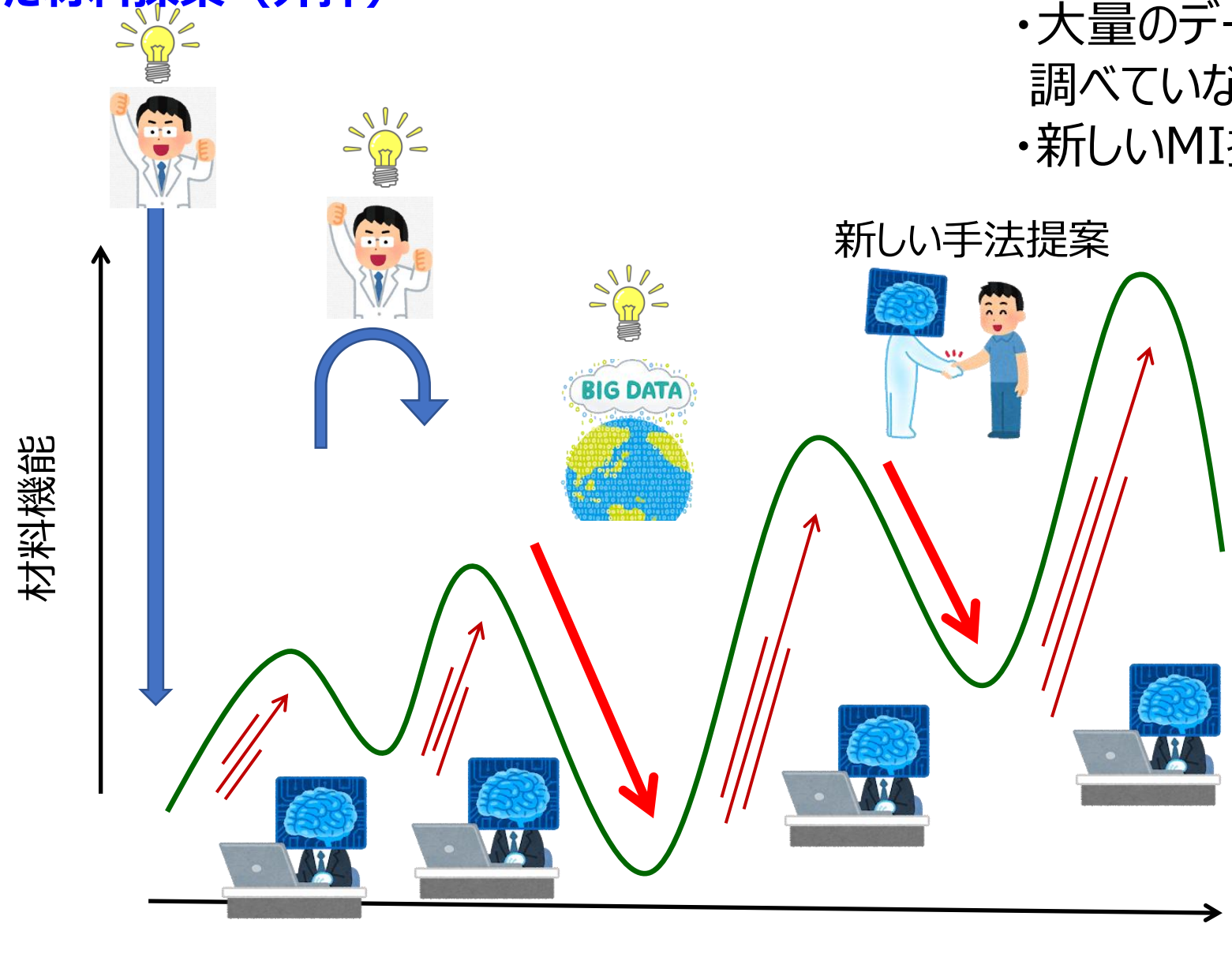
データを使った材料探索（内挿）

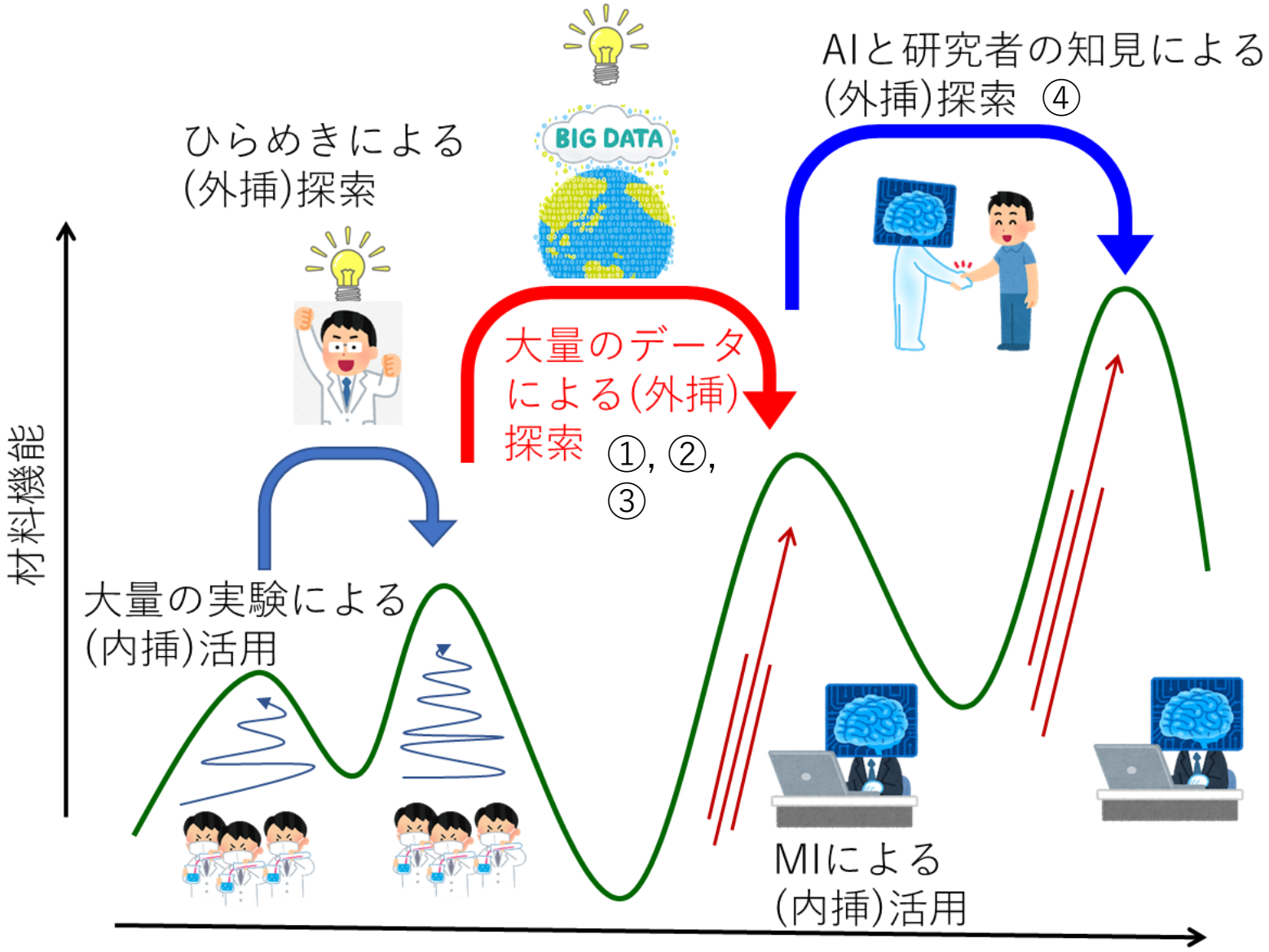


MIにより、実験による最適化にかかる時間を短縮する。
次のアイディアを実現するまでのタイムサイクルを短縮。

データを使った材料探索（外挿）

・大量のデータ蓄積により、これまで調べていない材料系を提案
・新しいMI探索手法を提案





①PEFCデータベース
②自律自動実験データ
③自然言語処理による論文データ
④PEFCに関するMI解析ノウハウと解析手法開発

⇒新規材料の設計指針提案

最近のMI研究の動向（データのため方）

- 材料研究を加速するためにMI解析を利用する。
- **データのため方＋解析の仕方　両方の理解必要**

自分のデータだけ使う

自分のグループのデータだけ使う

外部データと合わせて使う

項目名(語彙)や計測の仕方は仲間内だけで統一すればよかった。

しかし、それだと**データ数が少ない。**

データ数が多いほど、予測精度はあがり、新しい材料を見つけられる可能性高まる。

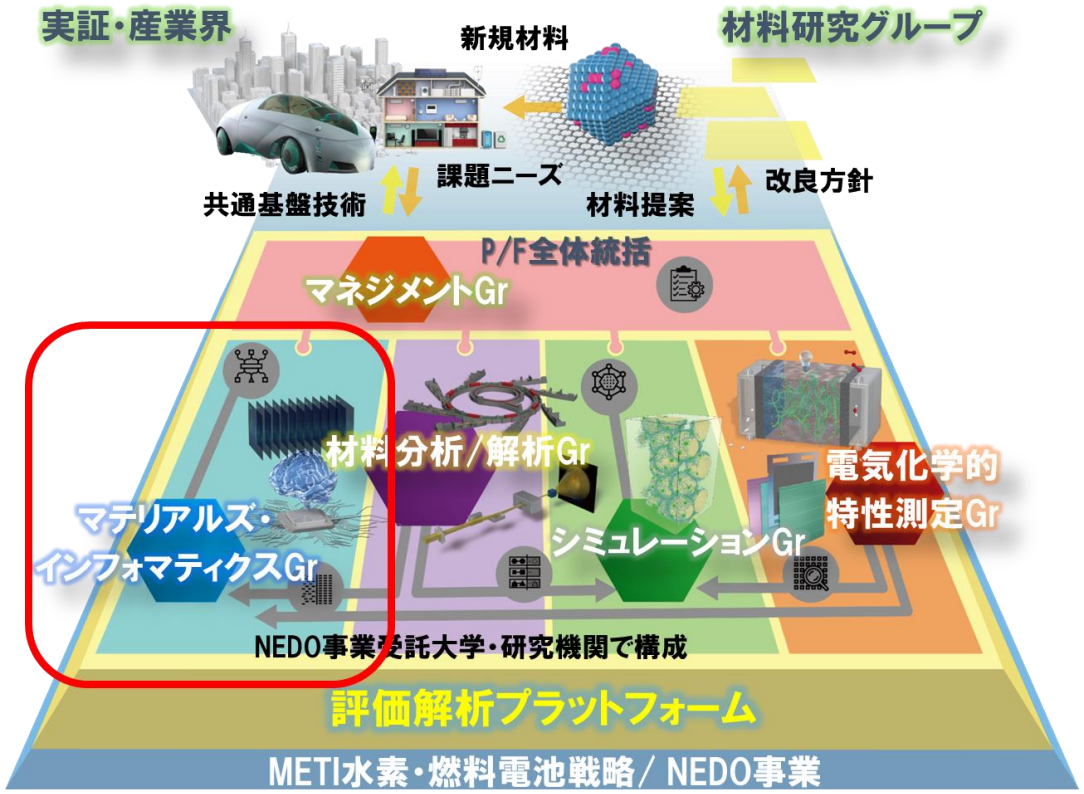
そこで、

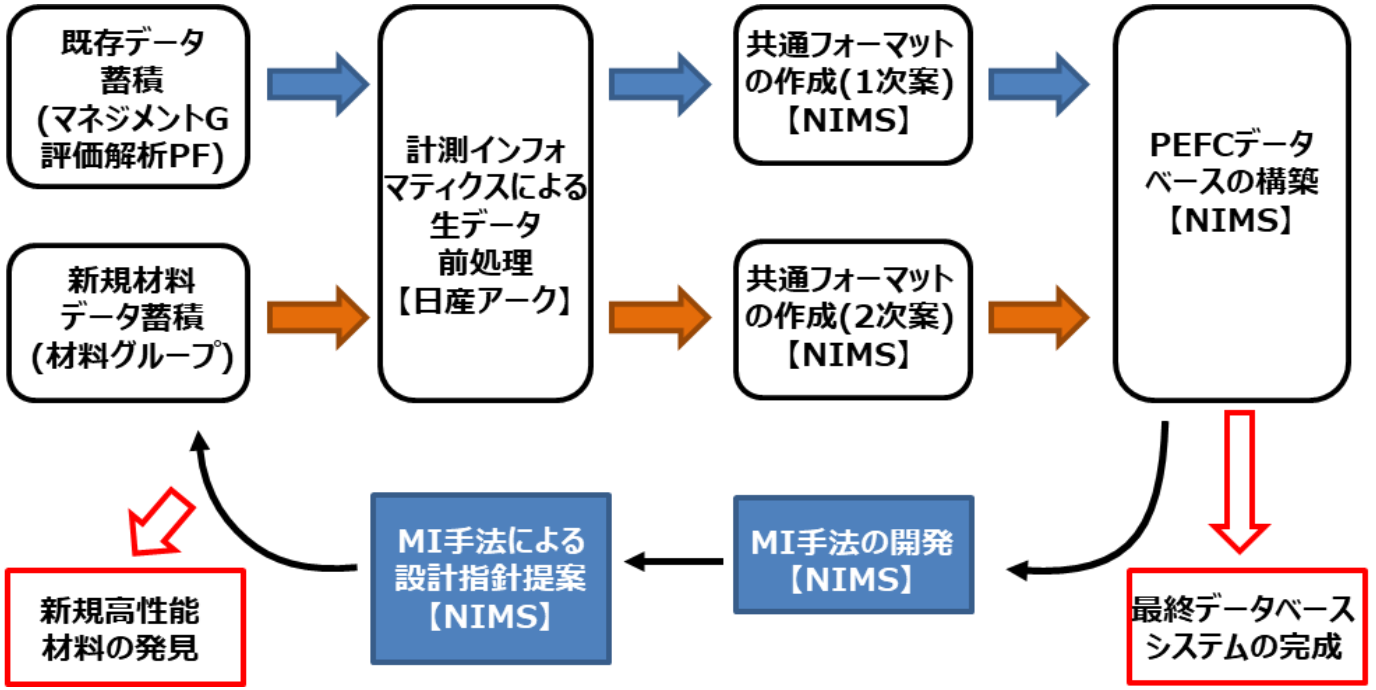
・語彙のリスト化（語彙の類義語を辞書化）

・計測ルールの共通化（理想的な計測範囲を設定）

が、分野ごとに行われている。

MIX Platformについて（背景）





・NEDO燃料電池事業にてデータを蓄積するために開発
・ウェブアプリに機能を統合。自動的にDBにアクセス。
・データ蓄積時に、簡単なMI解析や実験解析を行うことが可能

MIX Platformについて（全体像）

MIX Platform: 初のPEFC専用データ解析・蓄積プラットフォーム

1．PEFC関連データを蓄積するだけでなくデータ解析が可能

- **電気化学測定データ**
- **スペクトルデータ**
- **画像データ**

本日は、現在開発が一番進んでいる2項目について説明します。
（今後アップデートしていきます。）

2．蓄積したデータのMI解析機能

3．PEFC関連の生データだけでなく解析後のデータを、実験条件などのメタデータ付きで蓄積が可能

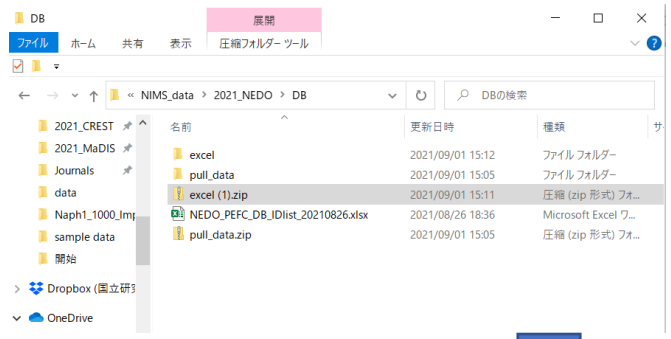
MIX Platformで出来ること

ウェブブラウザ（Chrome推奨）でURLにアクセス

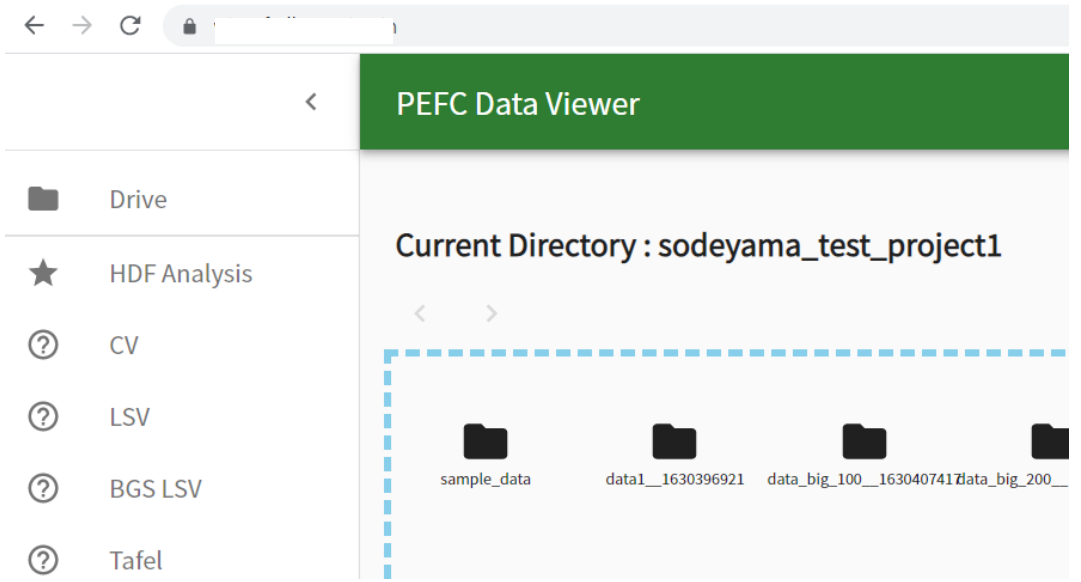
→データをアップロードする画面が表示
　アップロード後、以下の機能が利用可能。

1．生データの読込と実験データ解析
2．統一的な形でデータを保管
**　　→PEFCデータベースへ**
3．簡単なMI解析を手元で実施

エクスプローラー



ウェブアプリに
ドラッグ＆ドロップ



MIX Platformで出来ること（データの読込と保存）

現在

ファイル形式
A

ファイル形式
C

ファイル形式
B

ファイル形式
D

変換
DB格納

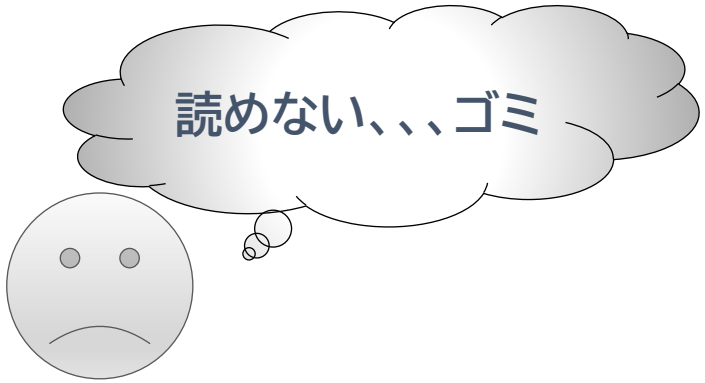
HDF※形式

提供者のメリット①:
生データを自動読み込み
実験データ解析可能

※ 生データと実験条件
（メタデータ）を合わせて保存

提供者のメリット③:
簡単なMI解析ができる

20年後



提供者のメリット②:
データを死蔵させない

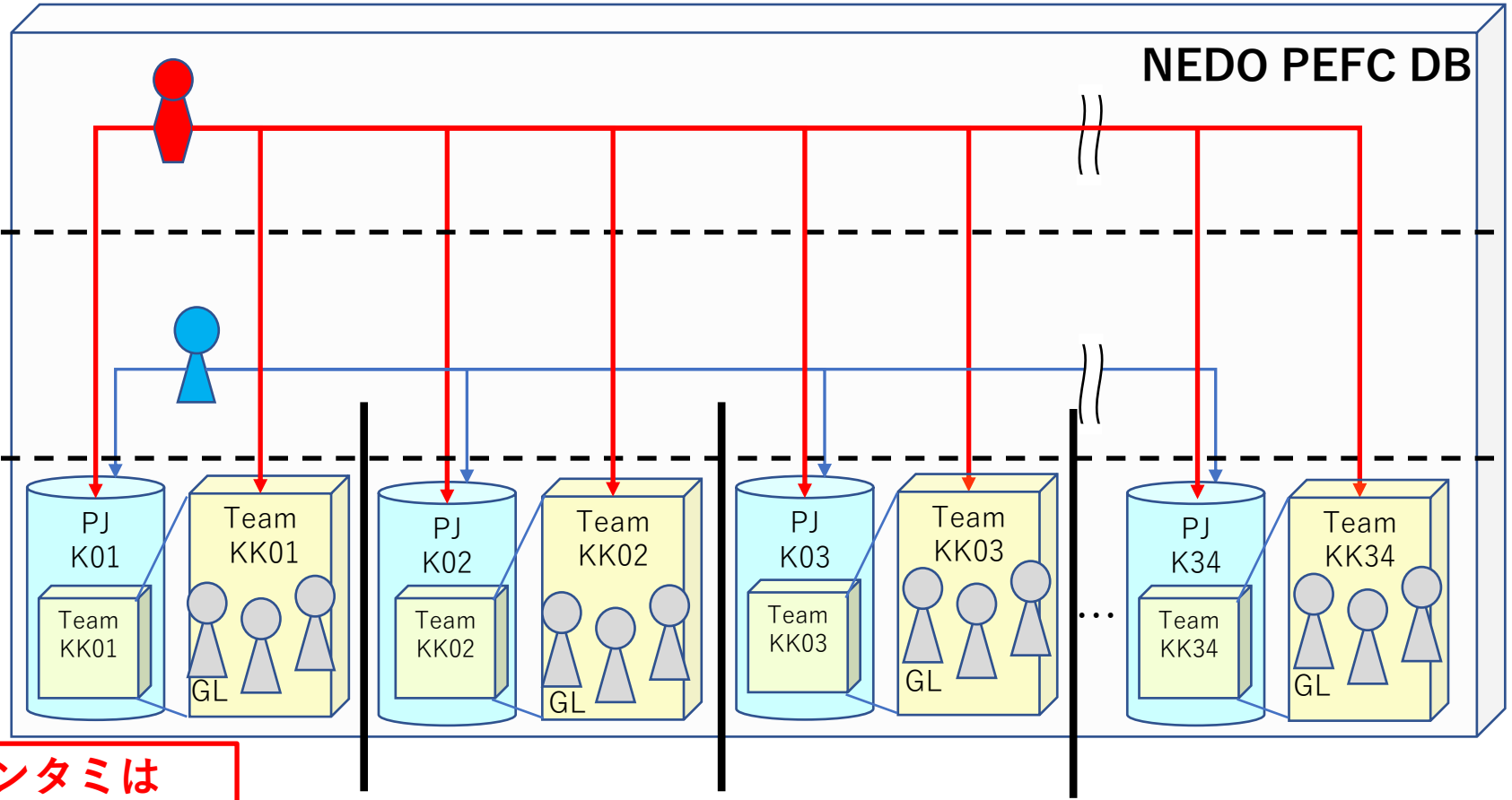
日本（NEDO事業）のメリット:
知の集約/ 材料設計に活用

MIX Platformで出来ること（DB）

PF マネジメントグループ(竹内)
〇メンバーの追加/削除と権限付与
〇プロジェクトの作成・削除
〇データの閲覧・編集・消去

PF MIグループ（袖山）
〇データの閲覧(編集は不可)

材料研究グループ（34事業のGL）
〇自分のPJのデータのみ閲覧可能
〇ＰＪ内でデータの書き込み・消去
〇グループ内の利用者をリスト化→読み書きできるユーザーを管理者が追加
×他のプロジェクトの閲覧



・異なるプロジェクトとのデータのコンタミはありません。
・自課題以外にデータにアクセスできるのは、現在袖山と竹内の二人だけです。

・運用は、基本的にMETI公開のガイドラインに沿って行います。

・PJ間のデータ(例えばK01とK02)を合わせてMI解析する場合がある。
(事前に事業者の許可を取ります。)

・システム管理者(外部運用ベンダー)とNIMSの間で秘密保持事項を契約書に記載

MIX Platformで出来ること（電気化学解析）

- **現在の電気化学測定結果の解析**

  １．実験機器から(機器付属のソフトを用いて)生データを得る。

  ２．測定者がExcelなどを用いて電気化学解析手法のプログラムを作る。

  ×測定者によりプログラム異なる。⇒出てくる値が異なる。

  ３．測定者がExcelなどを用いて電気化学解析手法を実施する。

  ×自分でプログラムを書いていない場合、使い方を覚える必要がある。

MIX Platformで出来ること（電気化学解析）

- **MIX platformによる電気化学測定結果の解析**

１．実験機器から得られる生データを読み込み可能。

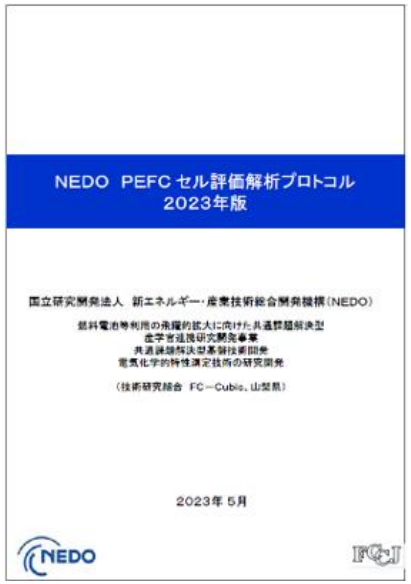
➢データの描画実施。HDFファイルに自動変換

➢現状は、フォーマットを限定。

２．FC-Cubicのプロトコルでよく使われる解析コードを実装。

３．データ蓄積時にクリックするだけで解析可能。

MIX Platformで出来ること（電気化学解析）



最新版のNEDO 評価解析プロトコルには以下のMEAの評価方法が記載されている。

（3）電極触媒・電極触媒層

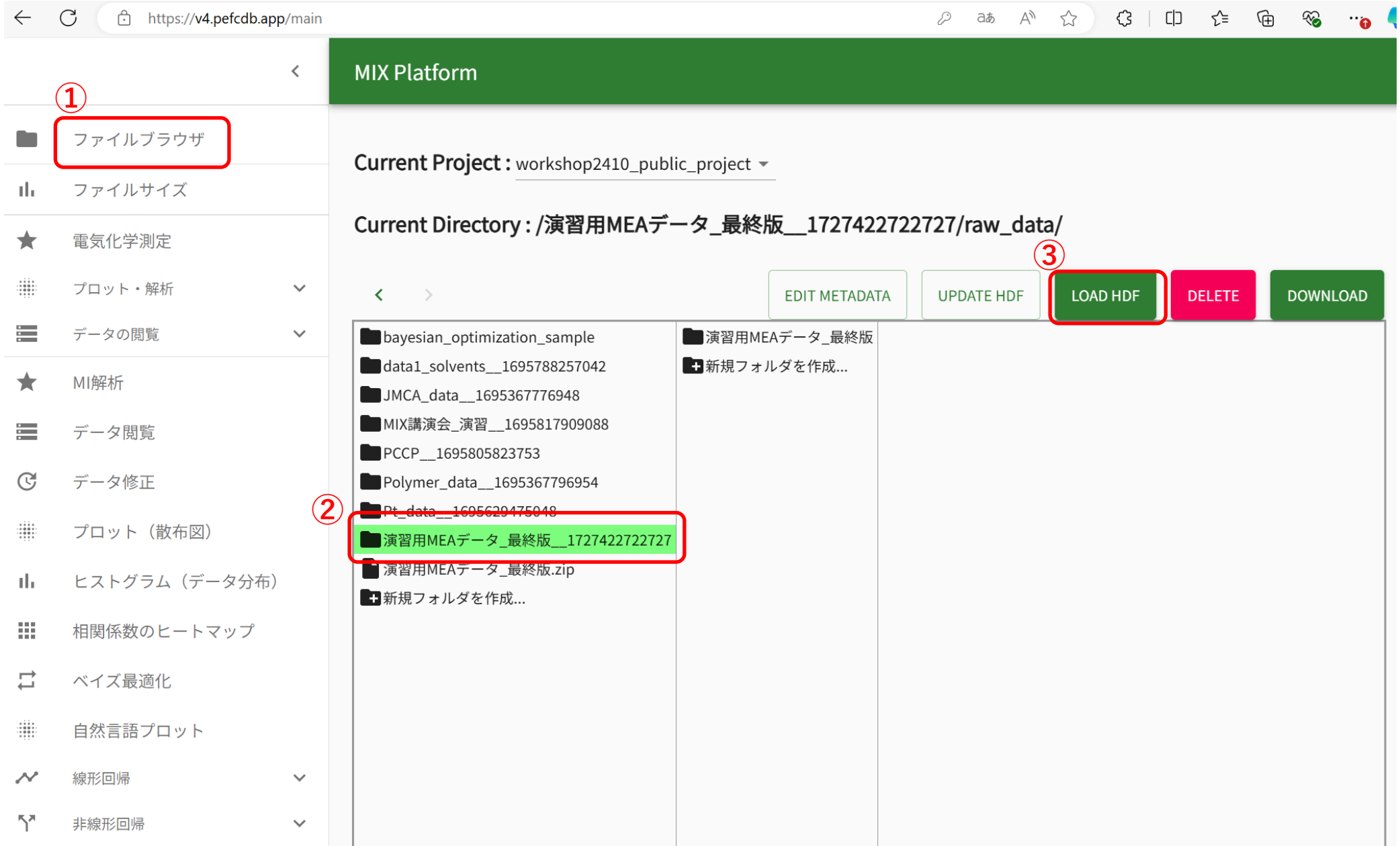
A. 評価方法

・C -1	試験名：I-V測定方法・過電圧分離解析方法
・C- 2（III-3-4）	試験名：ECA測定方法
・C- 3	試験名：クロスオーバー電流測定方法
・C- 4（III-3-2）	試験名：ORR活性評価方法
・C- 5	試験名：HOR性能（アノード触媒）
・C- 6	試験名：酸素拡散抵抗
・C- 7	試験名：プロトン輸送抵抗
・C- 8	試験名：高温評価（セル温度120℃まで）

電気化学的特性測定GrがMEA評価として標準的に実施している評価方法のうちよく利用される以下のものをMIX platformに実装した。

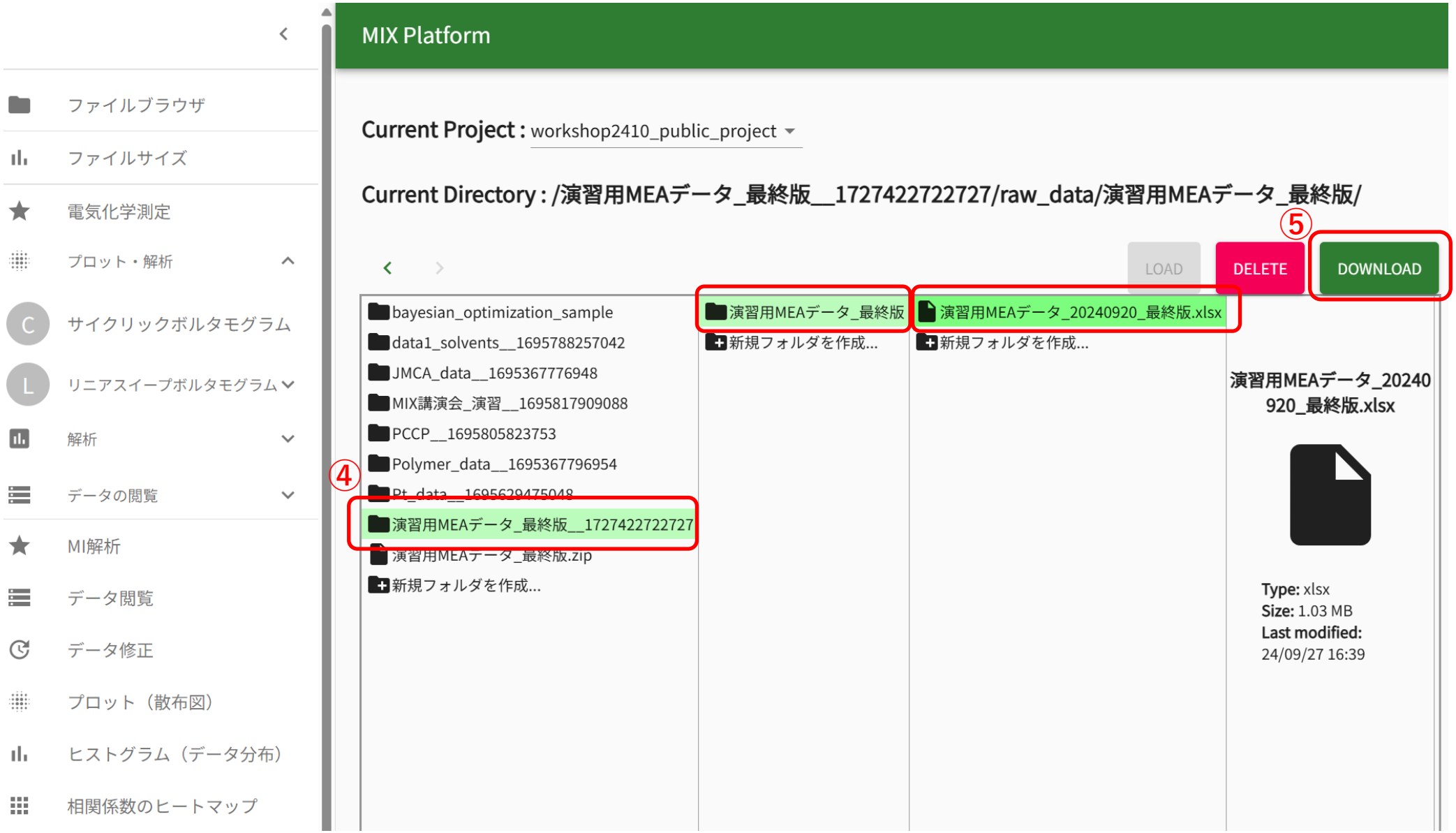
- ECSA評価（CV測定）
- LSV測定の解析手法（クロスオーバー電流(COC)測定方法）
- *I-V*測定測定の解析手法（I-V測定方法・過電圧分離解析方法）
- ORR活性測定の解析手法（ORR活性評価方法）

MIX Platformで出来ること（電気化学解析）



①「ファイルブラウザ」をクリック
②「演習用MEAデータ」をクリック
③「LOAD HDF」をクリック

MIX Platformで出来ること（電気化学解析）



④「演習用MEAデータ」…「演習用MEAデータ.xlsx」をクリック
⑤「DOWNLOAD」をクリック
⇒今日使うExcelファイルがローカルPCにダウンロードされる。

MIX Platformで出来ること（電気化学解析）

A	B	C	D	E	F	G
時間[sec]	電圧[V]	電流[A]	電流密度[mA/cm2]	サイクル		
24.00259939	0.49965489	0.004702222	1.56740729	1		
24.00809939	0.49942899	0.001690158	0.563386072	1		
24.01809939	0.4989104	0.001724736	0.574911911	1		
24.02809939	0.49838904	0.001586235	0.52874507	1		
24.03809939	0.49790987	0.001563894	0.521297985	1		
24.04809939	0.49739963	0.001526072	0.508690642	1		
24.05809939	0.49690184	0.001556885	0.518961737	1		
24.06809939	0.49639705	0.001539803	0.513267728	1		
24.07809939	0.49591225	0.001525499	0.508499789	1		
24.08809939	0.49539247	0.001532964	0.510988042	1		
24.09809939	0.49490616	0.001517367	0.505788877	1		
24.10809939	0.49441218	0.001545508	0.515169288	1		
24.11809939	0.49389553	0.001554033	0.518010979	1		
24.12809939	0.49340263	0.001474073	0.491357757	1		
24.13809939	0.49289465	0.00147367	0.491223416	1		
24.14809939	0.49241871	0.001515829	0.505276371	1		
24.15809939	0.49189886	0.001537979	0.512659823	1		
24.16809939	0.49139088	0.001426772	0.475590612	1		
24.17809939	0.49089423	0.001517494	0.505831276	1		
24.18809939	0.49039876	0.001500783	0.500261005	1		
24.19809939	0.48991707	0.001439019	0.4796729	1		
24.20809939	0.48940638	0.001467807	0.489268907	1		
24.21809939	0.48891461	0.001530854	0.510284666	1		

CV-metadata_40deg　CV-data_40deg　LSV-metadata_80deg

相関係数のヒートマップ

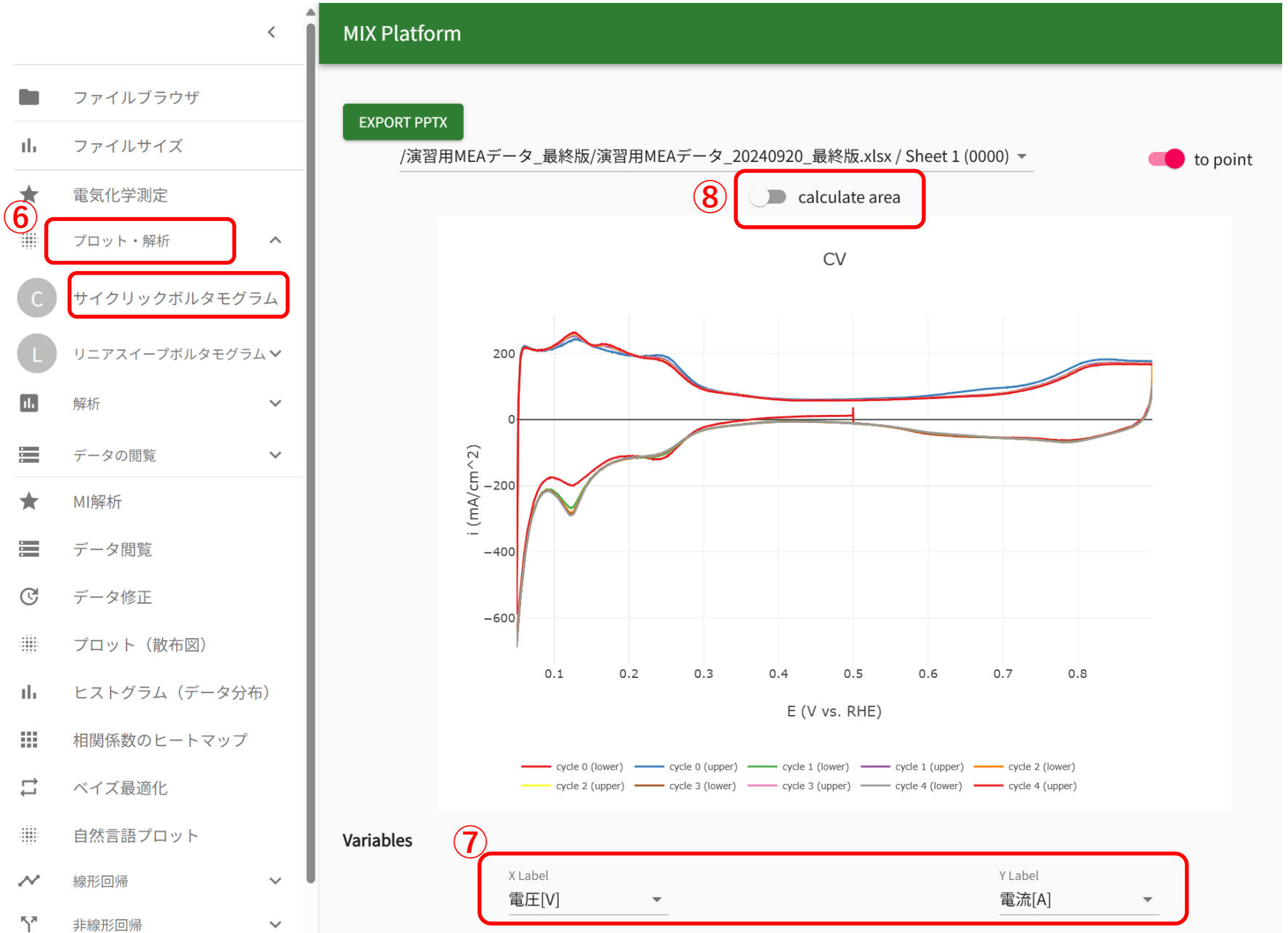
④「演習用MEAデータ」…「演習用MEAデータ.xlsx」をクリック
⑤「DOWNLOAD」をクリック
⇒今日使うExcelファイルがローカルPCにダウンロードされる。

←実験データファイルの中身を確認

Sheet1: CV
Sheet3: LSV (COC)
Sheet5: ORR
Sheet7: IV

が記載されている。

MIX Platformで出来ること（電気化学解析）

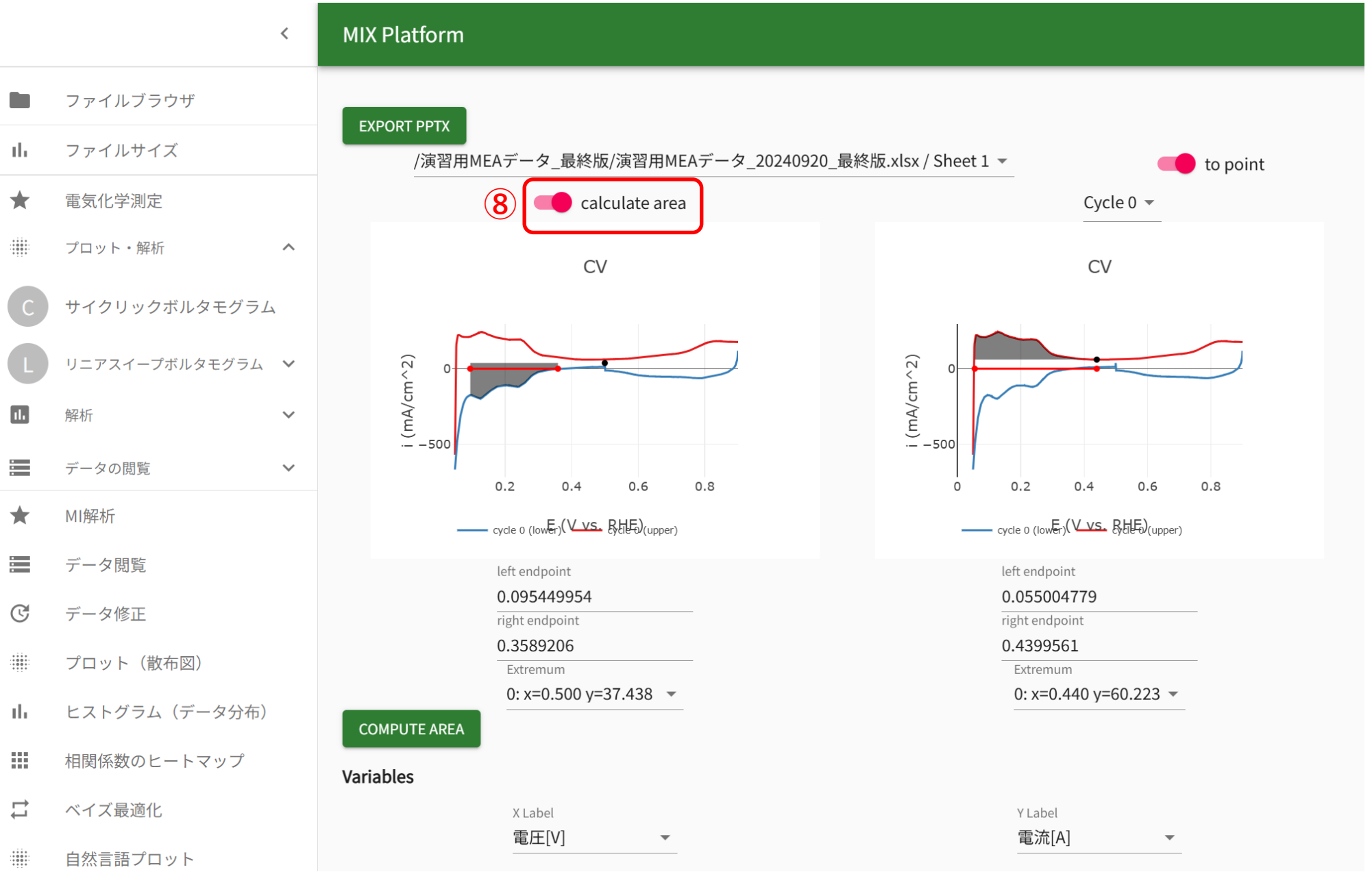


⑥「プロット・解析」、「サイクリックボルタモグラム」をクリック

⑦ X Labelを「電圧」、Y Labelを「電流」にセットする。

⑧Calculate areaをオンにする。

MIX Platformで出来ること（電気化学解析）

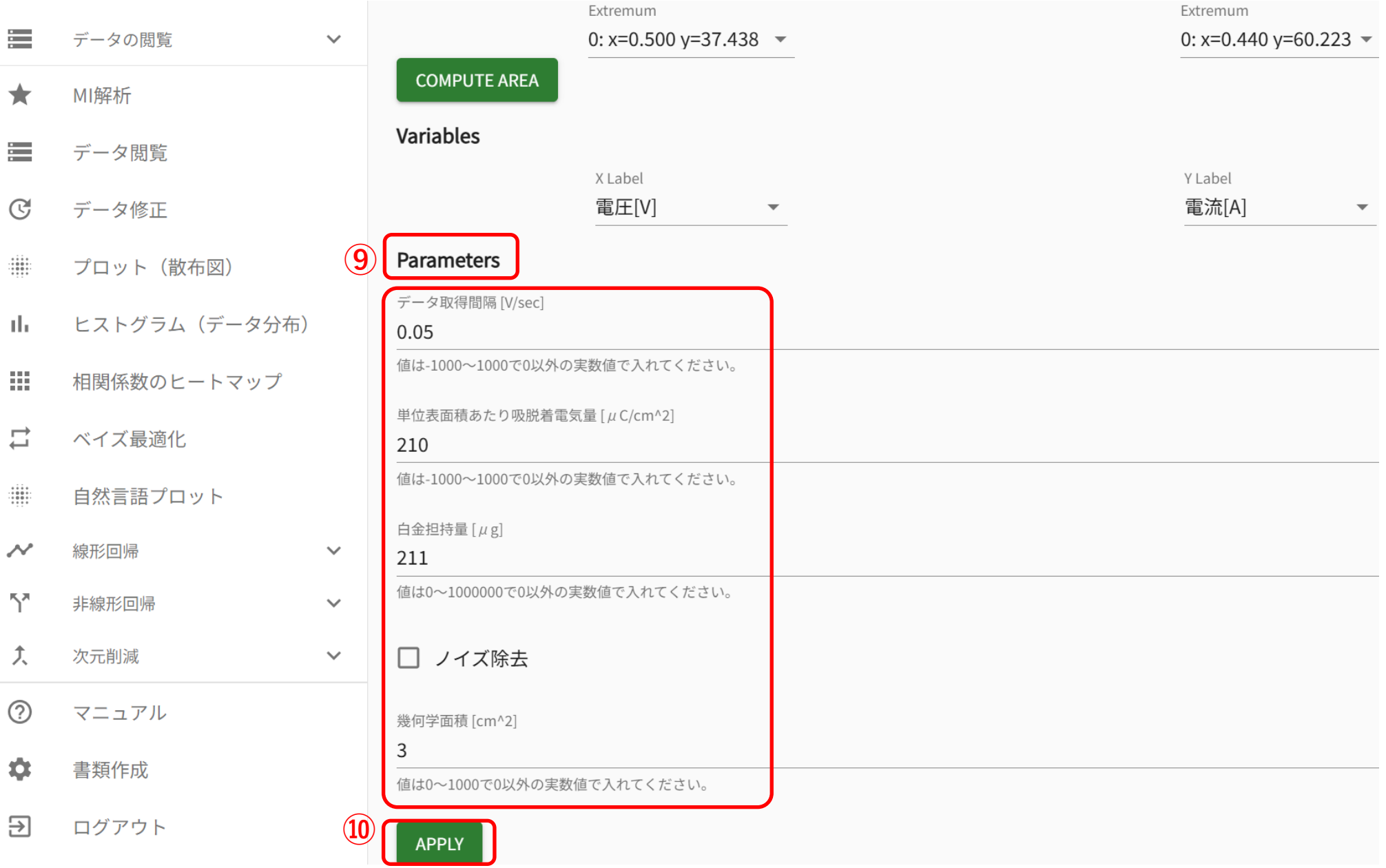


⑥「プロット・解析」、「サイクリックボルタモグラム」をクリック

⑦ X Labelを「電圧」、Y Labelを「電流」にセットする。

⑧Calculate areaをオンにする。
⇒左のような画面になる。
⑨下の方にあるParametersをセットする。（次ページ）

MIX Platformで出来ること（電気化学解析）

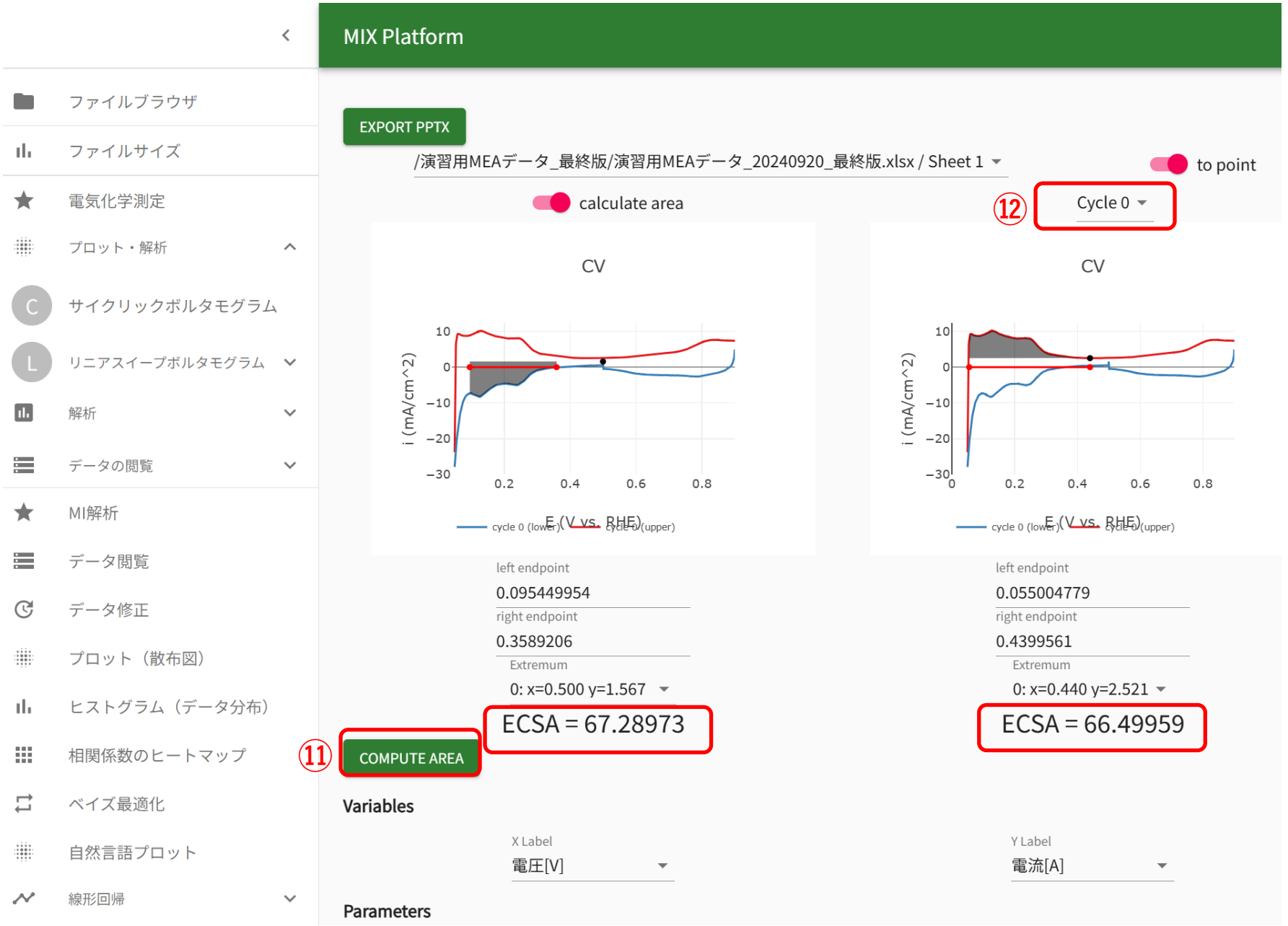


⑧Calculate areaをオンにする。
⇒左のような画面になる。

⑨Parametersを実験条件に合わせて入力する。（今回は左記のとおりの数値を入力）

⑩「APPLY」をクリック

MIX Platformで出来ること（電気化学解析）



⑩「APPLY」をクリック

⑪「COMPUTE AREA」をクリック
⇒ECSAの値が計算される。

⑫「Cycle 0」を他のサイクルに変更してCOMPUTE AREAをクリックするとサイクルごとのECSAが計算される

MIX Platformで出来ること（LSV測定COC解析）

目的

- 水素クロスオーバー電流 j_{XO}の見積もり
  →電解質膜のガスバリア特性、膜劣化時の指標
- 短絡抵抗 R_{SC}の測定→セル組立時に不備がないかなど

測定方法　評価解析プロトコル, C-3, p21

電位スイープ範囲・速度	0.2～0.5V vs. RHE の範囲を低電位側から 0.5mV/s でスイープする。
リーク電流の算出方法	0.2～0.5V の電位間をスイープしたときに得られる電位/電流密度曲線の 0.4～0.5V 近傍の直線部を 0V まで外挿したときの切片の電流密度が クロスオーバー電流となり、傾きの逆数が電子抵抗（短絡抵抗）となる。

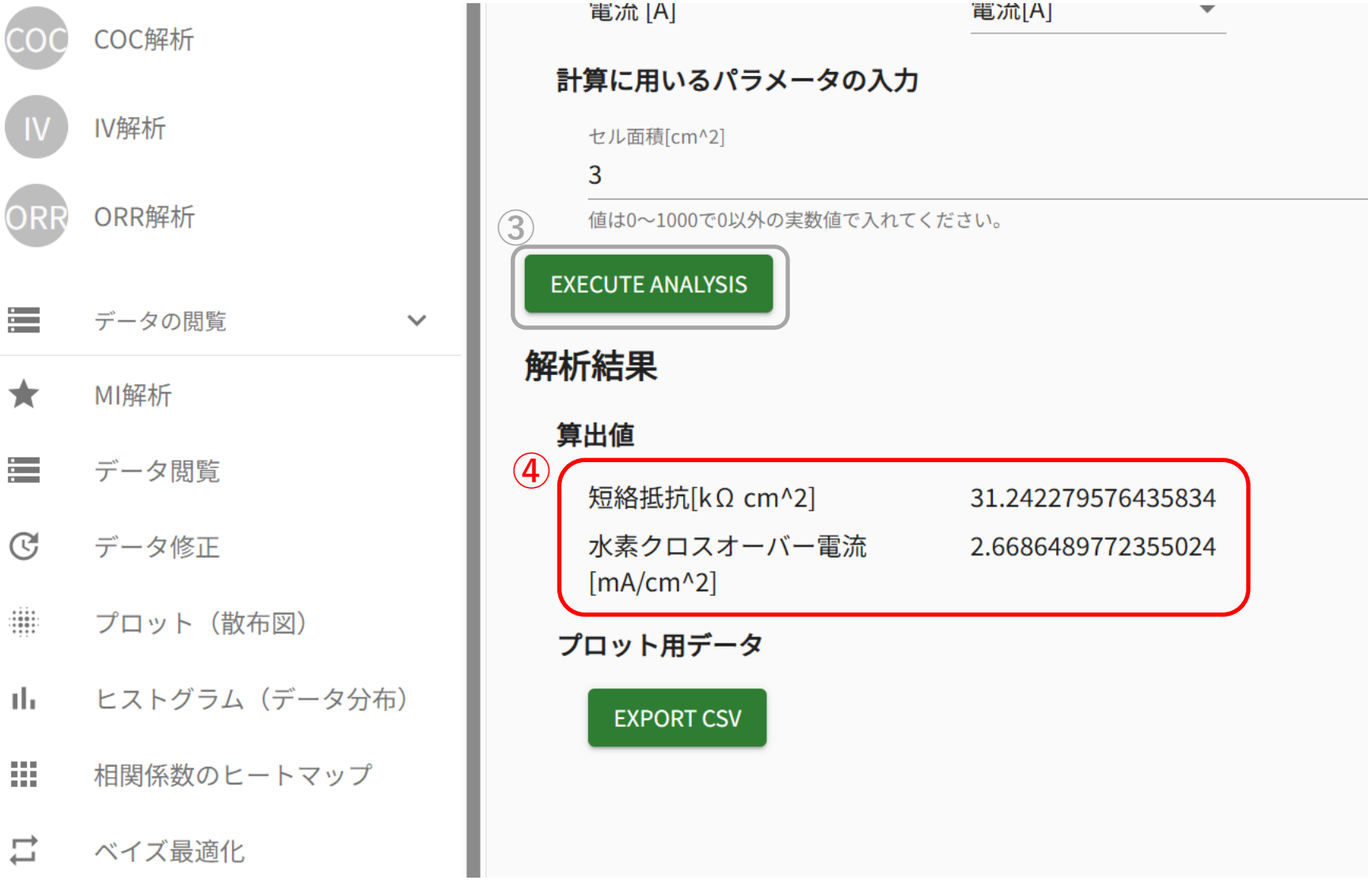
MIX Platformで出来ること（LSV測定COC解析）

①解析に用いるファイルで「Sheet3」を選択。

②解析に用いるデータの選択で、左記のように項目を選ぶ。（選択候補に挙がる項目は、Excelファイルの項目名が表示される）。

③「EXECUTE ANALYSIS」をクリック。

MIX Platformで出来ること（LSV測定COC解析）



①解析に用いるファイルで「Sheet3」を選択。

②解析に用いるデータの選択で、左記のように項目を選ぶ。（選択候補に挙がる項目は、Excelファイルの項目名が表示される）。

③「EXECUTE ANALYSIS」をクリック。

④短絡抵抗と水素クロスオーバー電流の値が計算される。

MIX Platformで出来ること（IV解析）

目的

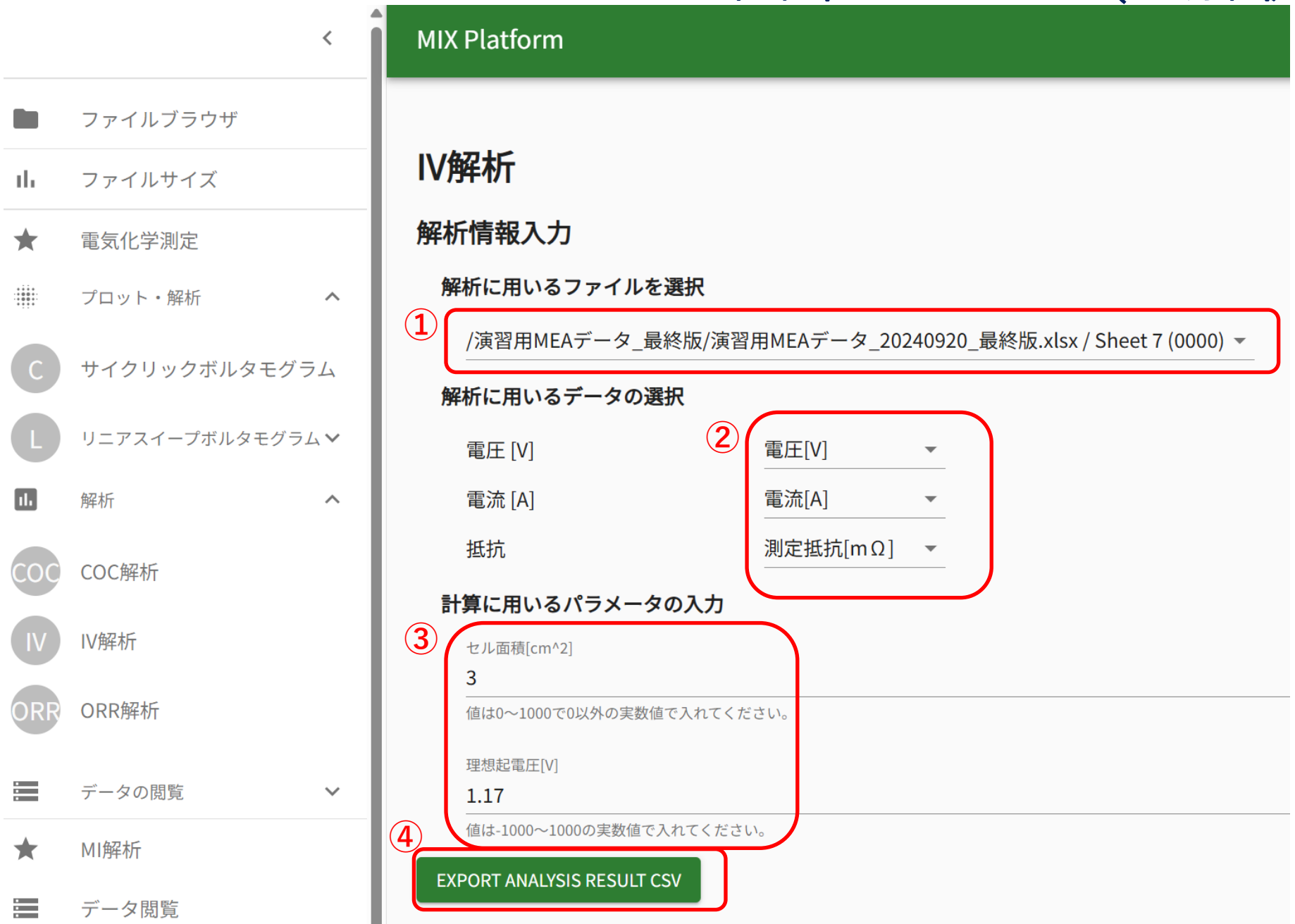
- MEAのIV特性評価→触媒、電解質膜、GDLなどの全構成要素の正味のIV性能
- 過電圧分離→触媒活性、電解質膜抵抗、酸素輸送抵抗の要素を分離

測定方法　評価解析プロトコル, C-1, p19

	1cm×1cmセル	1cm×3cmセル	JARI標準セル
電流密度と保持時間 低電流密度から 高電流密度へシフト	①** 2,4,6,10,20,50,75,100 mA/cm² (各5min) ②200,400,600,800,1000, 1500,2000mA/cm² (各10min)	① 2,4,6,10,20,50,75,100 mA/cm² (各5min) ② 200,400,600,800,1000, 1200 mA/cm² (各10min)	① 25,50,75,100 mA/cm² (各5min) ② 200,400,600,800,1000, 1200 mA/cm² (各10min)

⇐各電流密度での定常状態の電圧を計測

MIX Platformで出来ること（IV解析）



①解析に用いるファイルで「Sheet7」を選択。

②解析に用いるデータの選択で、左記のように項目を選ぶ。（選択候補に挙がる項目は、Excelファイルの項目名が表示される）。

③計算に用いるパラメータを左記のように入力。

④「EXPORT ANALYSIS RESULT CSV」をクリック。

MIX Platformで出来ること（ORR解析）

目的　・触媒性能の指標となる質量活性と面積比活性を算出する。

測定方法　評価解析プロトコル, C-4, p22~23

	1cm×1cmセル	JARI標準セル
電流密度と保持時間 高電流密度から 低電流密度へシフト	① 0 mA/cm² （O_2導入後2min） ② 1000, 200, 100, 80, 60, 40, 20 mA/cm²（各15min） 必要に応じて ③ 10, 6, 4, 2 mA/cm²（各5min）	① 0 mA/cm² （O_2導入後2min） ② 1000, 200, 100, 80, 60, 40, 20 mA/cm²（各15min）

MIX Platformで出来ること（ORR解析）

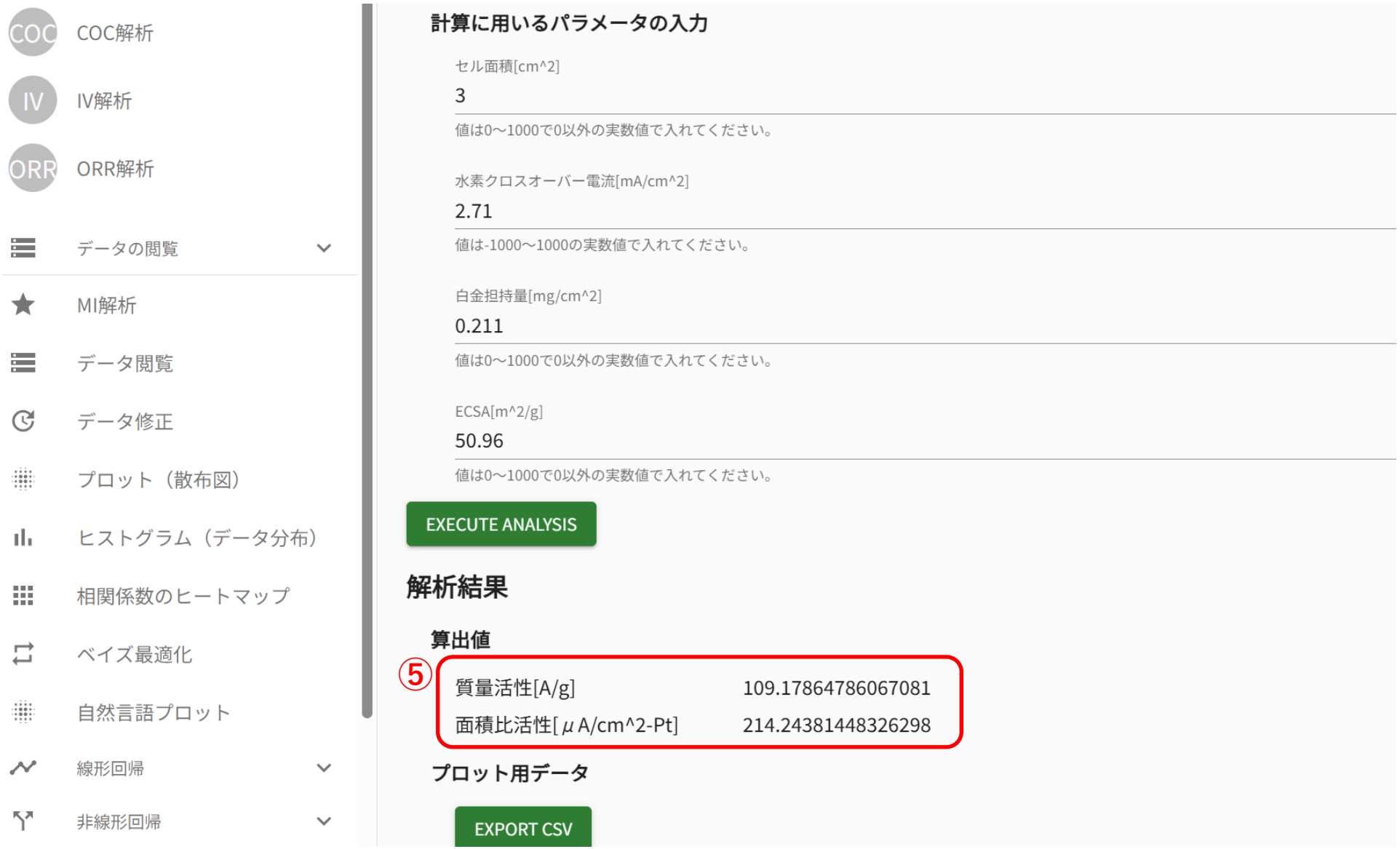
①解析に用いるファイルで「Sheet5」を選択。

②解析に用いるデータの選択で、左記のように項目を選ぶ。（選択候補に挙がる項目は、Excelファイルの項目名が表示される）。

③計算に用いるパラメータを左記のように入力。

④「EXECUTE ANALYSIS」をクリック。

MIX Platformで出来ること（ORR解析）



①解析に用いるファイルで「Sheet5」を選択。

②解析に用いるデータの選択で、左記のように項目を選ぶ。（選択候補に挙がる項目は、Excelファイルの項目名が表示される）。

③計算に用いるパラメータを左記のように入力。

④「EXECUTE ANALYSIS」をクリック。

⑤質量活性と面積比活性の値が計算される。

MIX Platformで出来ること（MI解析）

1．表形式のデータを読み込み（.csvファイル）

➢既にきれいな形に揃っているデータを想定。

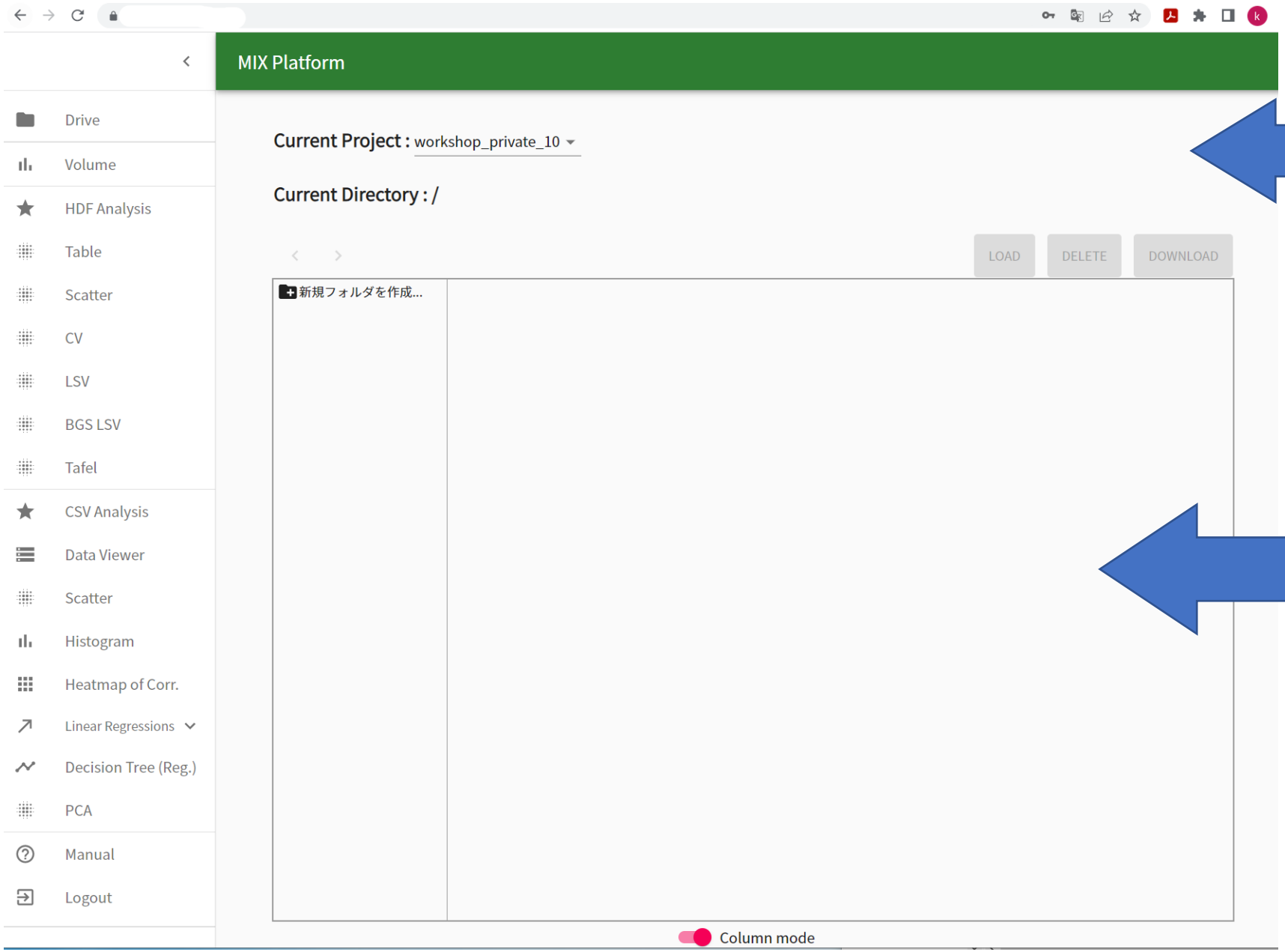
2．データのMI解析が可能

➢散布図、ヒストグラム分布、相関係数のヒートマップを
ウェブアプリ上で描画可能。

➢線形回帰（重回帰、LASSO、Ridge回帰)を利用可能。
欲しい物性値の予測と、何が予測に効くかの要因解析を
ウェブアプリ上で実施。

1．非線形回帰手法(**決定木**、**Random Forest**、**XGBoost**)の利用
2．**全状態探索**（線形回帰）手法の利用
3．**ベイズ最適化**の利用
4．主成分分析、UMAPの利用
5．その他（データ上書き機能、PPTファイル出力機能...）

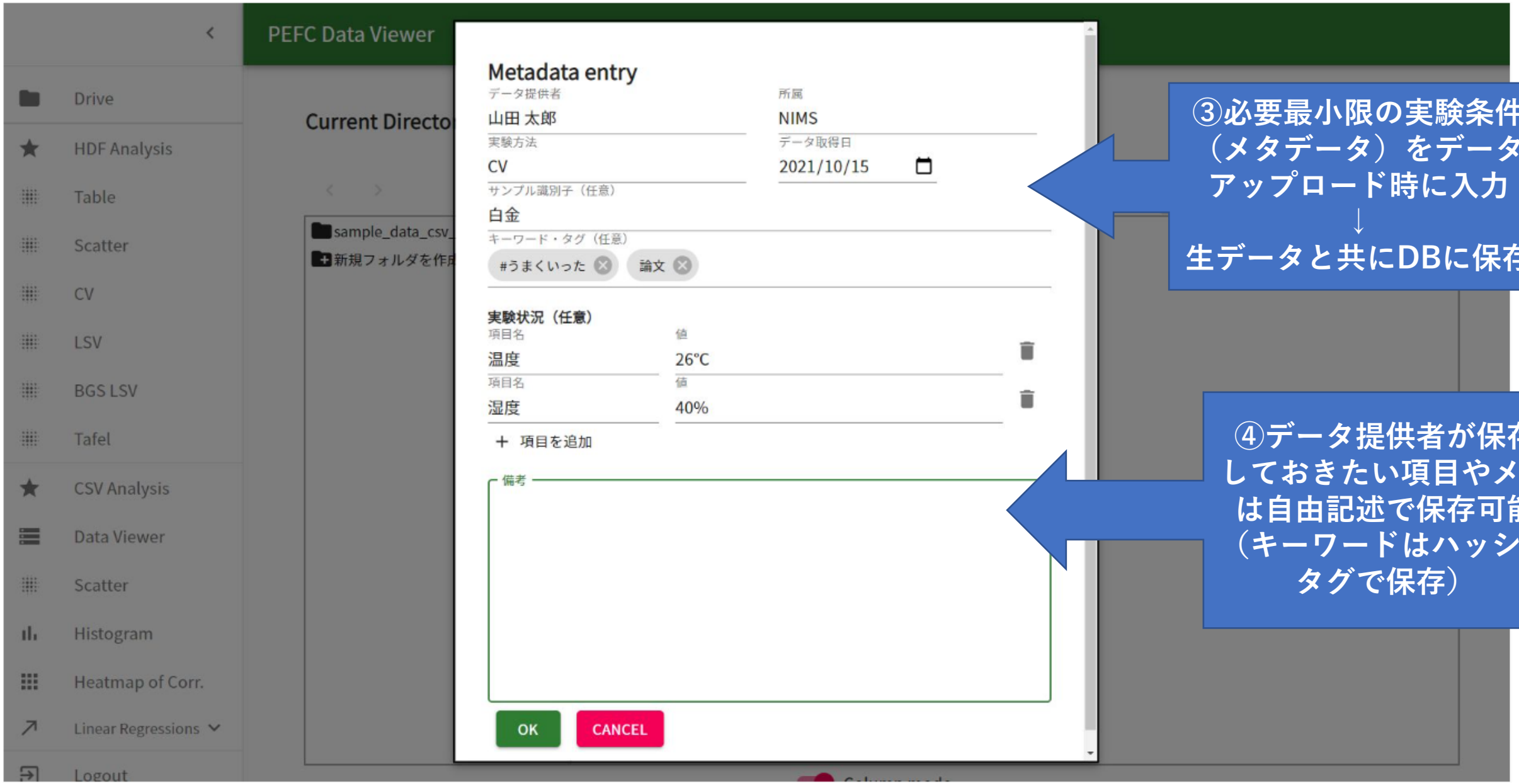
MIX Platformでできること（DBへの蓄積１）



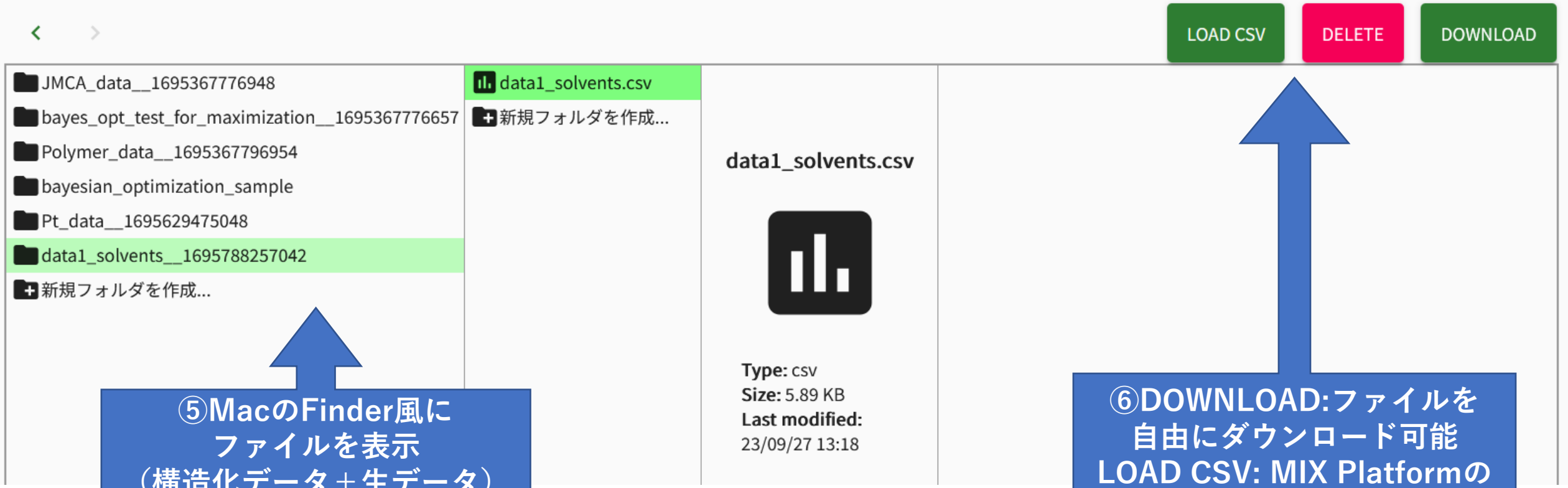
①ウェブブラウザから
アクセス可能
（windowsやMacで
利用可能）

②Drag & dropで
データを簡単に
アップロードできる
（テキストやエクセル
ファイルに対応）
↓
自動的にファイルの中身
を読み取って構造化

MIX Platformでできること（DBへの蓄積２）



MIX Platformでできること（DBへの蓄積３）

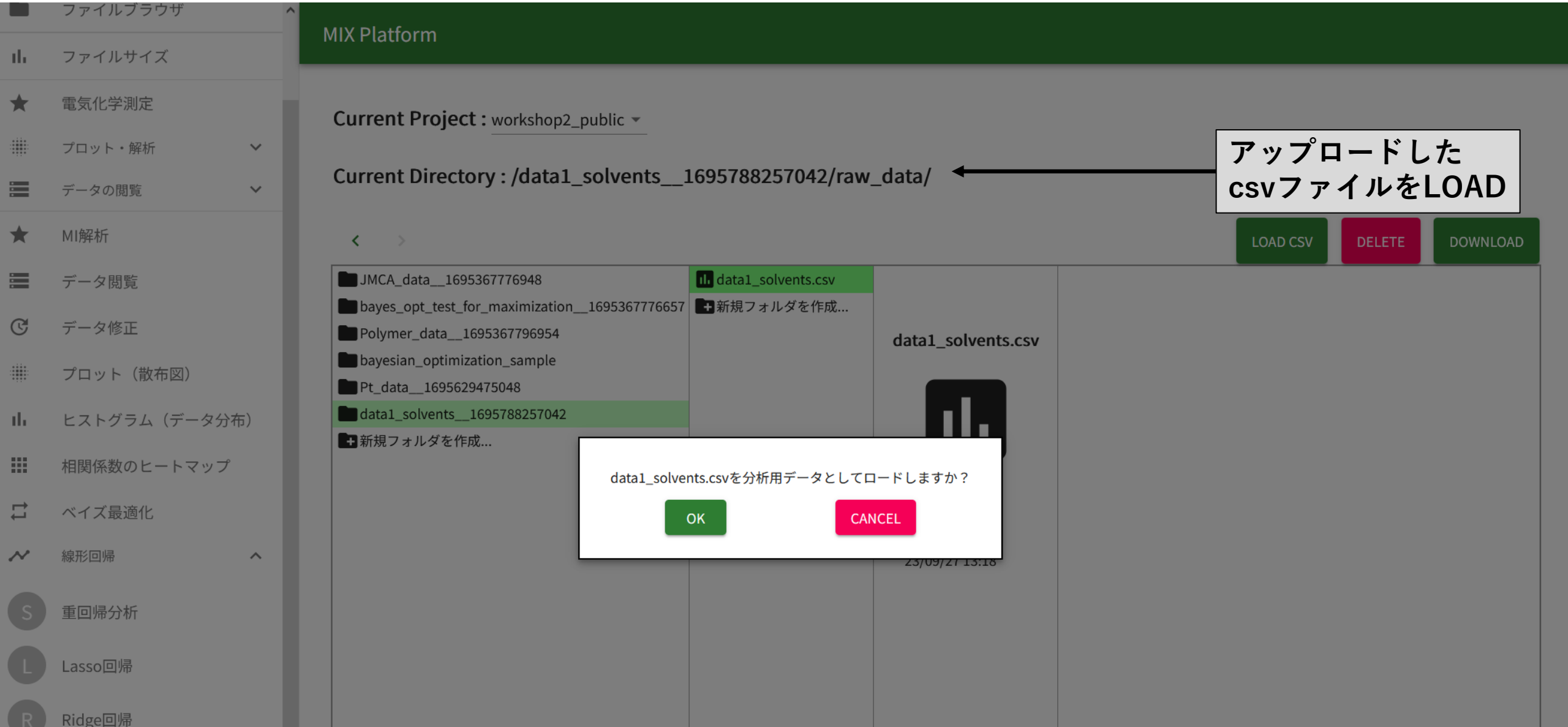


⑤MacのFinder風に
ファイルを表示
（構造化データ＋生データ）

⑥DOWNLOAD:ファイルを
自由にダウンロード可能
LOAD CSV: MIX Platformの
MI解析機能を利用可能

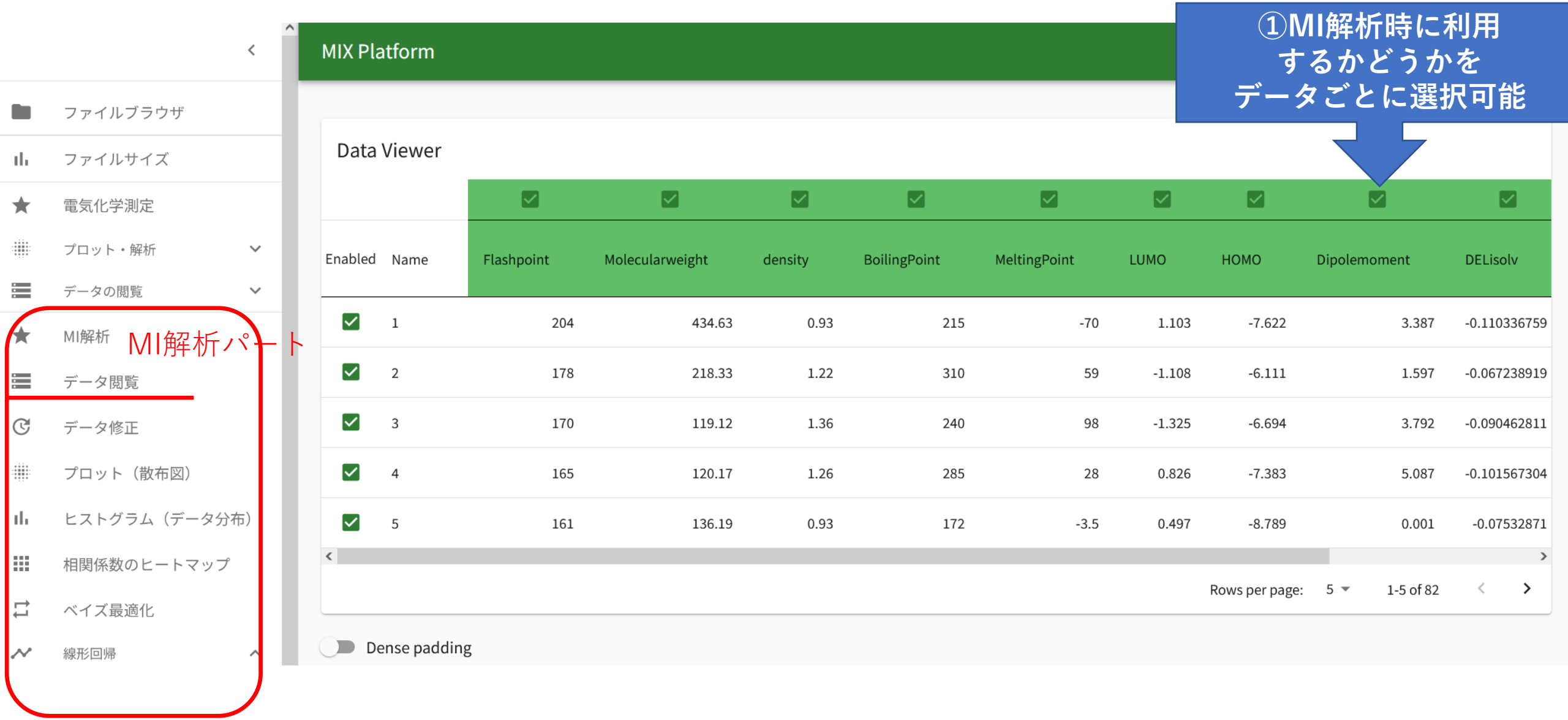
MIX Platformでできること（MI解析1）

CSVファイルの読込(MI解析はcsvファイルのみ対応)



MIX Platformでできること（MI解析2）

Data Viewer（アップロードしたデータの確認、特徴量の選択/非選択）



MIX Platformでできること（MI解析3）

散布図の描画(Scatter)

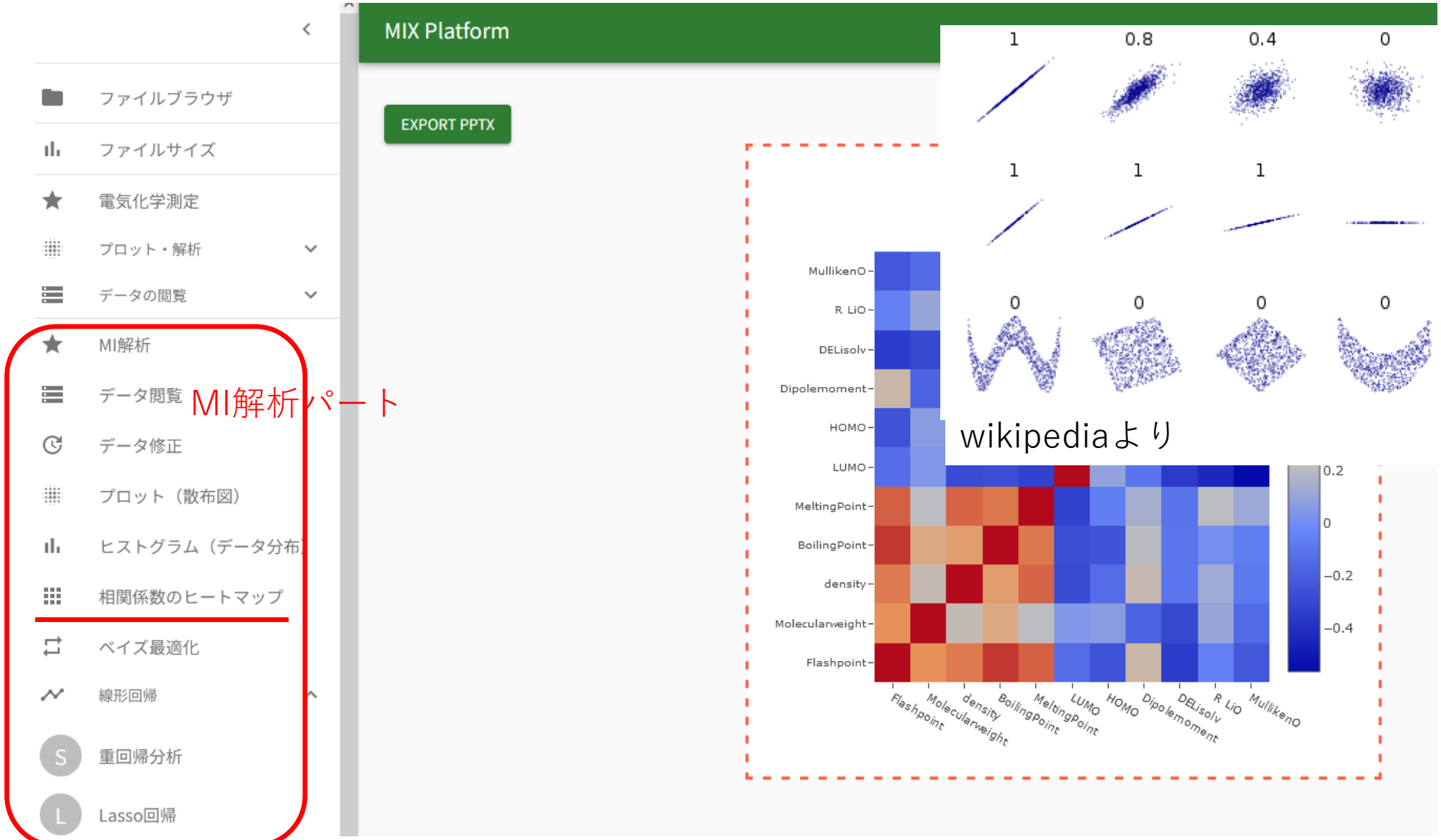
MIX Platformでできること（MI解析4）

ヒストグラム

アップロードデータの分布を描画

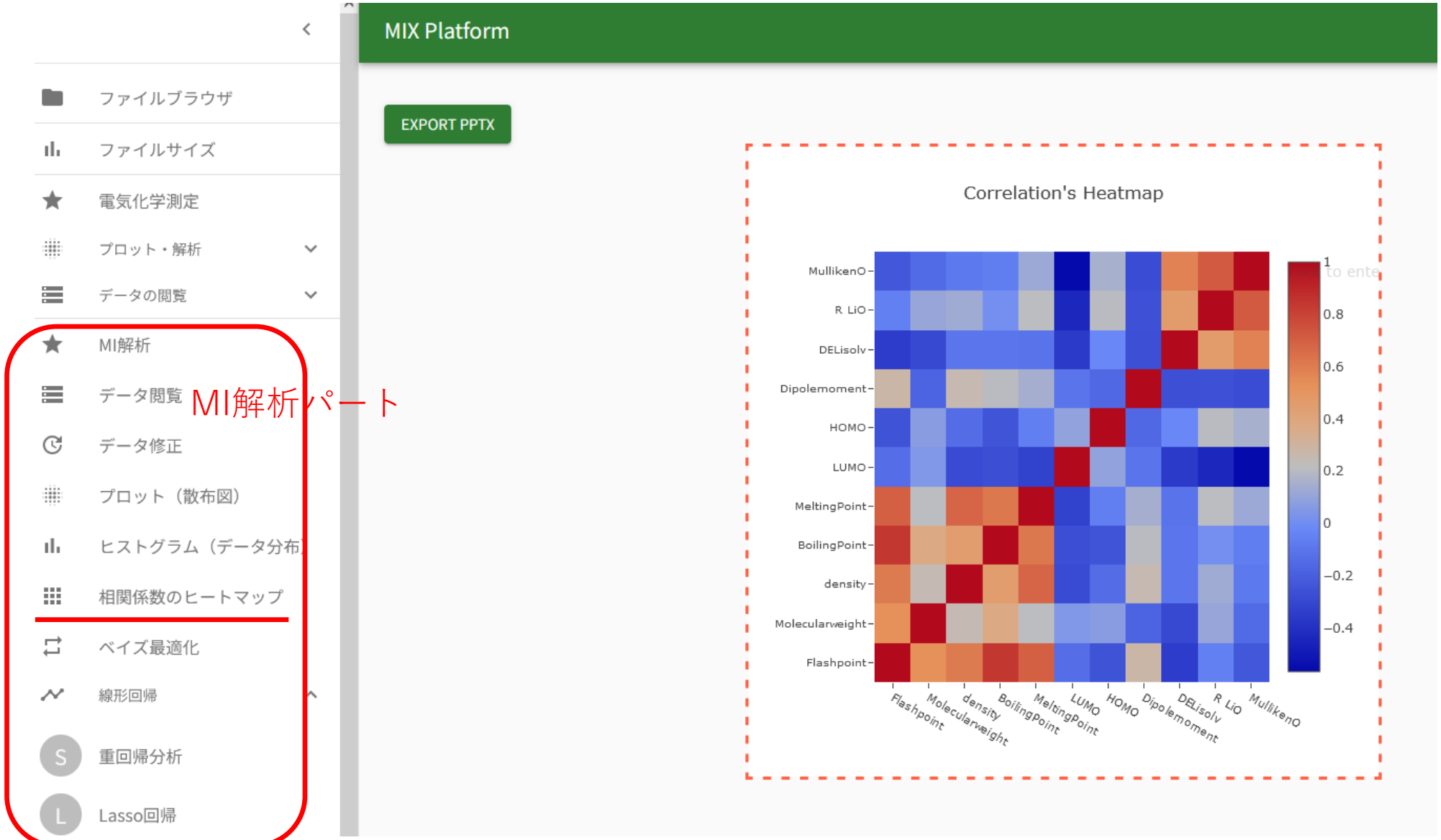
MIX Platformでできること（MI解析5）

相関係数のヒートマップ(Pearsonの相関係数)



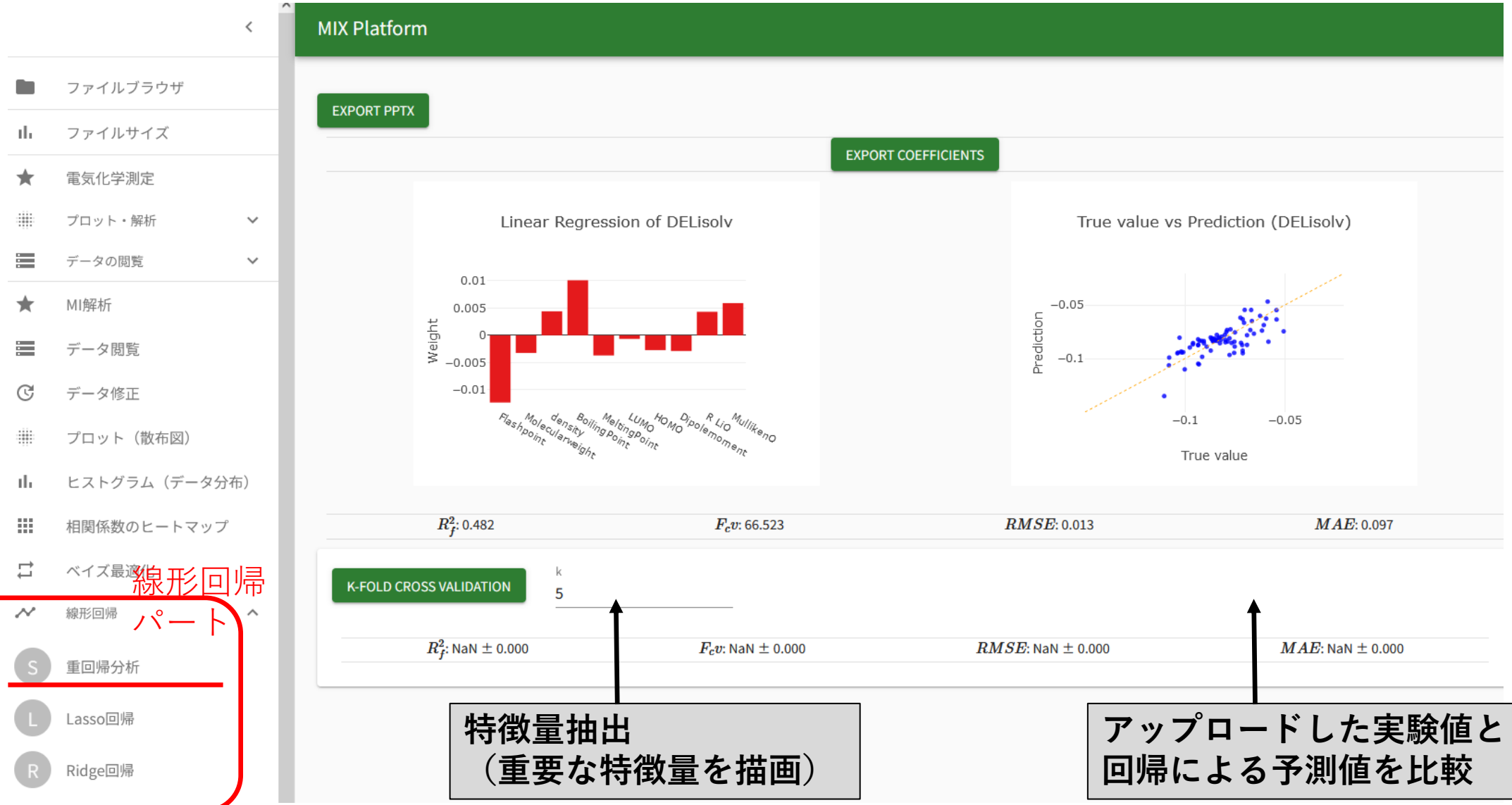
MIX Platformでできること（MI解析5）

相関係数のヒートマップ(Pearsonの相関係数)



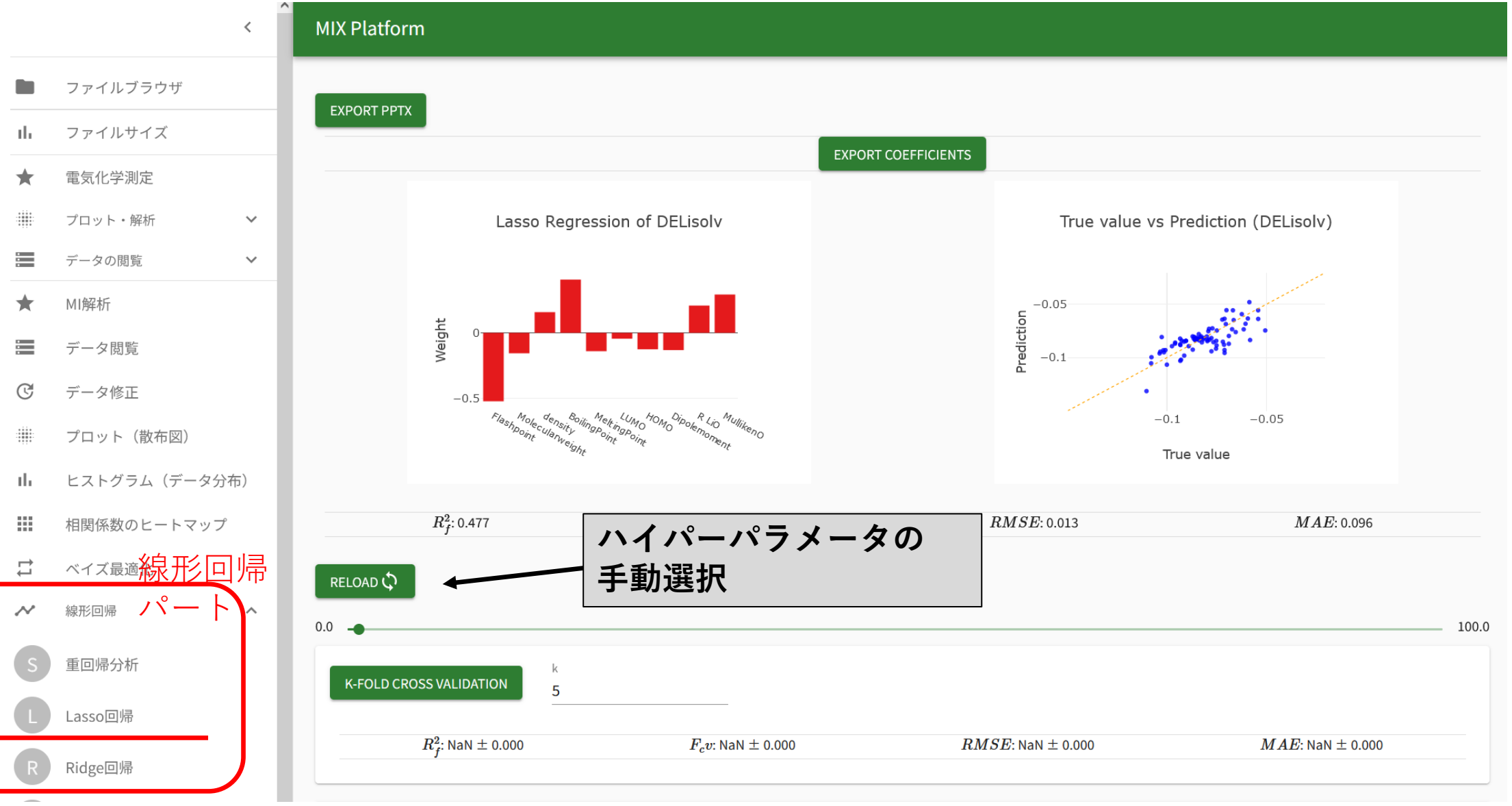
MIX Platformでできること（MI解析6）

線形回帰分析(重回帰分析)



MIX Platformでできること（MI解析7）

線形回帰分析(LASSO分析)

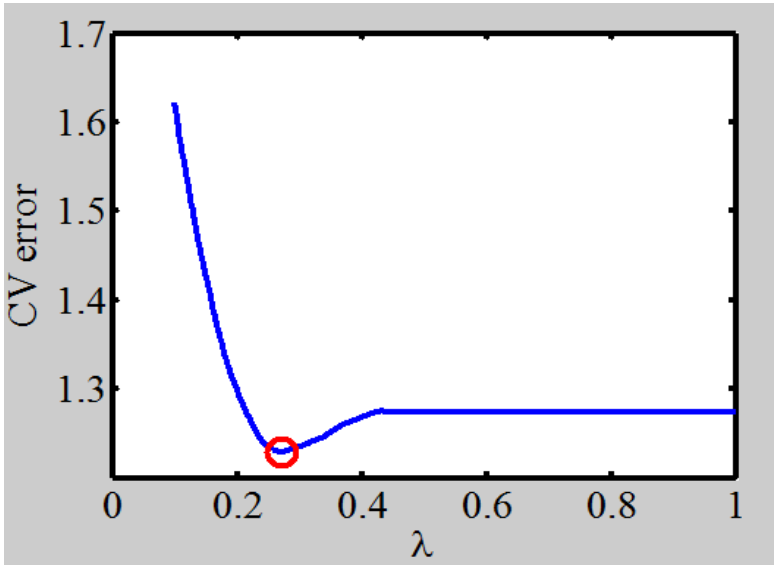


MIX Platformでできること（MI解析7）

線形回帰分析(LASSO分析)

LASSO (least absolute shrinkage and selection operator)

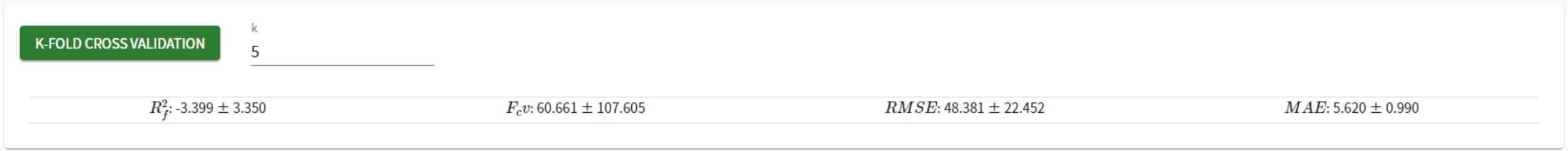
$$E = \sum_{\mu=1}^{103} (z^{\mu} - \sum_{i=1}^{10} w_i x_i^{\mu})^2 + \lambda \sum_{i=1}^{10} |w_i|$$



・パラメータλを変化させることで予測精度を上げる

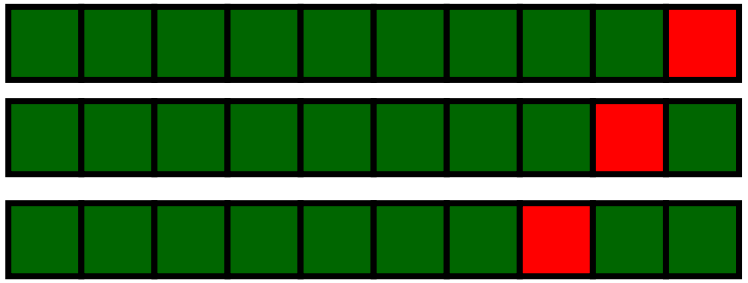
MIX Platformでできること（MI解析7）

線形回帰分析(LASSO分析)　クロスバリデーション



Cross Validation (CV)

1 10 20 30 40 50 60 70 80 90 100



：回帰用（学習用）データ

：誤差確認用（テスト用）データ

100点あるデータを全て使って予測モデルを作成
→新しいデータの予測性能がわからない。

100点データがあるうちの、90点を使って予測モデル作成。
残り10点を使って予測モデルにあてはめ、精度検証。

どの10点を残すかを変えながら10パターン計算を行う。

MIX Platformでできること（MI解析7）

線形回帰分析(LASSO分析)　ハイパーパラメータ最適化

PARAMETER OPTIMIZATION

α min	α max	α count
0	100	100

number of α values to try

k	Holdout ratio	Target loss
5	0.2	RMSE

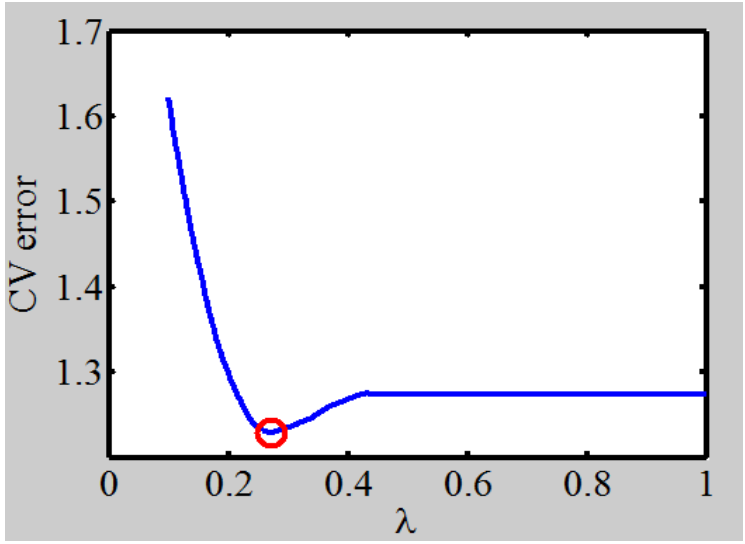
cross validation

Holdout

R_f^2: -3.520	F_cv: -41.628	$RMSE$: 55.849	MAE: 6.452

Best cross-validation error

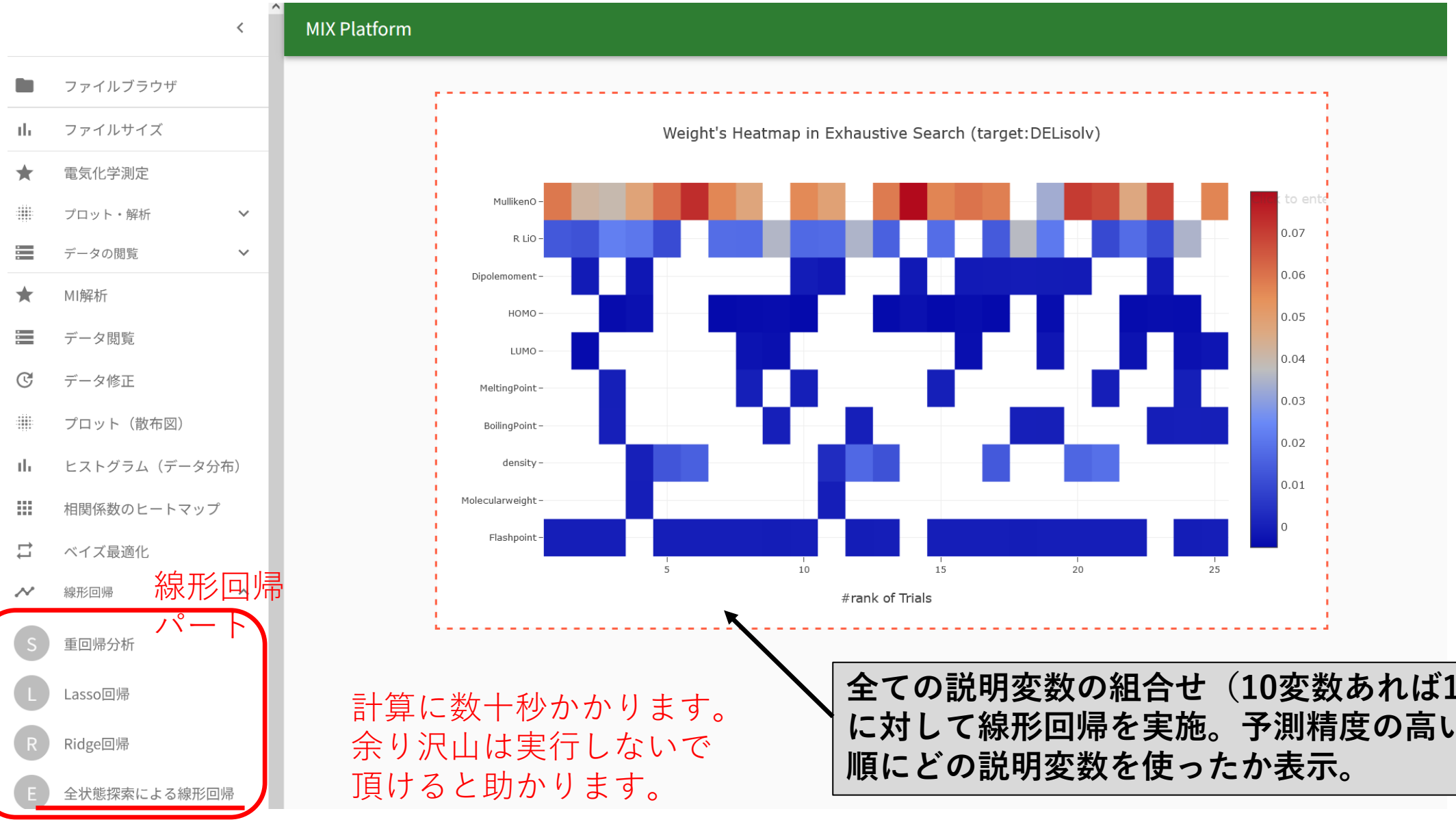
R_f^2: -3.409 ± 1.095	F_cv: 7.290 ± 78.316	$RMSE$: 28.590 ± 7.313	MAE: 4.555 ± 0.331



・パラメータλを変化させることで予測精度を上げる
・ λをクロスバリデーションによるCV errorの最小値をとるλとして選ぶ

MIX Platformでできること（MI解析8）

線形回帰分析(全状態探索)



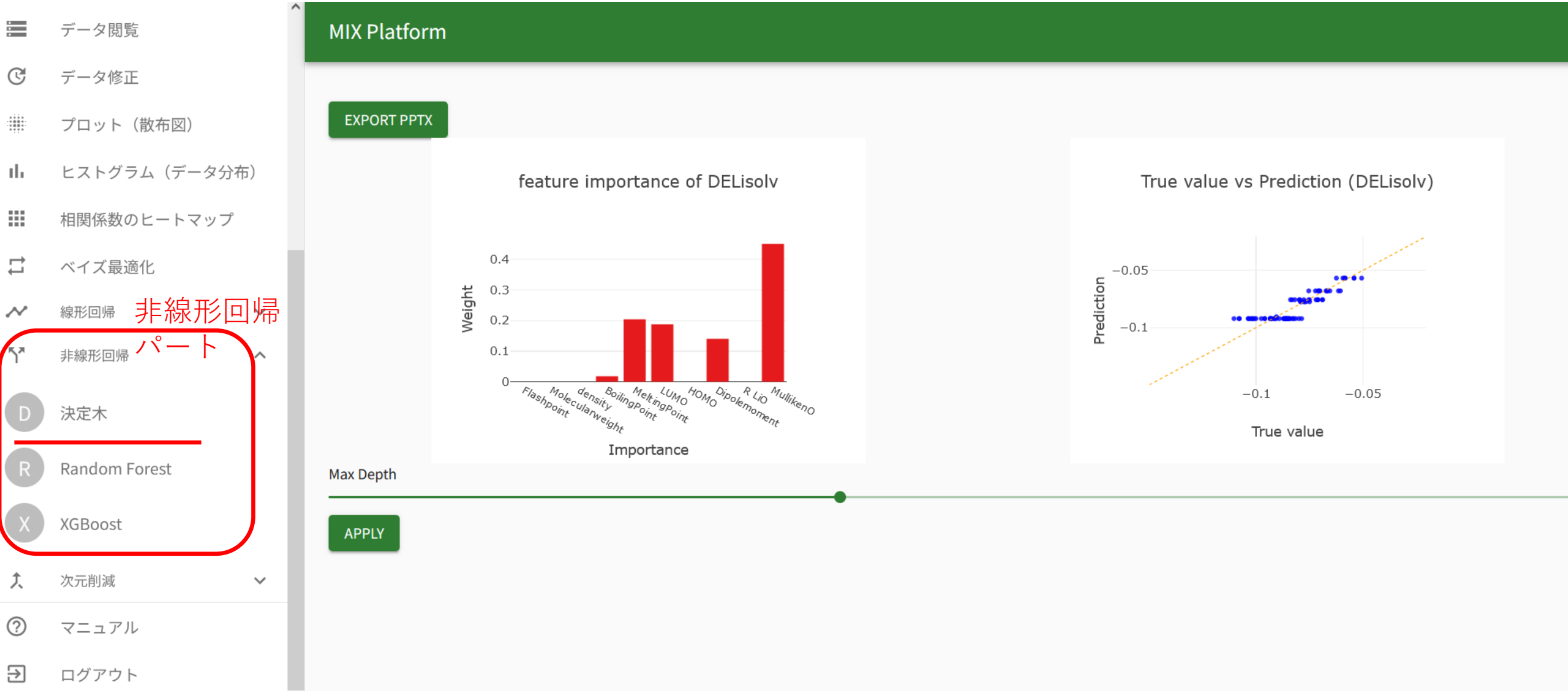
線形回帰パート

計算に数十秒かかります。
余り沢山は実行しないで
頂けると助かります。

全ての説明変数の組合せ（10変数あれば1023通り）に対して線形回帰を実施。予測精度の高いものから順にどの説明変数を使ったか表示。

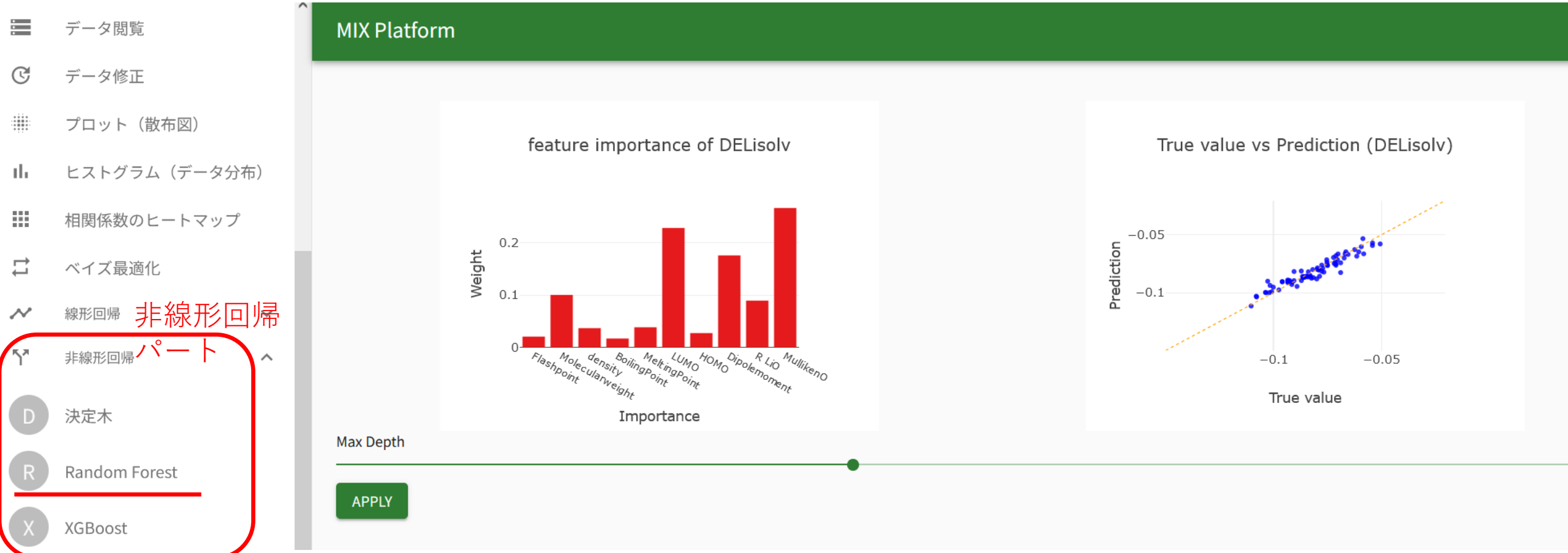
MIX Platformでできること（MI解析9）

非線形回帰分析(決定木)



MIX Platformでできること（MI解析10）

非線形回帰分析(Random Forest)



MIX Platformでできること（MI解析11）

非線形回帰分析(XGBoost)

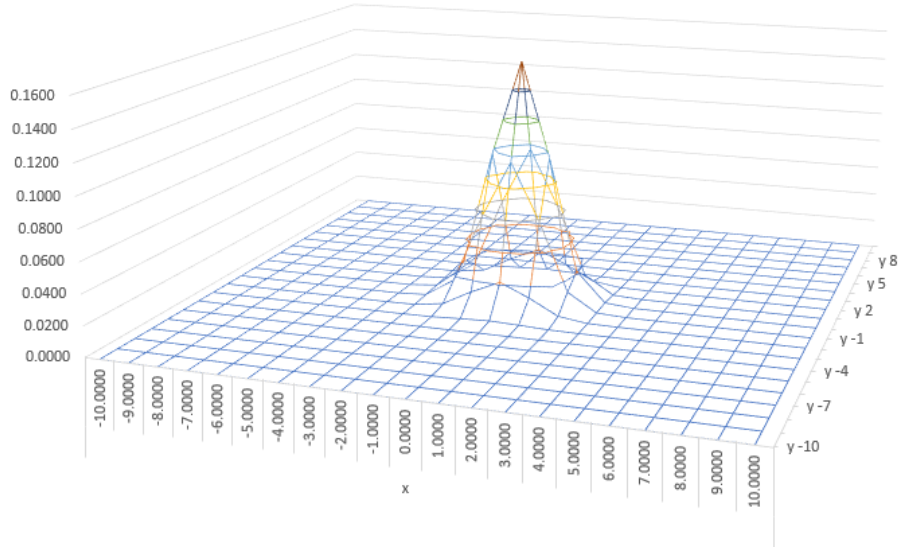
MIX Platformでできること（MI解析12）

ベイズ最適化

1．少数の実験値から、次に実験すべきものを提案
2．その提案値を元に実験
3．新規実験値も予測に利用して、再度次に実験すべきものを提案

このサイクルを満足する物性値が得られるまで続ける。

例：２次元空間における最大値を探す



		x																				
		-10.0000	-9.0000	-8.0000	-7.0000	-6.0000	-5.0000	-4.0000	-3.0000	-2.0000	-1.0000	0.0000	1.0000	2.0000	3.0000	4.0000	5.0000	6.0000	7.0000	8.0000	9.0000	10.0000
y	-10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0011	0.0018	0.0011	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0029	0.0131	0.0215	0.0131	0.0029	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0011	0.0131	0.0585	0.0965	0.0585	0.0131	0.0011	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0018	0.0215	0.0965	0.1592	0.0965	0.0215	0.0018	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0011	0.0131	0.0585	0.0965	0.0585	0.0131	0.0011	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0029	0.0131	0.0215	0.0131	0.0029	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0011	0.0018	0.0011	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

MIX Platformでできること（MI解析12）

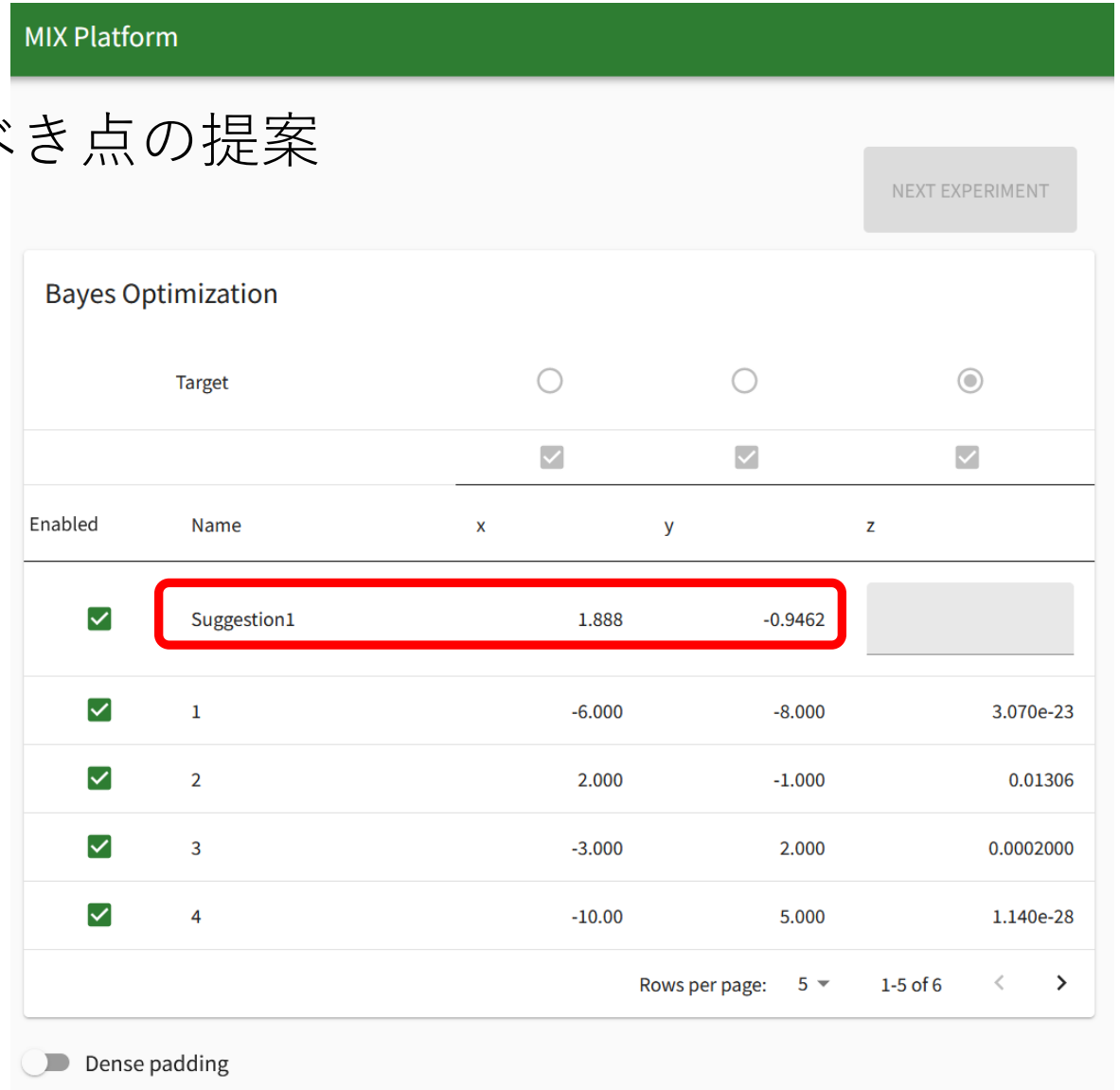
ベイズ最適化

例：２次元空間における最大値を探す

①. 初期値

x	y	z
-6	-8	3.07E-23
2	-1	0.013064
-3	2	0.0002
-10	5	1.14E-28
5	9	1.53E-24

②. 次の実験すべき点の提案



MIX Platformでできること（MI解析12）

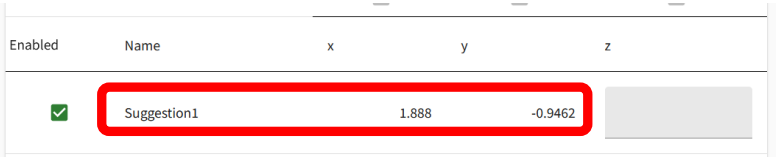
ベイズ最適化

例：２次元空間における最大値を探す

①. 初期値

x	y	z
-6	-8	3.07E-23
2	-1	0.013064
-3	2	0.0002
-10	5	1.14E-28
5	9	1.53E-24

②. 次の実験すべき点の提案



Enabled	Name	x	y	z
☑	Suggestion1	1.888	-0.9462	

③. 実験実施
（今回は実験の代わりにExcelで値を導出）

	A	B	C	D
1	試行回数	予測値x	予測値y	予測値z
2	0	-3	2	0.00023928
3	0	-10	5	1.14397E-28
4	0	5	9	1.52831E-24
5	0	-6	-8	3.0697E-23
6	0	2	-1	0.013064233
7	1	1.94	-0.9594	0.015300018
8	2			

MIX Platformでできること（MI解析12）

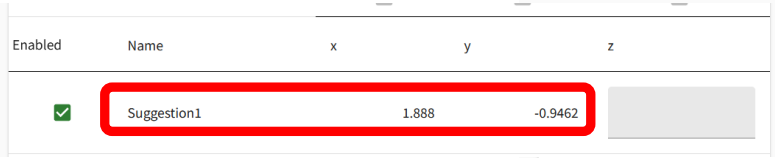
ベイズ最適化

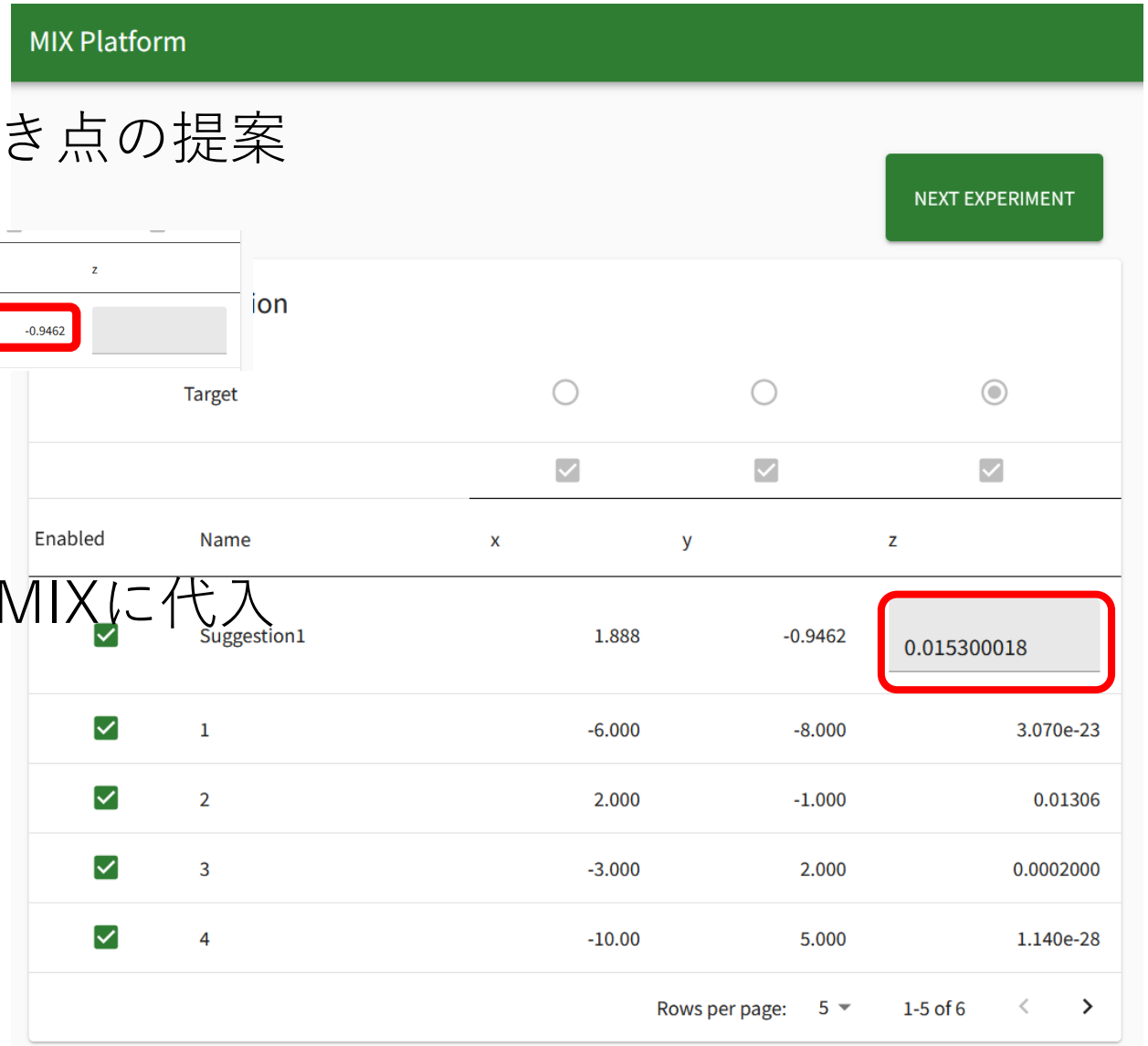
例：２次元空間における最大値を探す

①. 初期値

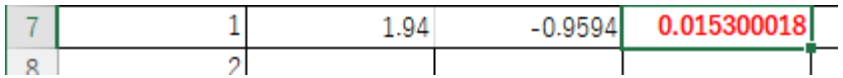
x	y	z
-6	-8	3.07E-23
2	-1	0.013064
-3	2	0.0002
-10	5	1.14E-28
5	9	1.53E-24

②. 次の実験すべき点の提案





③. 実験実施
（今回は実験の代わりにExcelで値を導出）



④. 実験値をMIXに代入

MIX Platformでできること（MI解析12）

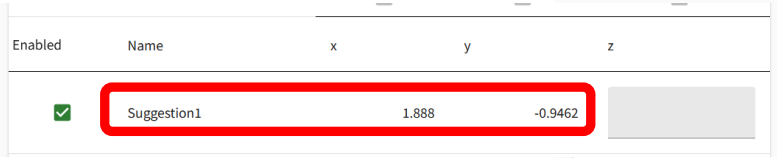
ベイズ最適化

例：２次元空間における最大値を探す

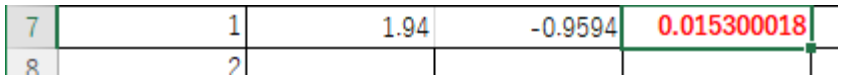
①. 初期値

x	y	z
-6	-8	3.07E-23
2	-1	0.013064
-3	2	0.0002
-10	5	1.14E-28
5	9	1.53E-24

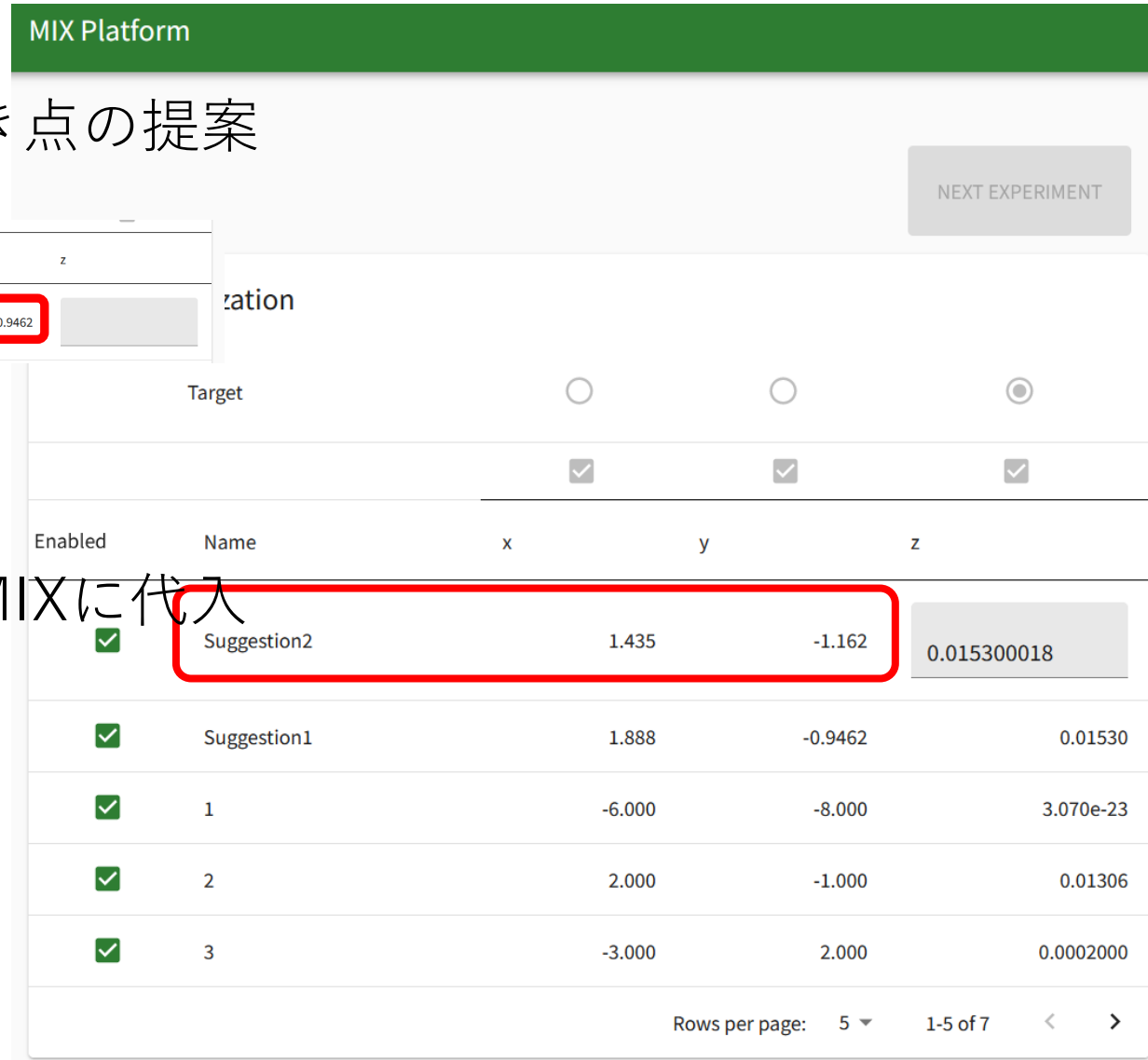
②. 次の実験すべき点の提案



③. 実験実施
（今回は実験の代わりにExcelで値を導出）



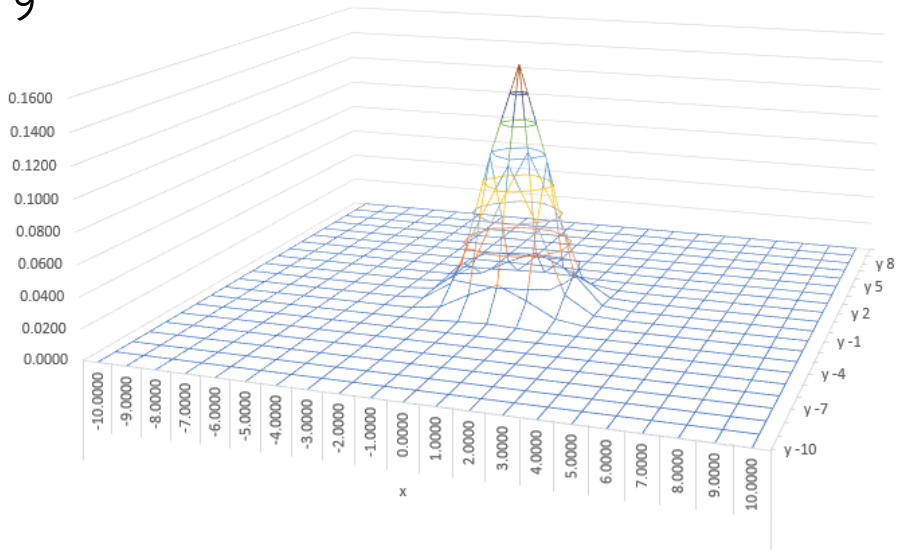
④. 実験値をMIXに代入
次の実験提案

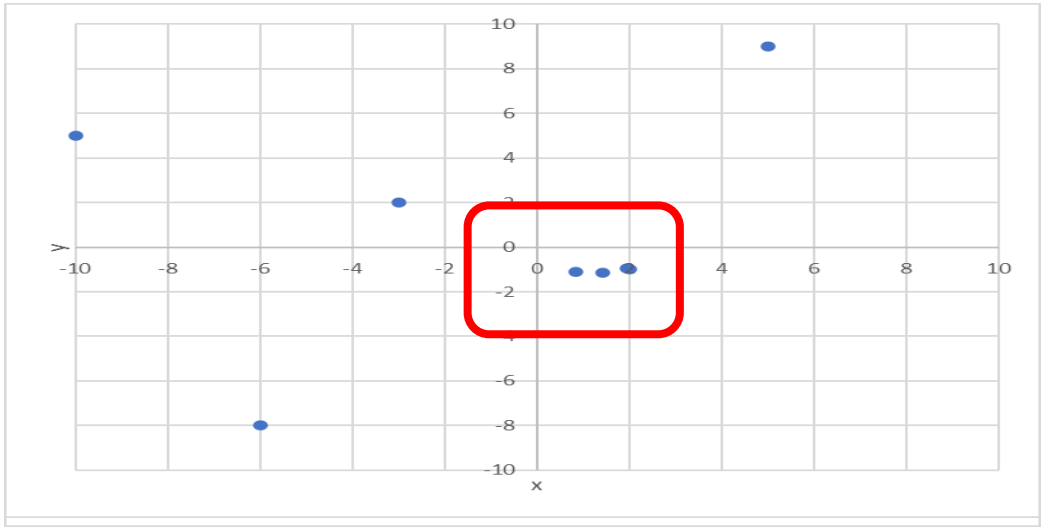


MIX Platformでできること（MI解析12）

ベイズ最適化

これを繰り返して大きな値をとるx, yの値を探す

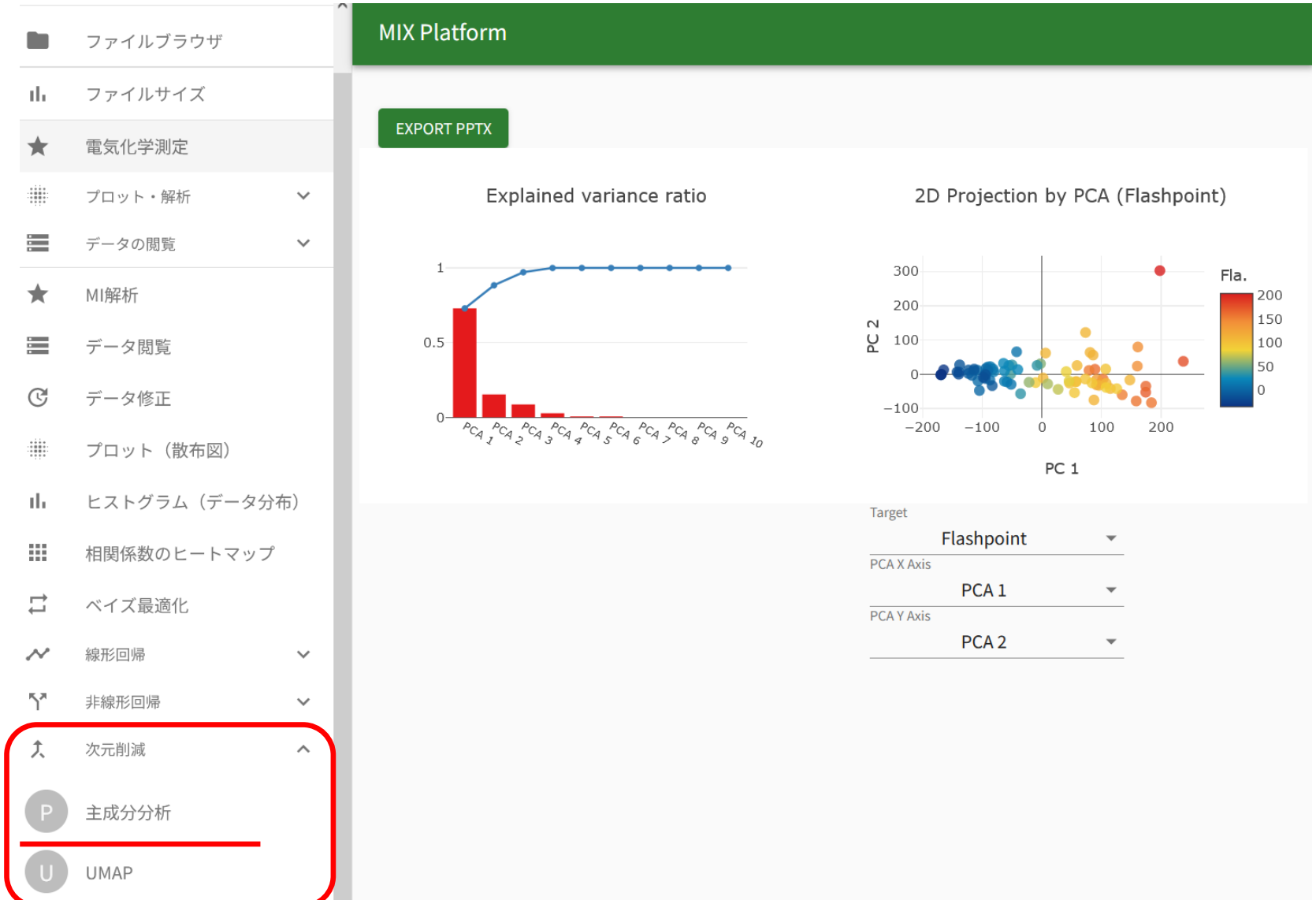




		x																				
		-10.0000	-9.0000	-8.0000	-7.0000	-6.0000	-5.0000	-4.0000	-3.0000	-2.0000	-1.0000	0.0000	1.0000	2.0000	3.0000	4.0000	5.0000	6.0000	7.0000	8.0000	9.0000	10.0000
y	-10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0011	0.0018	0.0011	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0029	0.0131	0.0215	0.0131	0.0029	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	-1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0011	0.0131	0.0585	0.0965	0.0585	0.0131	0.0011	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	0	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0018	0.0215	0.0965	0.1592	0.0965	0.0215	0.0018	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	1	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0011	0.0131	0.0585	0.0965	0.0585	0.0131	0.0011	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	2	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0029	0.0131	0.0215	0.0131	0.0029	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	3	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0002	0.0011	0.0018	0.0011	0.0002	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	4	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0001	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	5	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	6	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	7	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	8	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	9	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000
	10	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000	0.0000

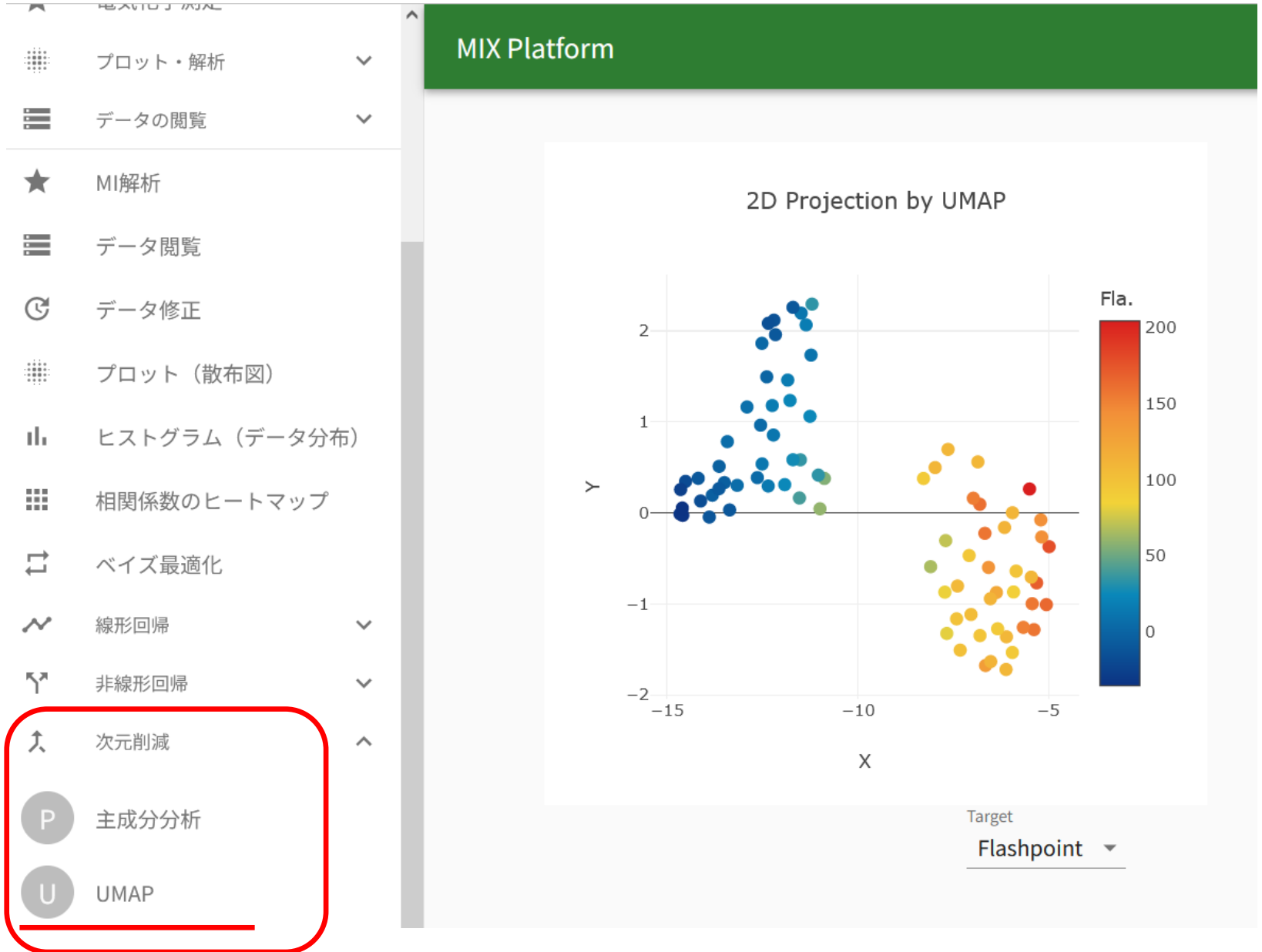
MIX Platformでできること（MI解析13）

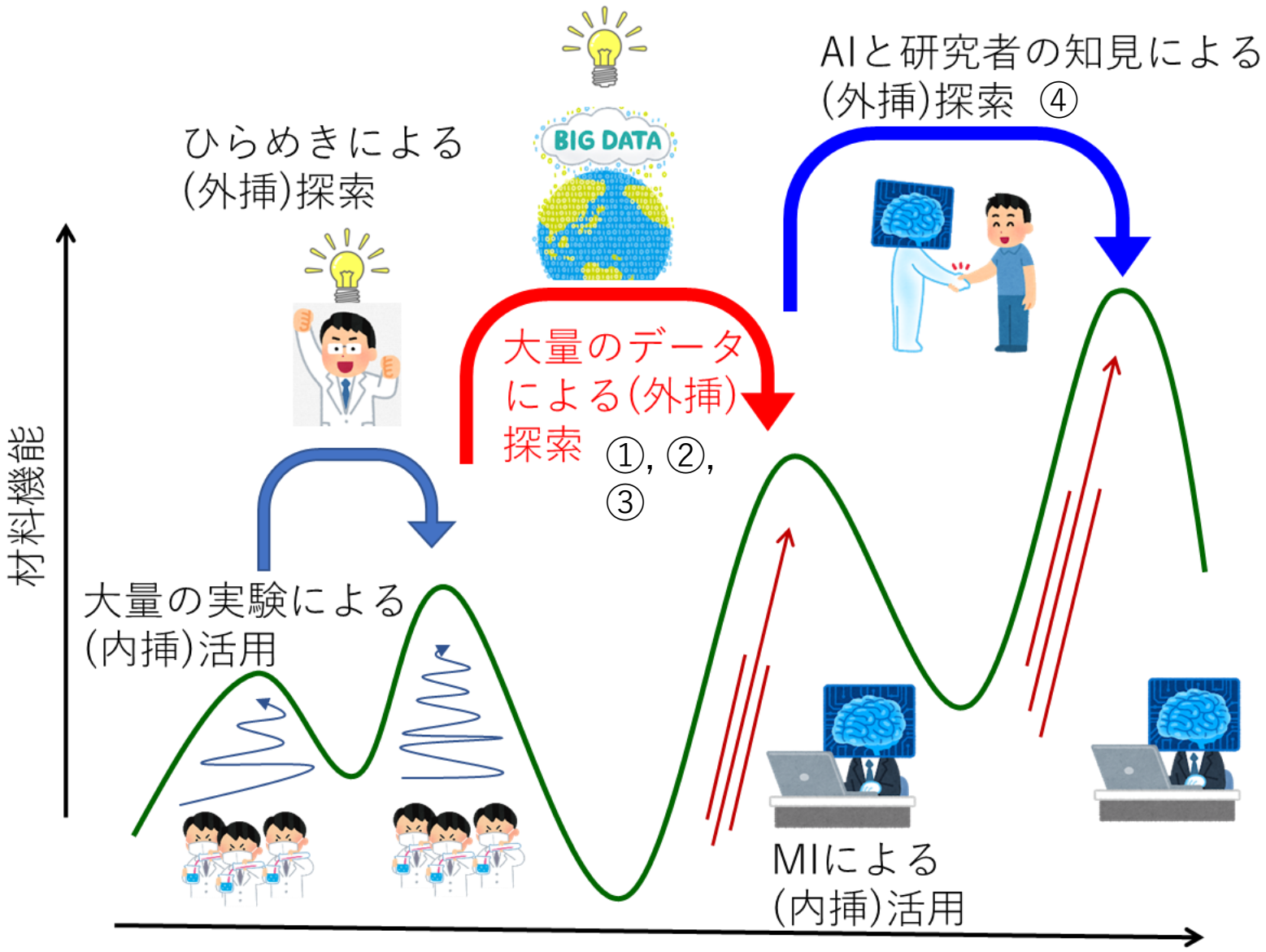
主成分分析



MIX Platformでできること（MI解析14）

UMAP





①PEFCデータベース
②自律自動実験データ
③自然言語処理による論文データ
④PEFCに関するMI解析ノウハウと解析手法開発

⇒新規材料の設計指針提案

自律・自動実験について：NIMSの自律自動実験装置の紹介

国立研究開発法人物質材料研究機構（NIMS）
マテリアル基盤センター

冨中 悟史、中田 謙吾

過去の講演内容

2022年　燃料電池材料の自律実験システムの設計

- ✓MIグループとしての自律実験
- ✓予測器のためのデータ処理について
- ✓自律実験のシステム設計の解説

2023年　燃料電池材料の自律実験システムの紹介

- ✓背景
- ✓NIMSのシステム（ELVES）の紹介
- ✓現状紹介（課題共有）

本日の講演内容

「燃料電池材料の自律実験システムの紹介」

✔全体のおさらい

✔動作の紹介

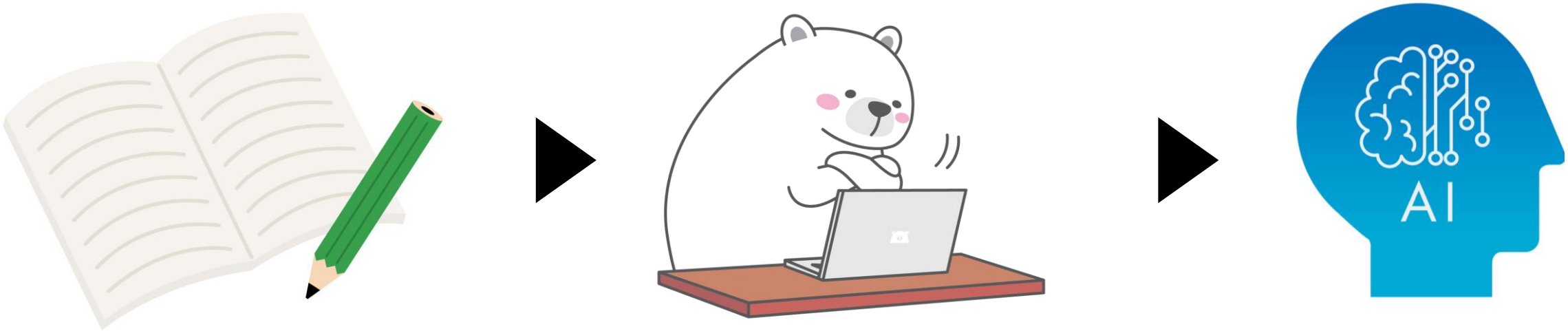
MIグループとしての自律実験の意義

- 均質かつ大量のデータ（データベース強化）
  →MIによる材料探索が加速
- 人が記録しきれないプロセス条件などもデータ化
- 並列化が容易
- **属人性を排除**（**特に燃料電池触媒では大きい**）

国全体の材料探索の効率化・迅速化

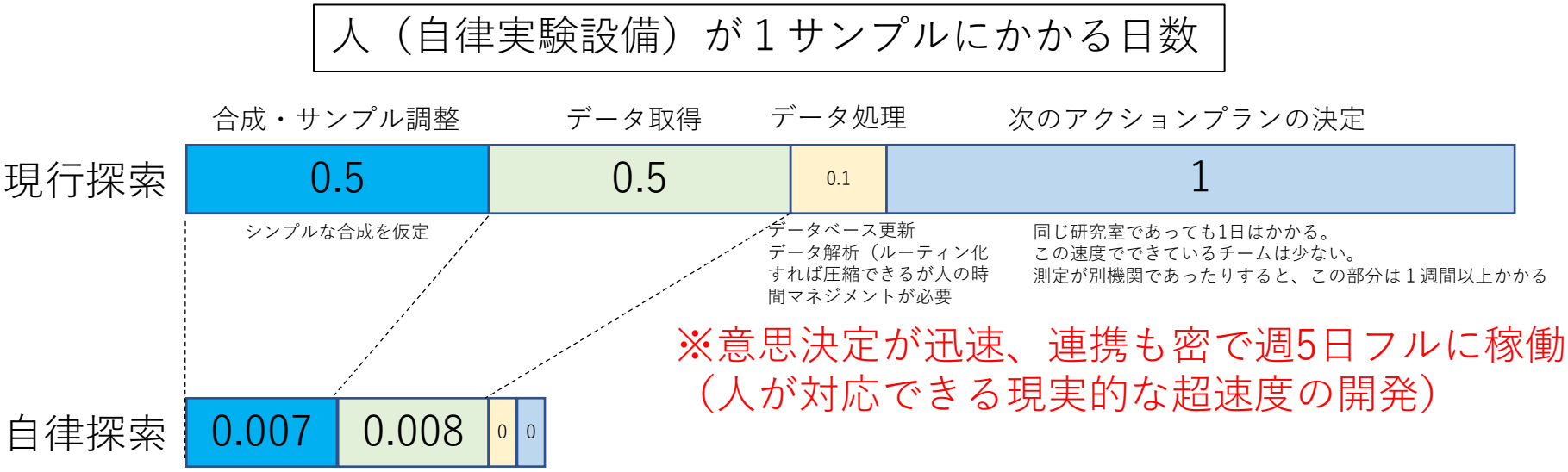
研究開発のDXの必要性

- ✓ 社会的要求に基づく研究開発は競争を伴う
- ✓ 既存の研究開発のやり方を変革することで**加速**
- ✓ 基礎研究≠役に立たないもの（DXで加速すべき研究とは何かを明確に）



燃料電池事業全体への加速効果

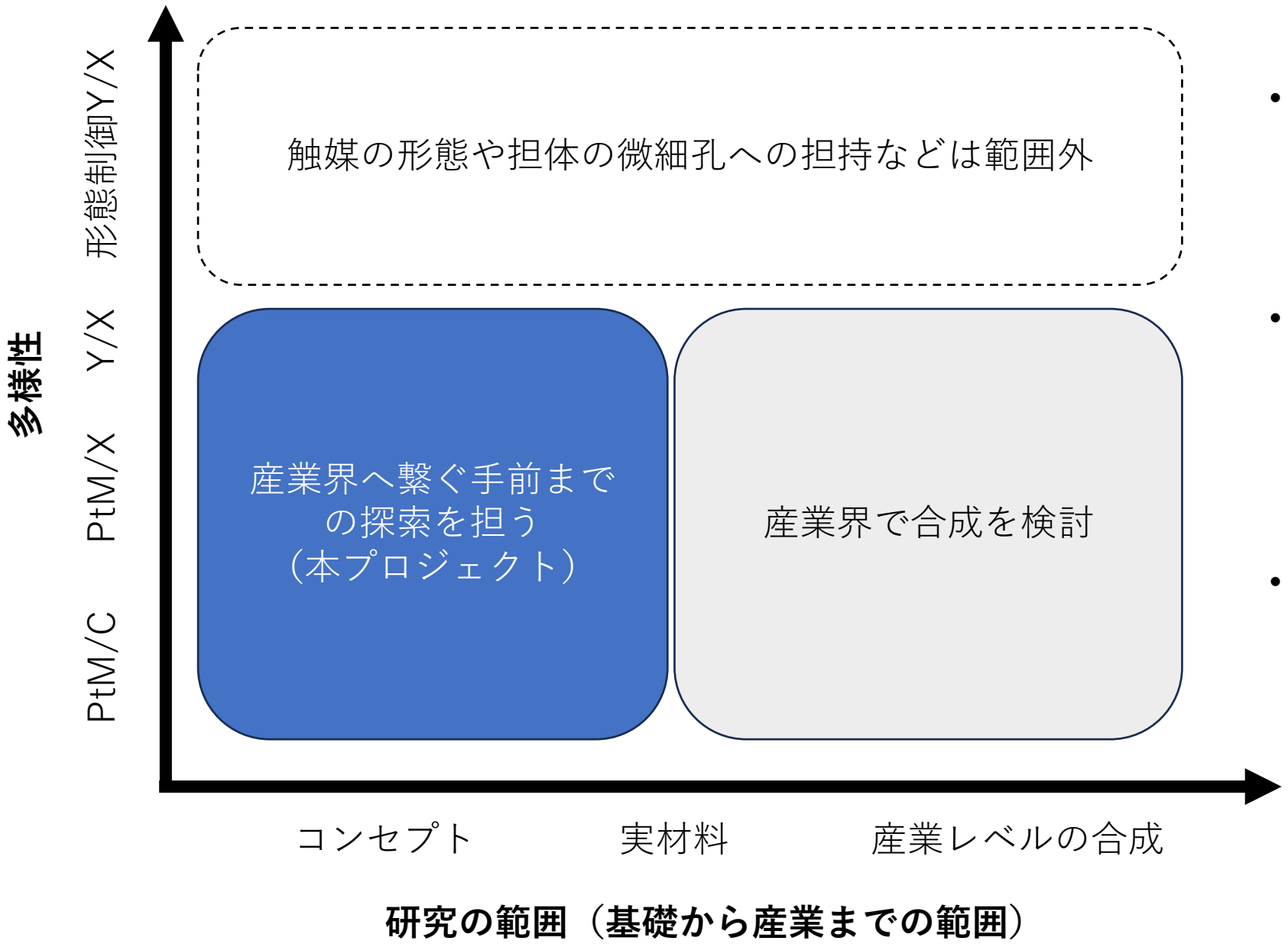
MI実施の視点でのメリット（前頁）に加え、材料探索スピードが少なく見積もって100倍加速の効果がある



自律探索は週７稼働とすると最低でも200倍は加速

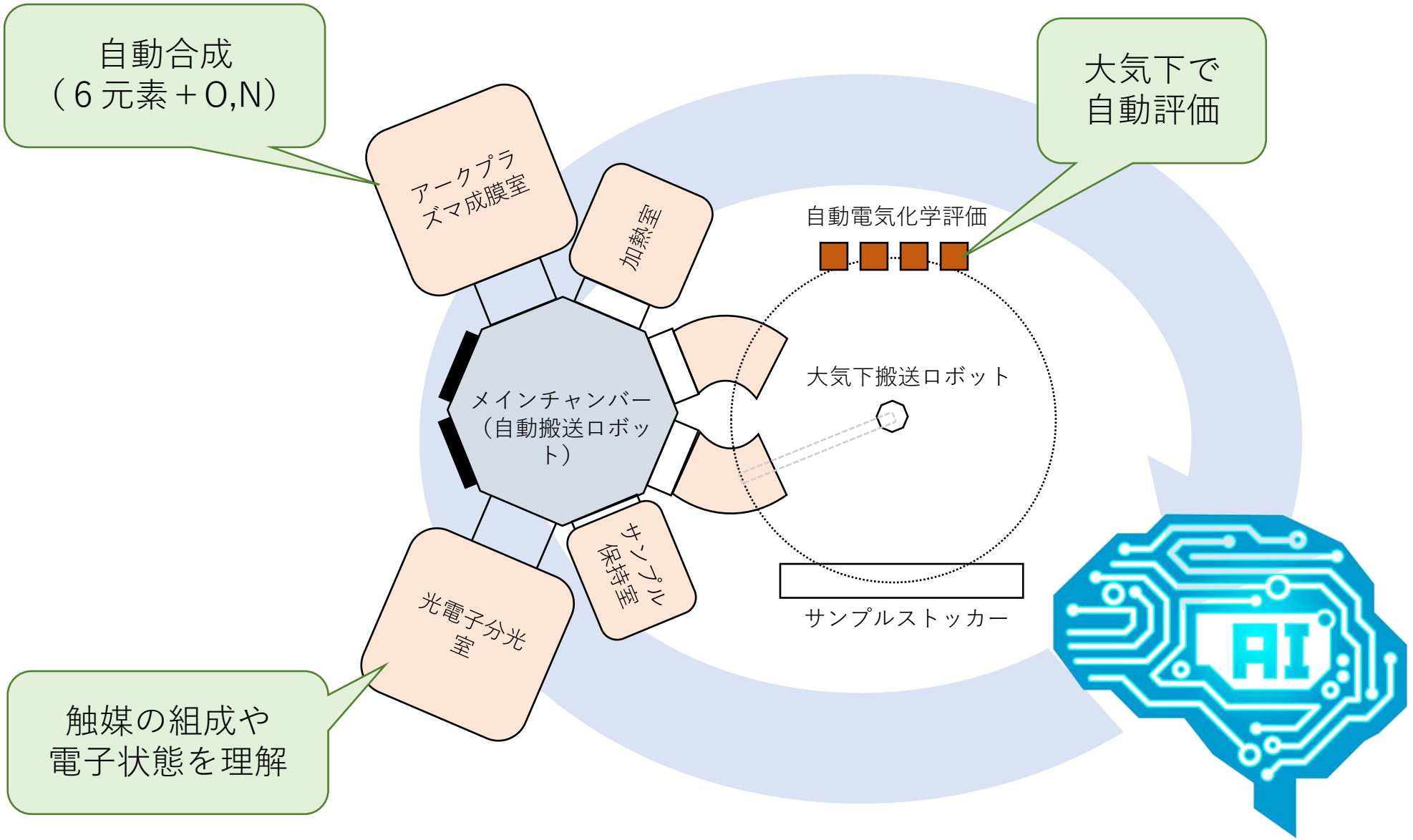
これは打席に立つ回数みたいなもので、AIによる次の１手の予測精度と人間の予測精度の打率は含めていない

加速アイテムの位置づけ



- 担体にナノ粒子を担持するという現行触媒材料のアーキテクチャの範囲で実材料となりえる手前までを高速で探索することを目標に設計。
- ナノ粒子として種々の合金を合成し、実作動に近い特性評価と光電子分光（化学状態や価電子情報の収集）により特性データと科学的な解析を行う。
- 産業界側での合成の最適化からは別のDX技術が必要だが、その手前までは広い範囲をカバー。

構想図：膨大な数のパラメータ空間を自動探索



ELVES：ロボットと人工知能の集合

ELVES
開発プロジェクト

「無人で研究を進める小人たち」

- 制御はPythonで自動化
- APIが無いものはNIMSで作製し公開
- データベース化（データプラットフォーム）
- 高速解析しMI解析（NIMSの専門家の知）
- AIをNIMSで構築し育成

統一規格品でなくても、インターフェースを作って連携させる。

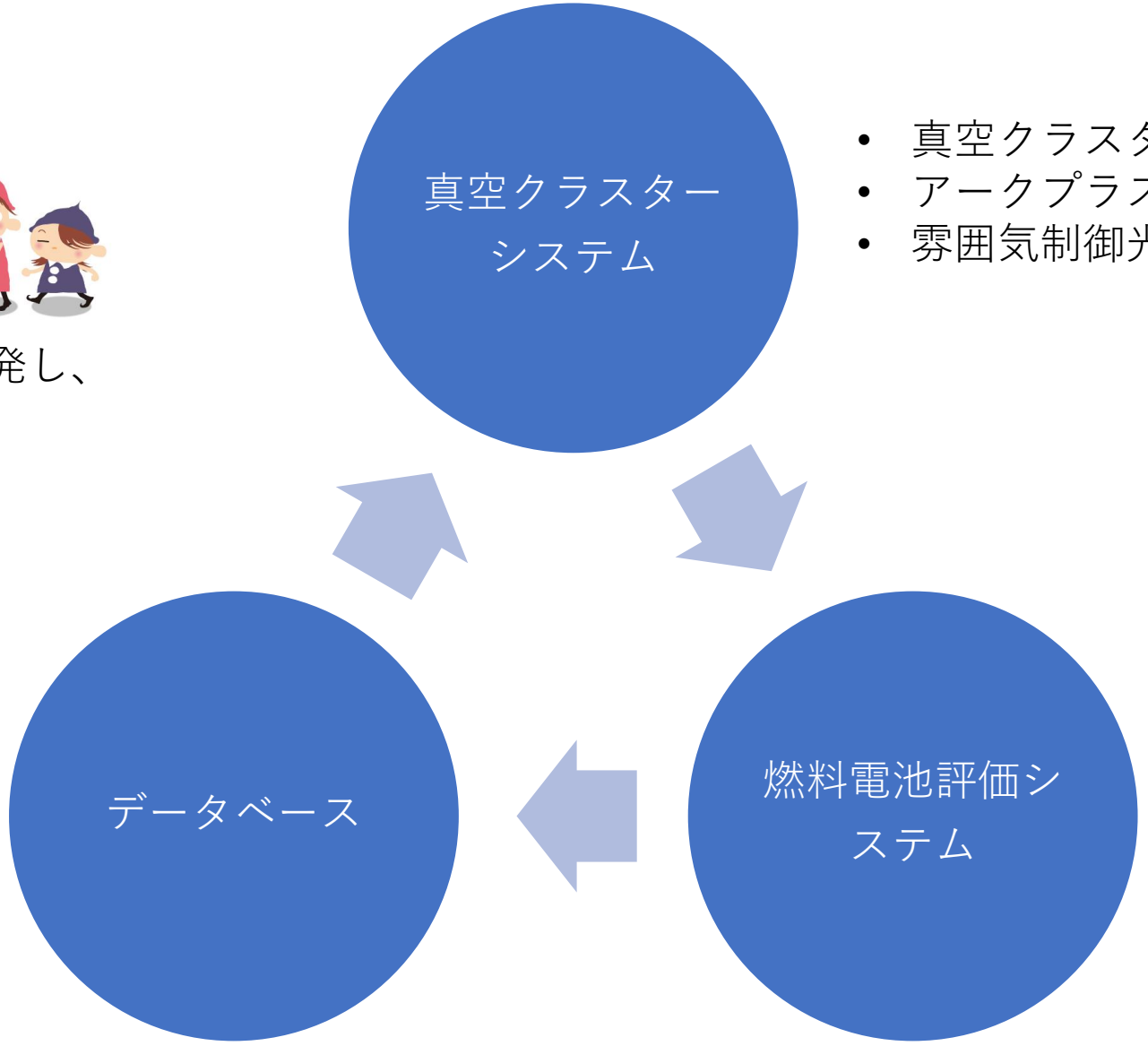
目的：研究のボトルネックを解消するために、実験の並列処理と異なる機器の直列化を実現する開発。

ELVESの主な要素

各要素のプログラムを開発し、
柔軟に高速処理を行う

GitLab

- サーバー

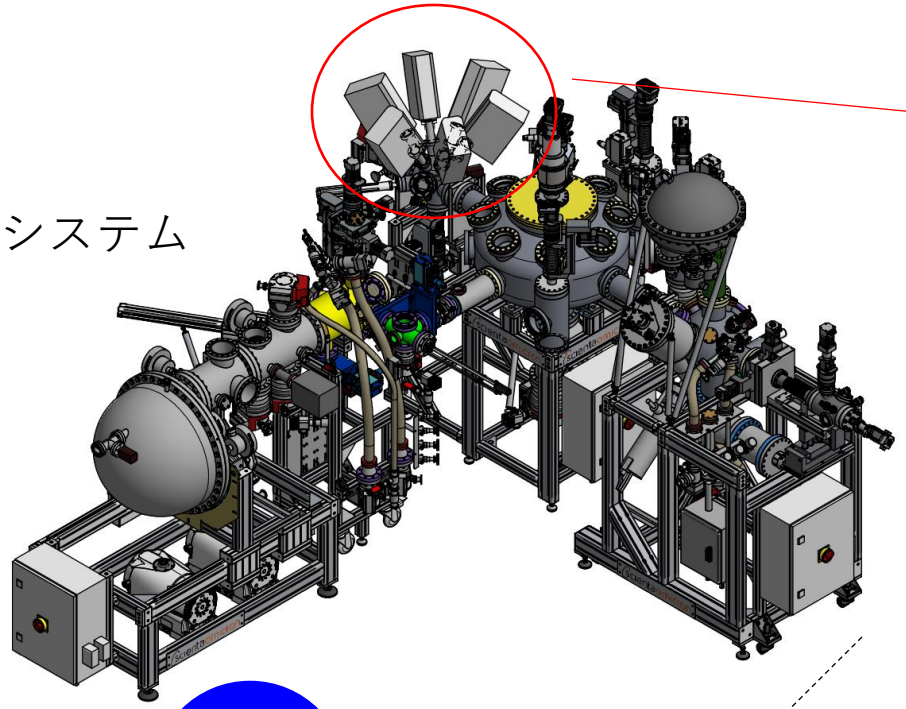


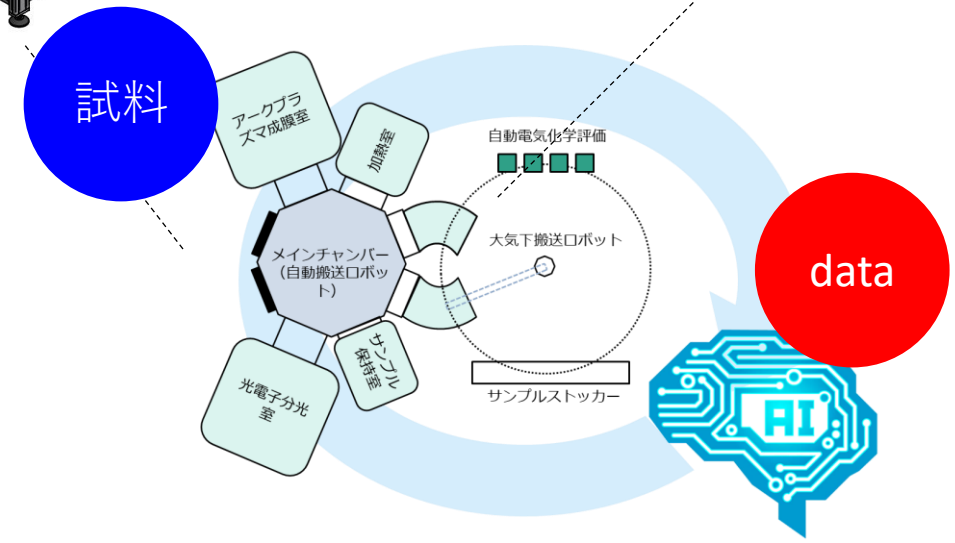
- 真空クラスターシステム
- アークプラズマ装置
- 雰囲気制御光電子分光装置

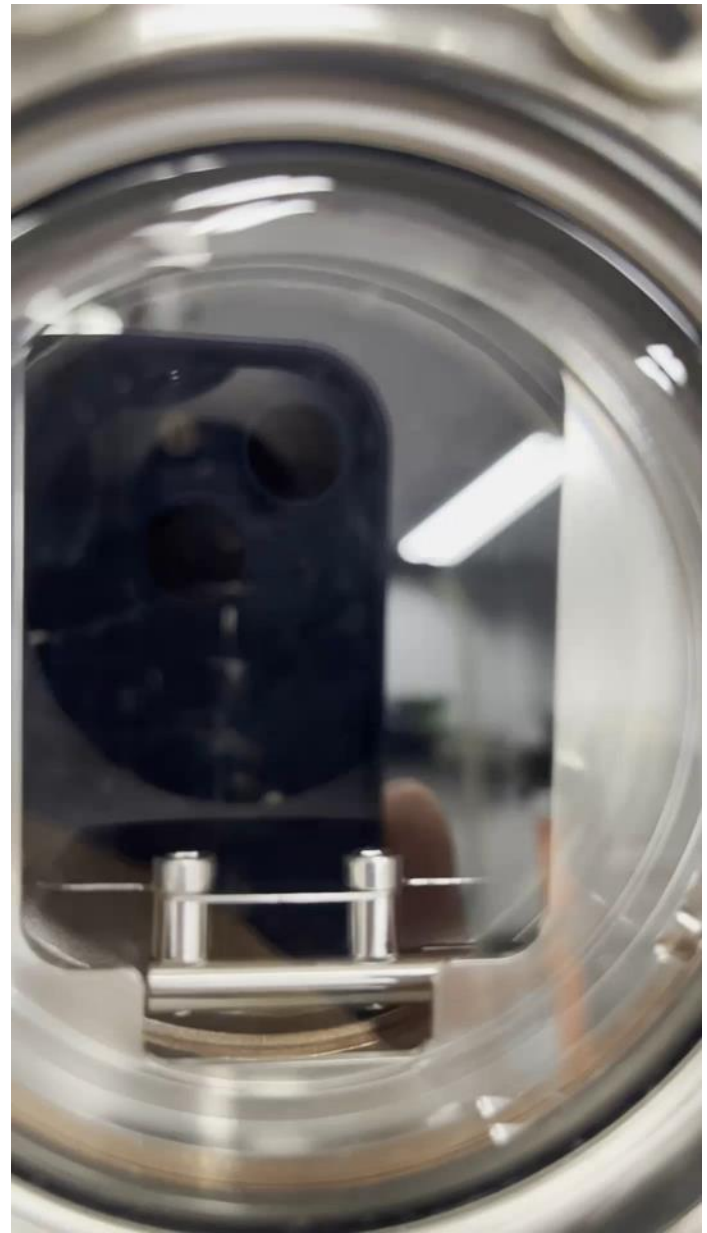
- 電気化学測定システム
- セル
- サーボプレス
- ガス制御
- 搬送ロボット

①アークプラズマによる白金の堆積

真空システム

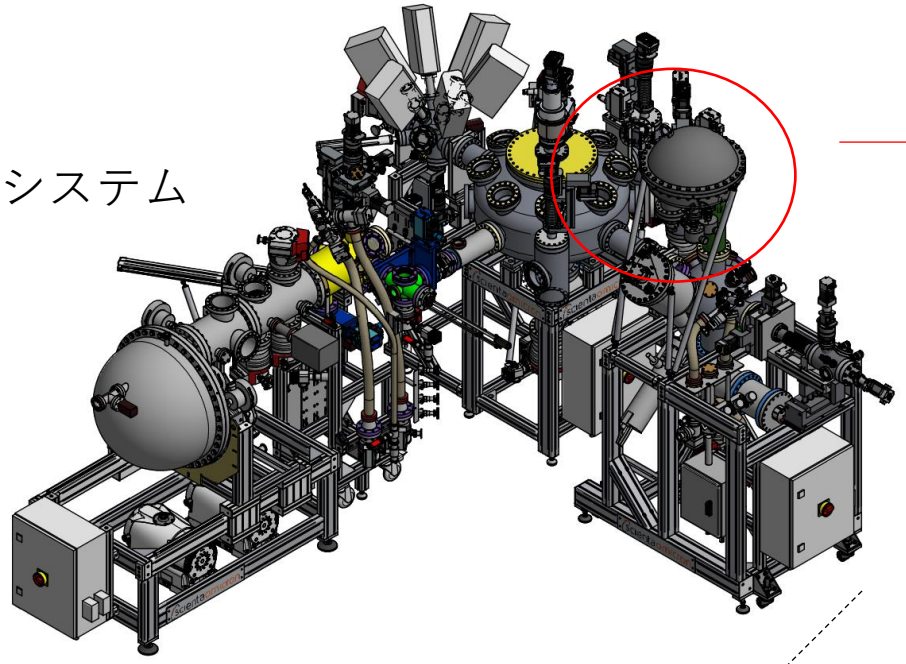


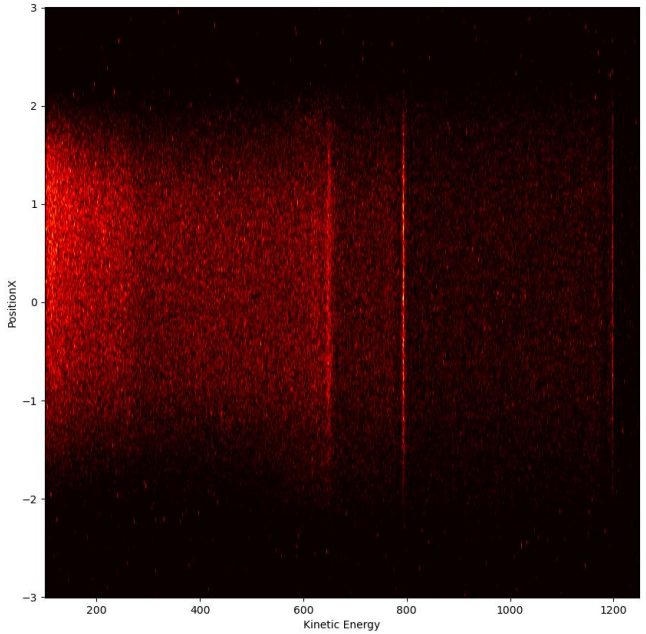




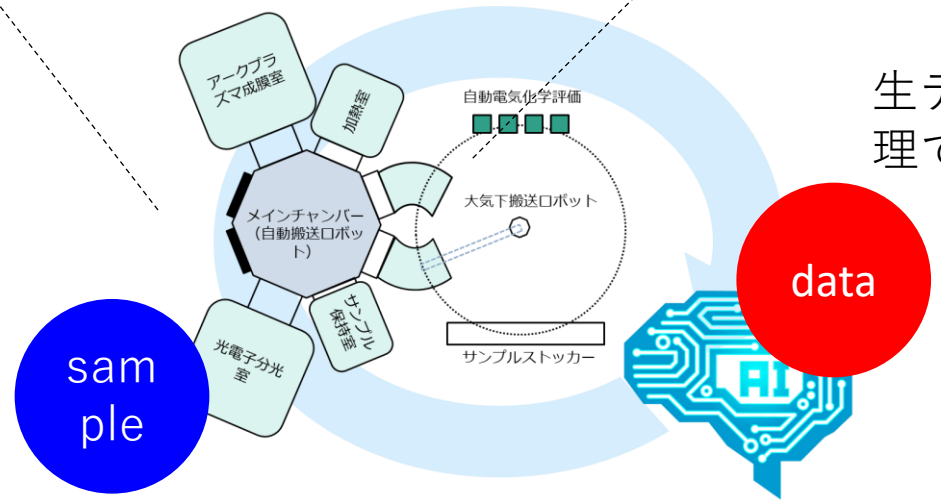
②光電子分光測定による分析

真空システム





Pt XPS 2d image



生データを記録しながら、高速処理で必要な情報を取り出す

```python
# coding: utf-8
#!/usr/bin/python

##################################################################
#   Description: XPS data extractor for ELVES
#   Version: 1.0.0
#   Last update: 2023.3.22
#   Developer: Satoshi Tominaka
#   Affiliation: National Institute for Materials Science
#   Contact: TOMINAKA.Satoshi@nims.go.jp
##################################################################
from __future__ import annotations
import copy, glob, os, re
#import numpy as np
#import arpes
#from arpes.io import load_data

#import peak
import h5py
from matplotlib import pyplot as plt


file_name = "h5_files/0231 test Pt_1Hz_100.h5"
data_folder = os.path.dirname(os.path.abspath(__file__))
print(data_folder)

def find_files():
    files = glob.glob(data_folder+"/"+file_name, recursive=True)
    return files

class XPS:
    def __init__(self):
        self.data_file = None #hdf5

    def read(self):
        self.file = h5py.File(self.data_file,"r")
```

ELVES Project

現状

- Pt触媒、PtCo触媒での試験中
- 日々、課題を１つずつ解決
- 類似の検討を進める方に向けた課題の取りまとめ

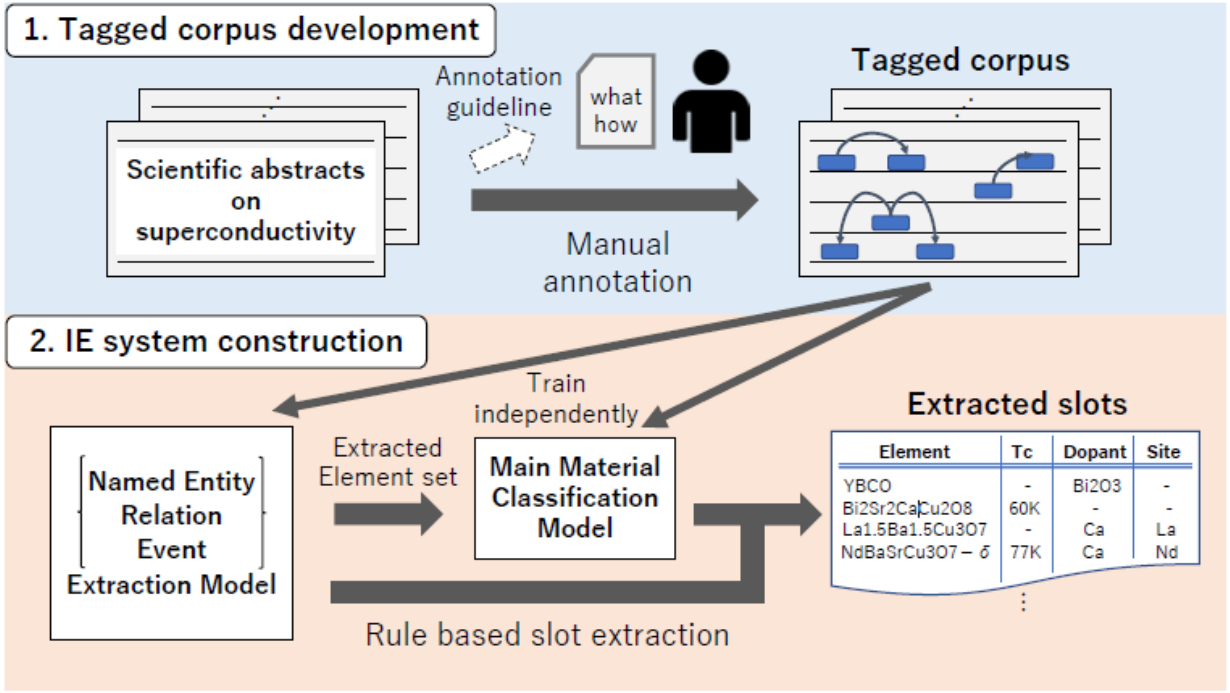
ありがとうございました。

謝辞：
この検討は、ＮＥＤＯの「マテリアルズ・インフォマティクスによる燃料電池材料の研究開発」として実施して得られたものです。

文献情報抽出技術による持続的なデータベース構築

SC-CoMIcs: A Superconductivity Corpus for Materials Informatics

Kyosuke Yamaguchi[1], Ryoji Asahi[2], Yutaka Sasaki[1]
[1]Toyota Technological Institute
2-12-1 Hisakata, Tenpaku-ku, Nagoya, 468-8511 Japan
[2]Toyota Central R&D Labs., Inc., 41-1, Yokomichi, Nagakute, Aichi 480-1192, Japan



文献、特許から組成ープロセスー物性をデータベース化

Yamaguchi, Asahi, Sasaki, Adv. Eng. Inf. 54, 101768 (2022)

Example of the corpus construction

Document collection → Annotation design → Manual annotation

ScienceDirect
Web APIs

1,000 abstracts
on superconductivity

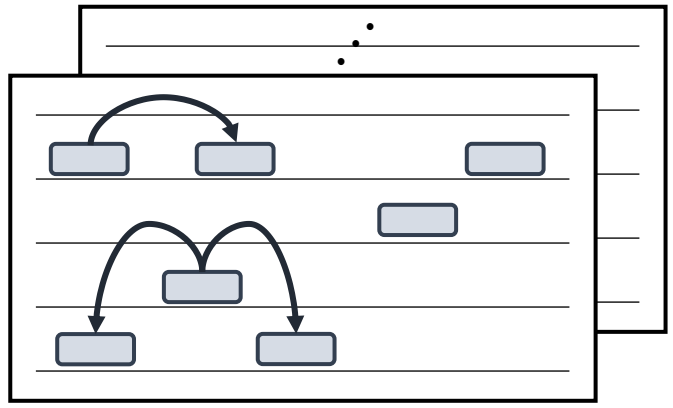
Named Entity
Characterization,
Process, Property,
Material, Element,
Value, SC

Main Material

Event
Doping
Dopant
Site

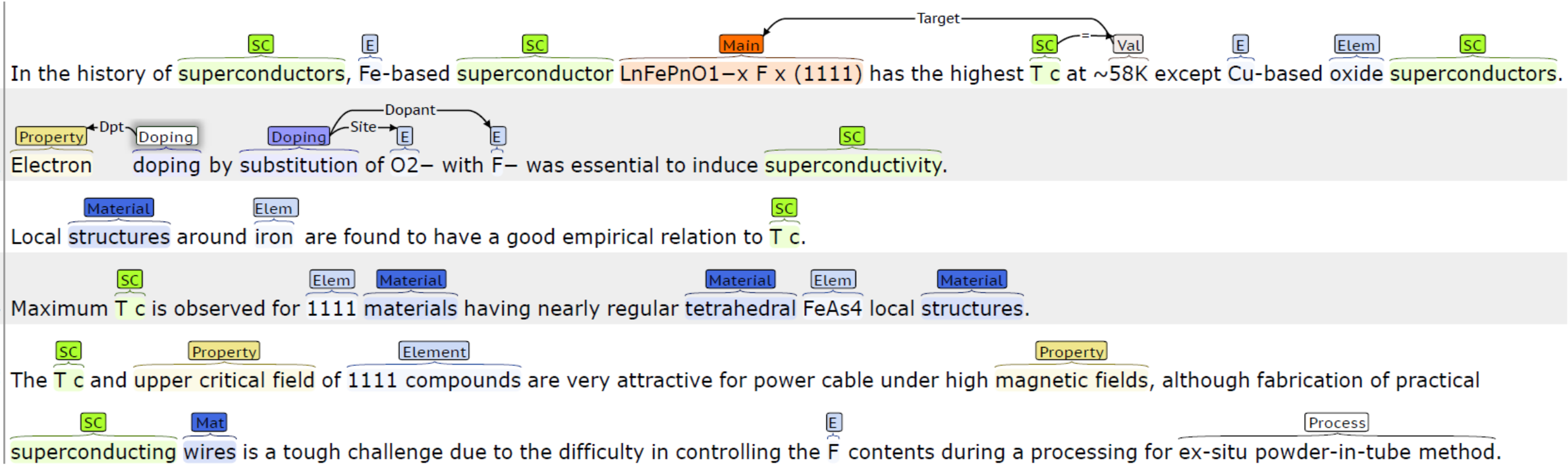
Relation
Equivalent
Target
Condition

Two workers



Tagged corpus

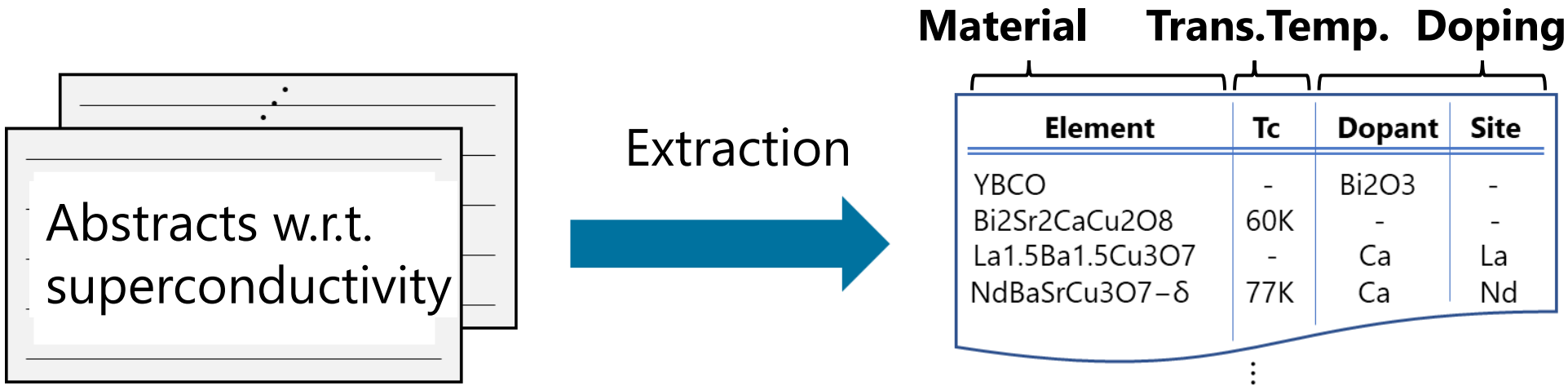
Annotation examples



brat rapid annotation tool: https://brat.nlplab.org/index.html

Extractions of superconducting materials and Tc

✓Construct a new corpus, SC-CoMIcs, for superconducting material information extraction and apply deep learning to confirm the usefulness of the corpus in the superconducting materials search.



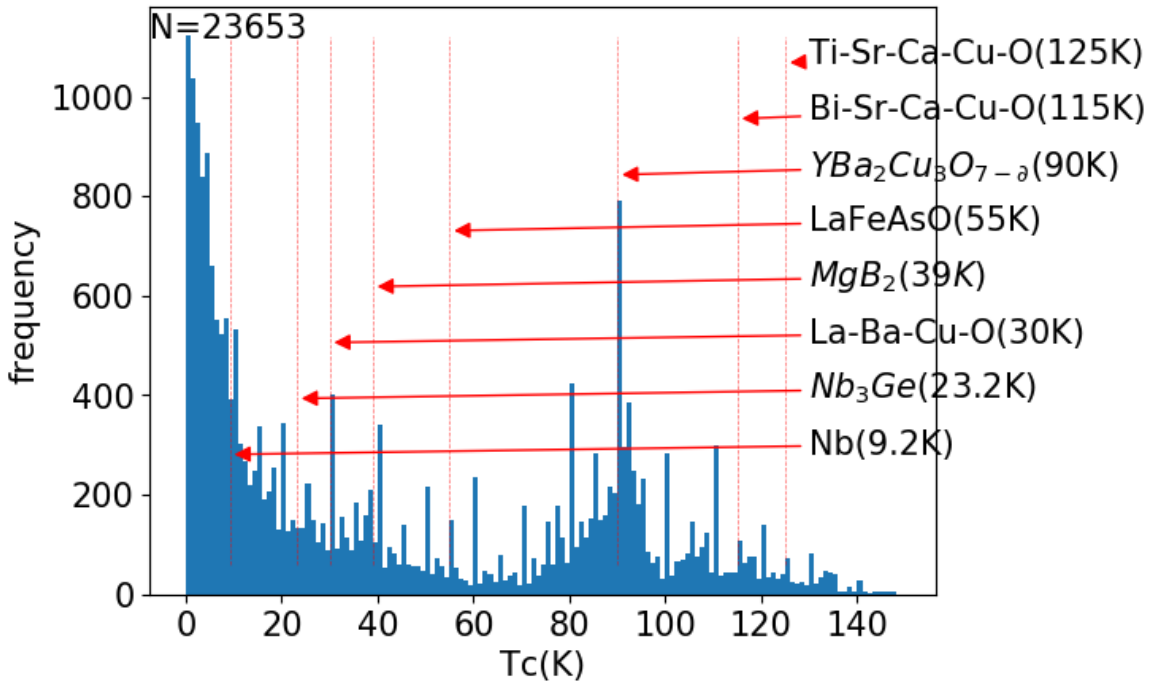
Extraction of the set of main material, phase transition temperature, and doping information

➡ **Prediction of characteristics of candidate materials**

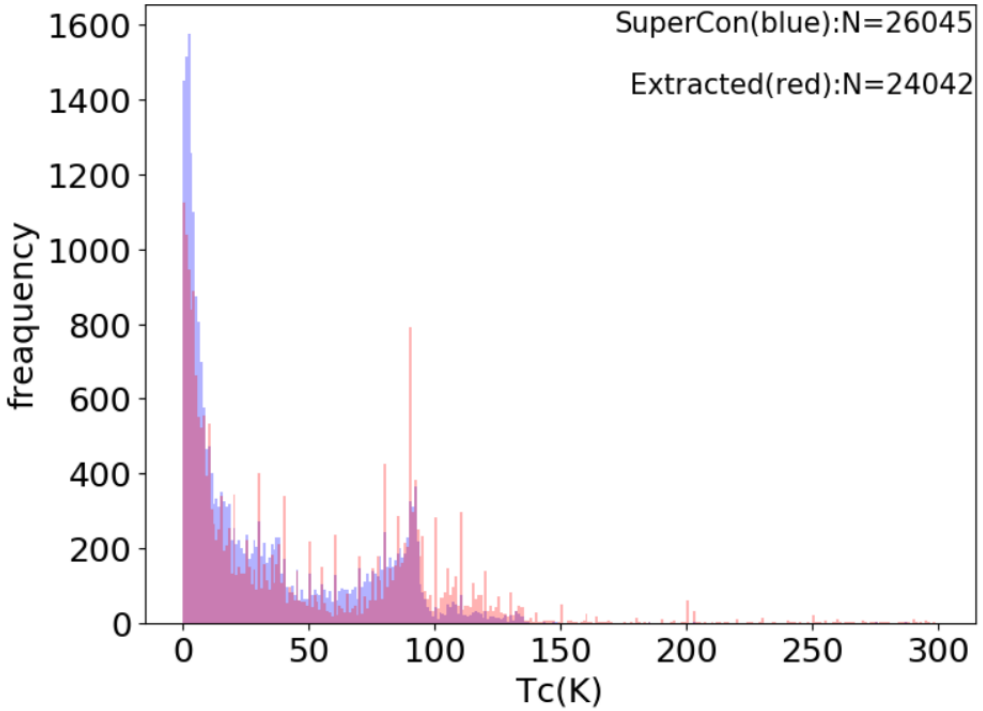
(Semi-)automatics population of meterials databases

Extraction of superconducting transition temperature (Tc)

extracted data



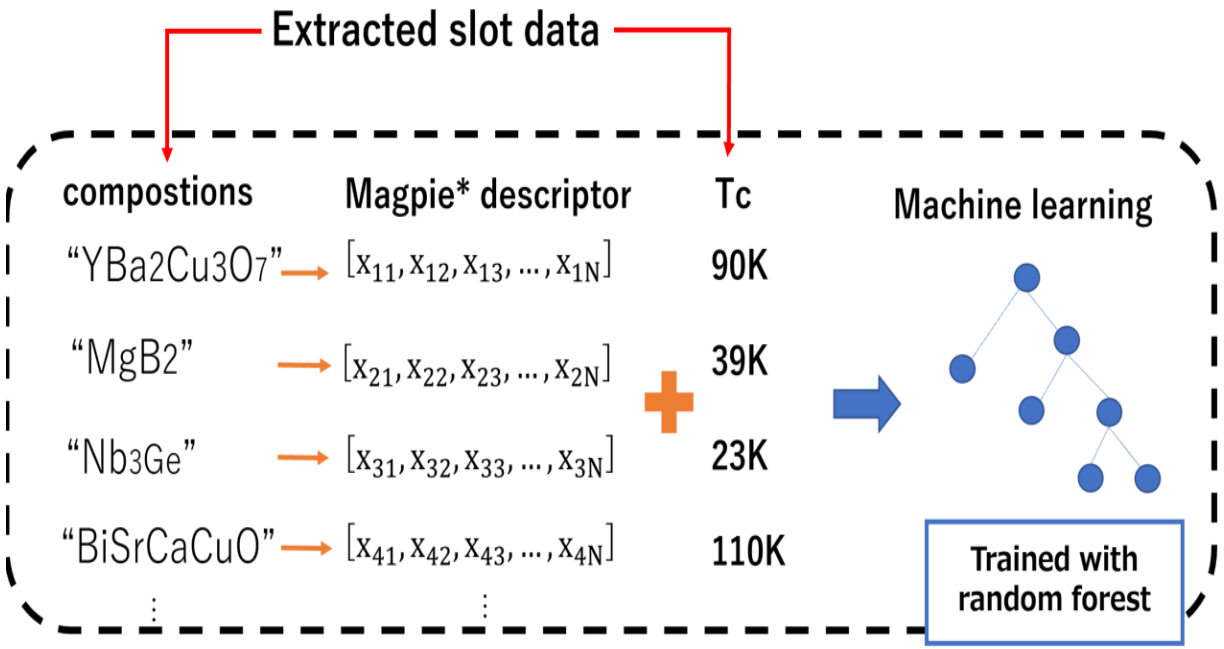
comparison with existing database (Supercon*)

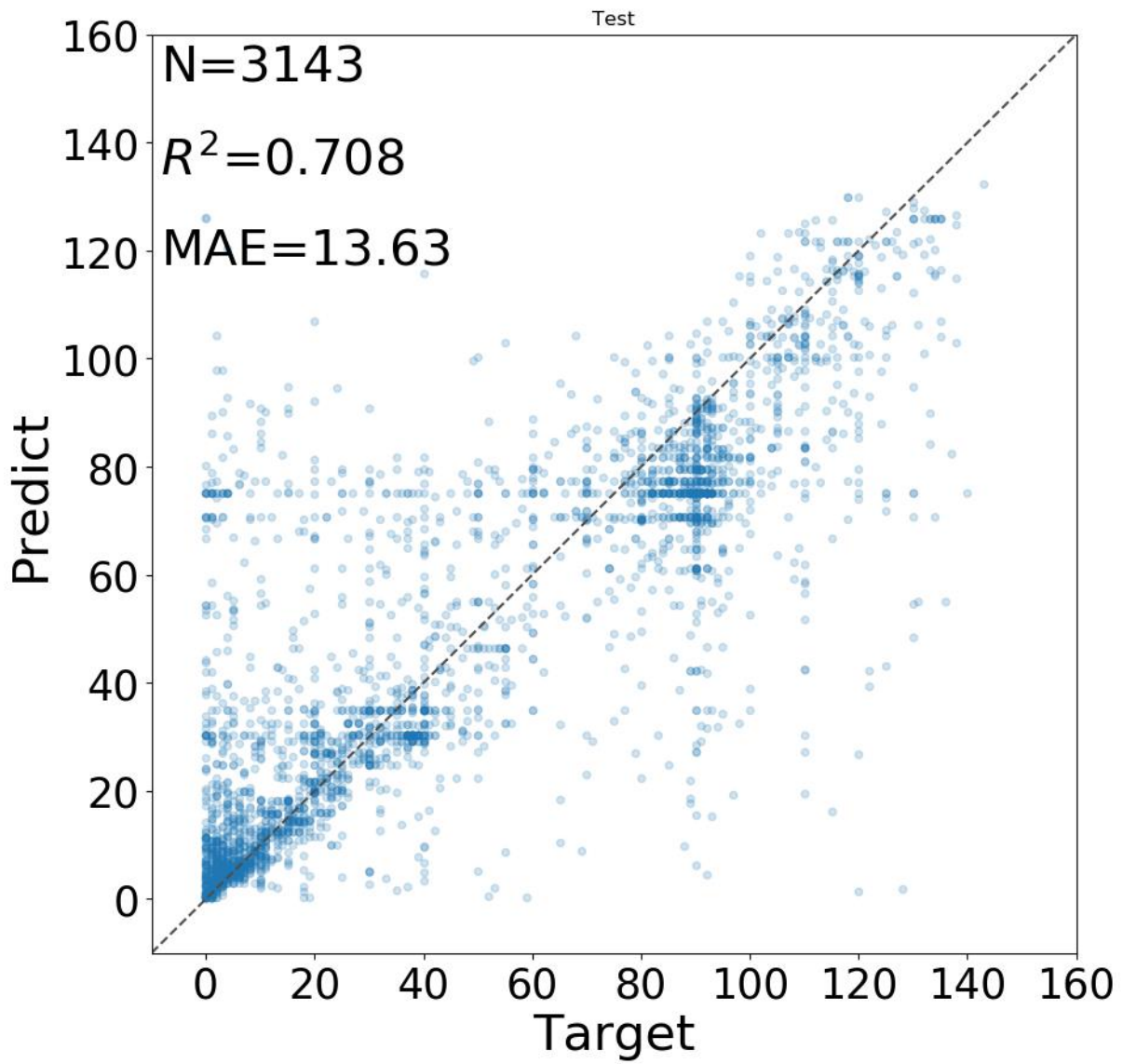


・Each peak corresponds to representative superconducting materials
・Automatically extracted data is consistent with the existing database Supercon* that has been manually constructed.

*https://mdr.nims.go.jp/collections/5712mb227

Performance of the Tc prediction model

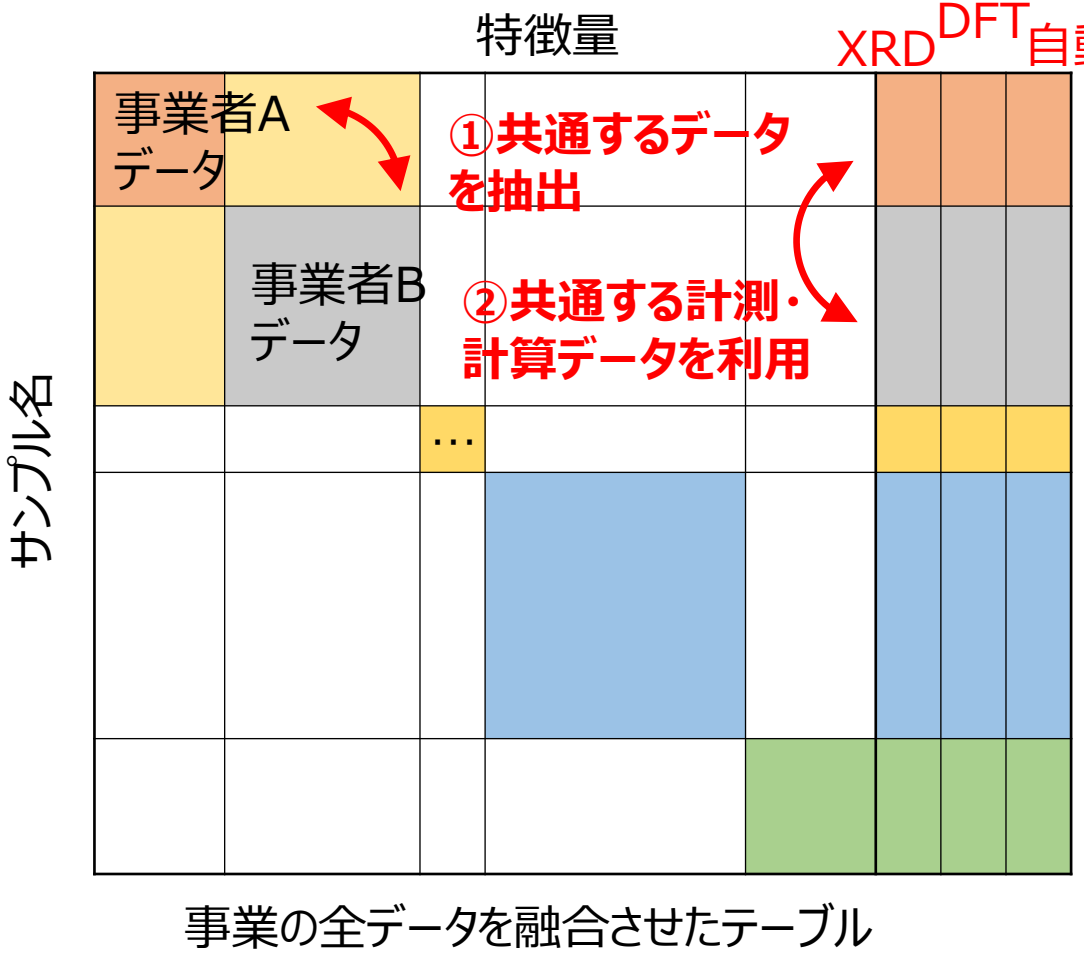




・**The model can predict Tc within an error of 14 K only using automatically extracted data.**

データ連携と検索

異なる研究者、分野のデータ連携→合体した大きな表を作成



自然言語処理による論文データ…

1. 本事業の全てのデータを融合させ、一枚のテーブルを作る。

2. 個別データ間の関連性を考慮していないため、非対角成分にデータが無く、データ解析ができない。

3. 非対角成分にデータを与えるため、
①個別データ間で共通するデータを抽出する。（データの共通リスト化）
②個別データ間で共通する計測手法や計算データを利用する。

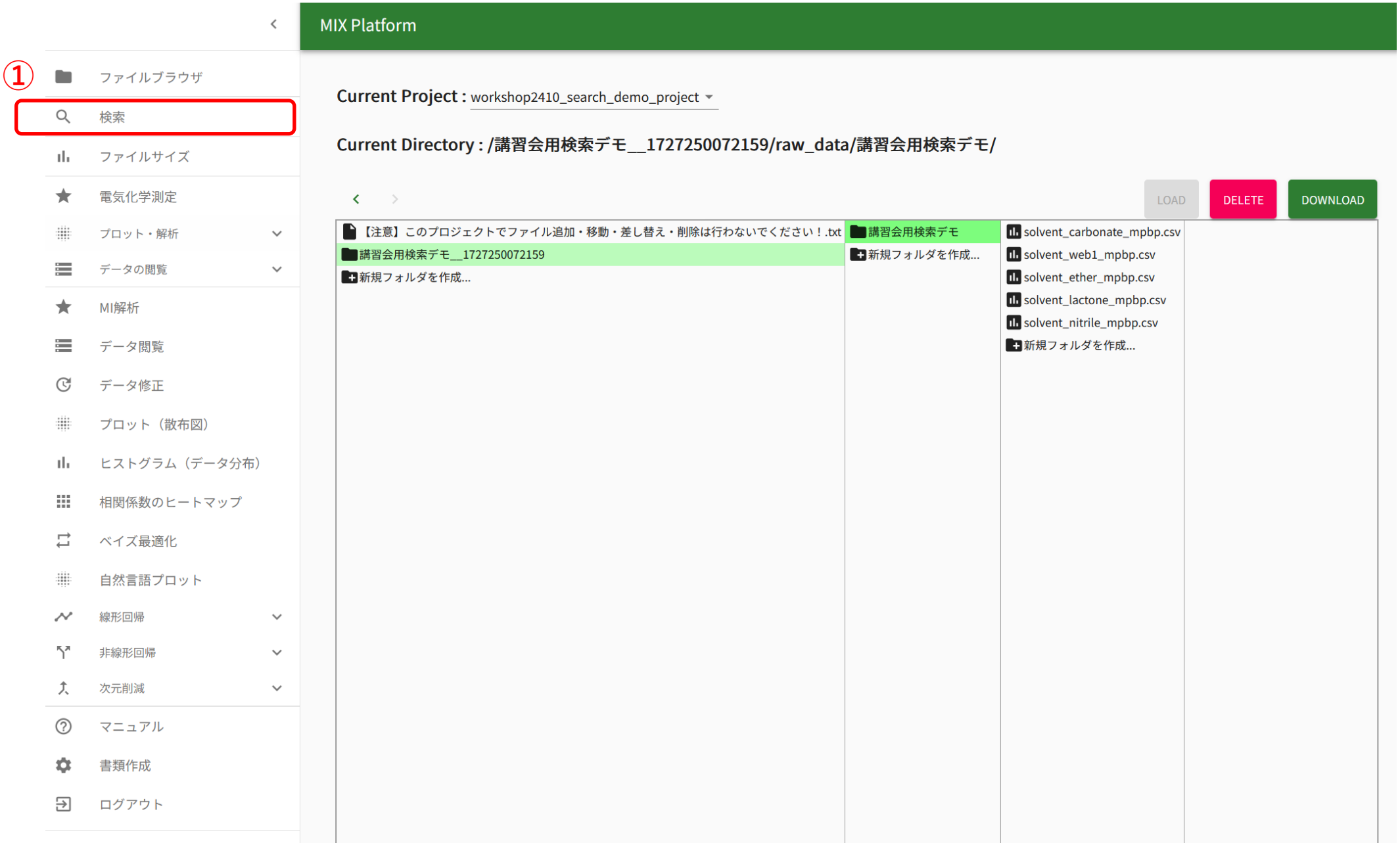
⇒大量に蓄積される、異なる性質のデータを利用した解析の実施が可能に。

⇒MIX platformに必要なデータを検索し、共通するデータを自動連携させてMI解析を実行する機能を搭載(中)

データ連携と検索

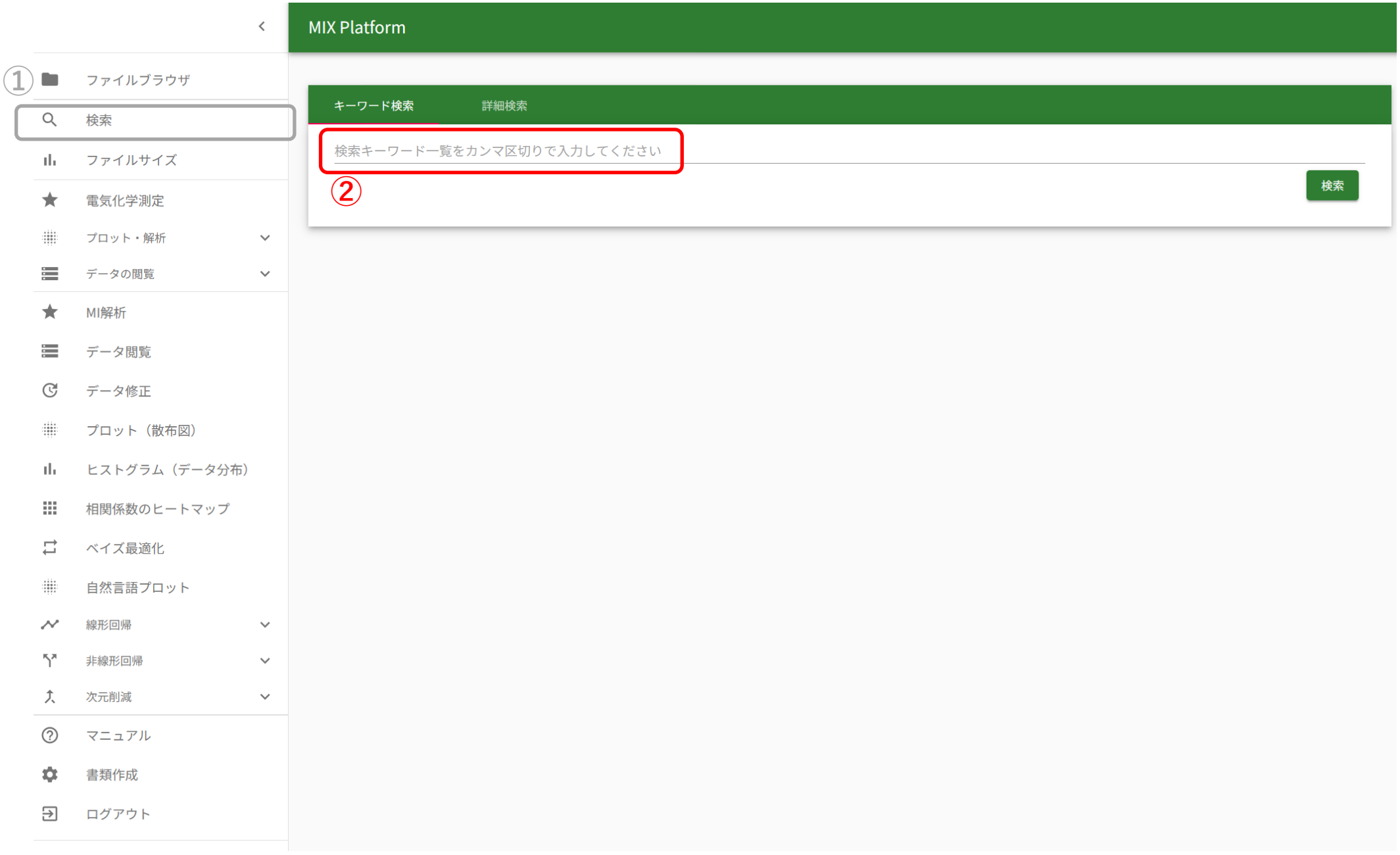
データ連携、検索パートは、まだ本番環境には実装されておらず、
以後デモバージョンの紹介となります。

データ連携と検索



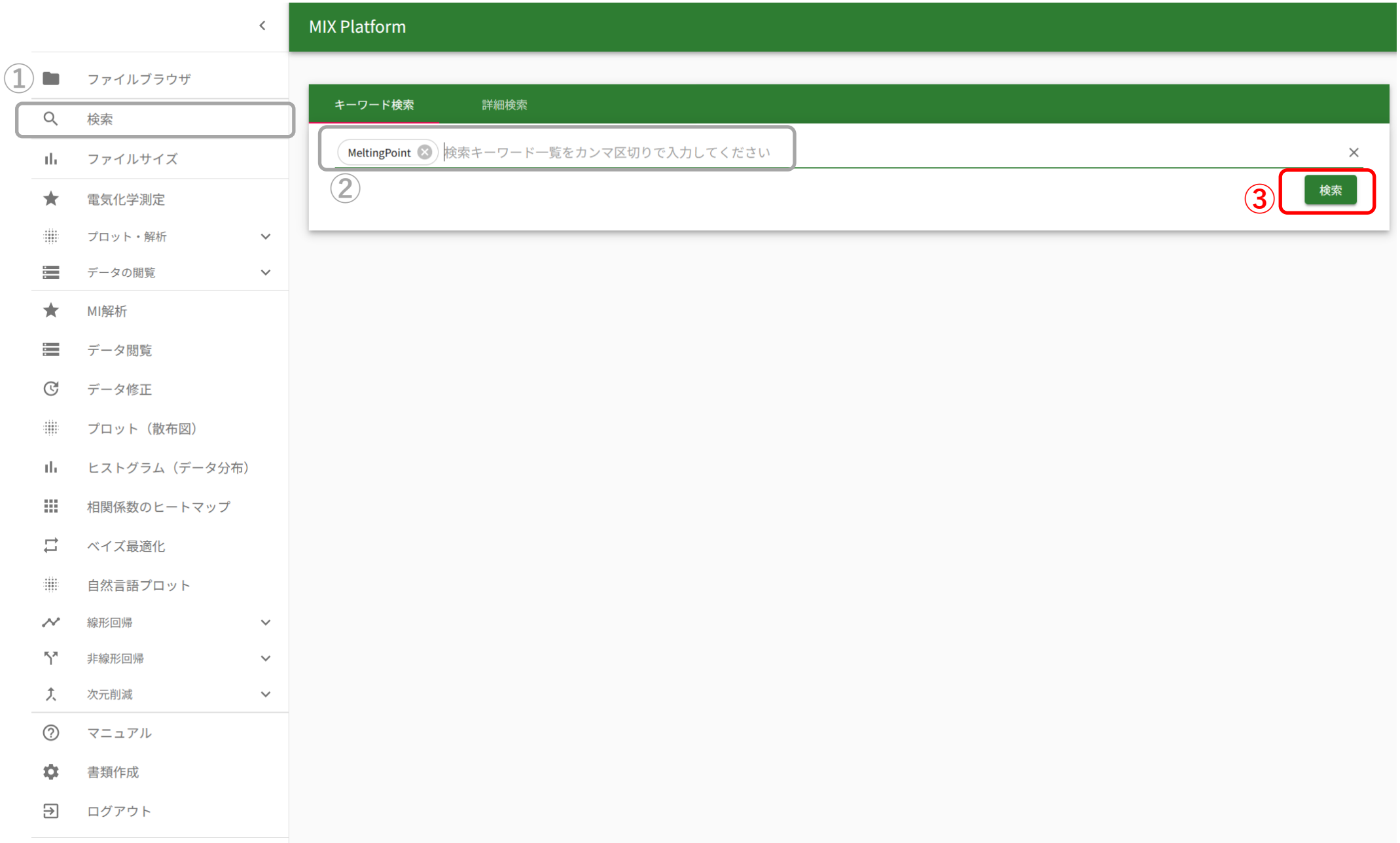
①「検索」をクリック

データ連携と検索



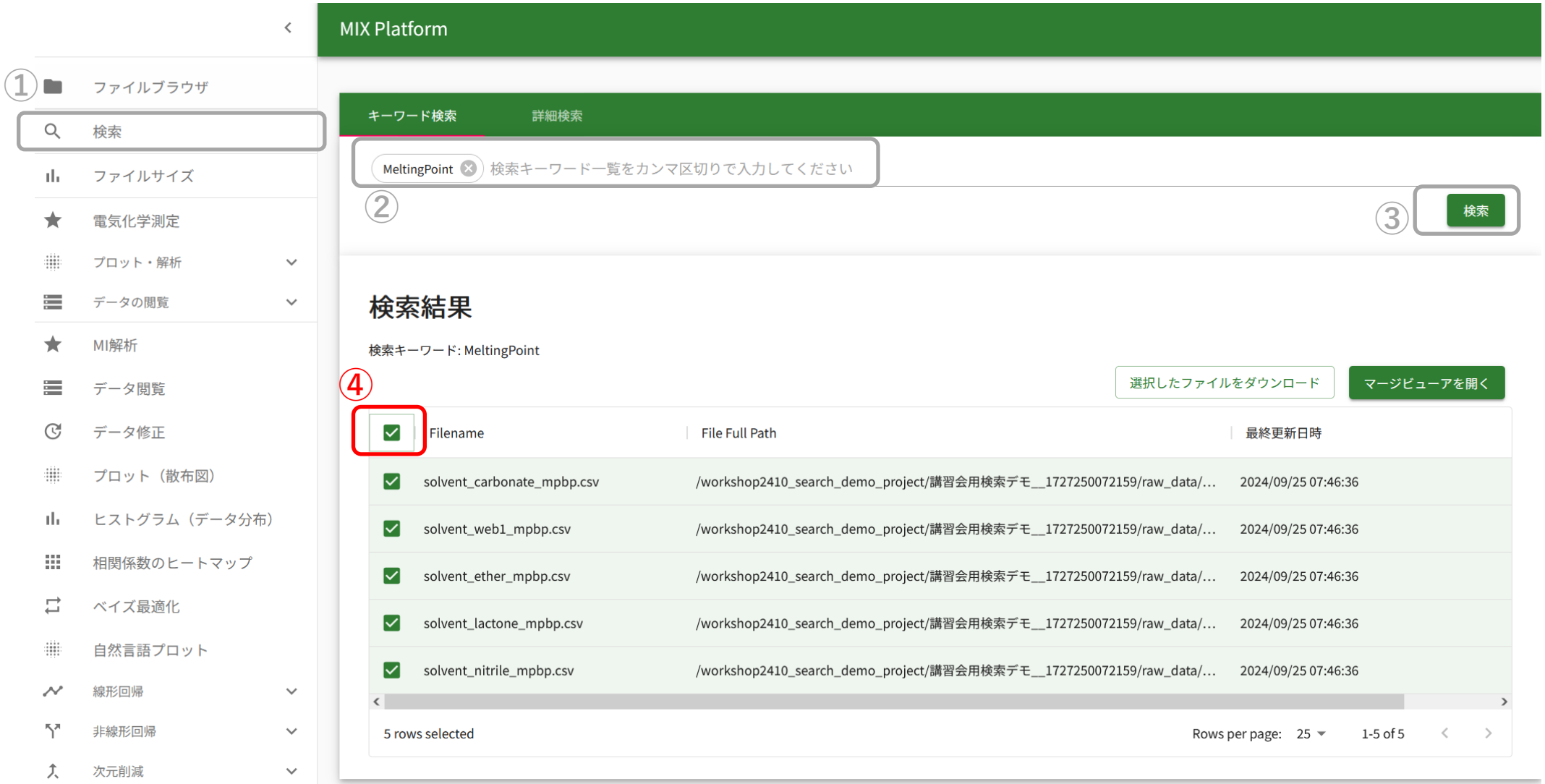
①「検索」をクリック
②にキーワードを入力
（入力しようとすると検索候補が自動表示）
⇒例えば
melting pointを選択

データ連携と検索



①「検索」をクリック
②にキーワードを入力
（入力しようとすると検索候補が自動表示）
⇒例えば melting pointを選択
③「検索」をクリック

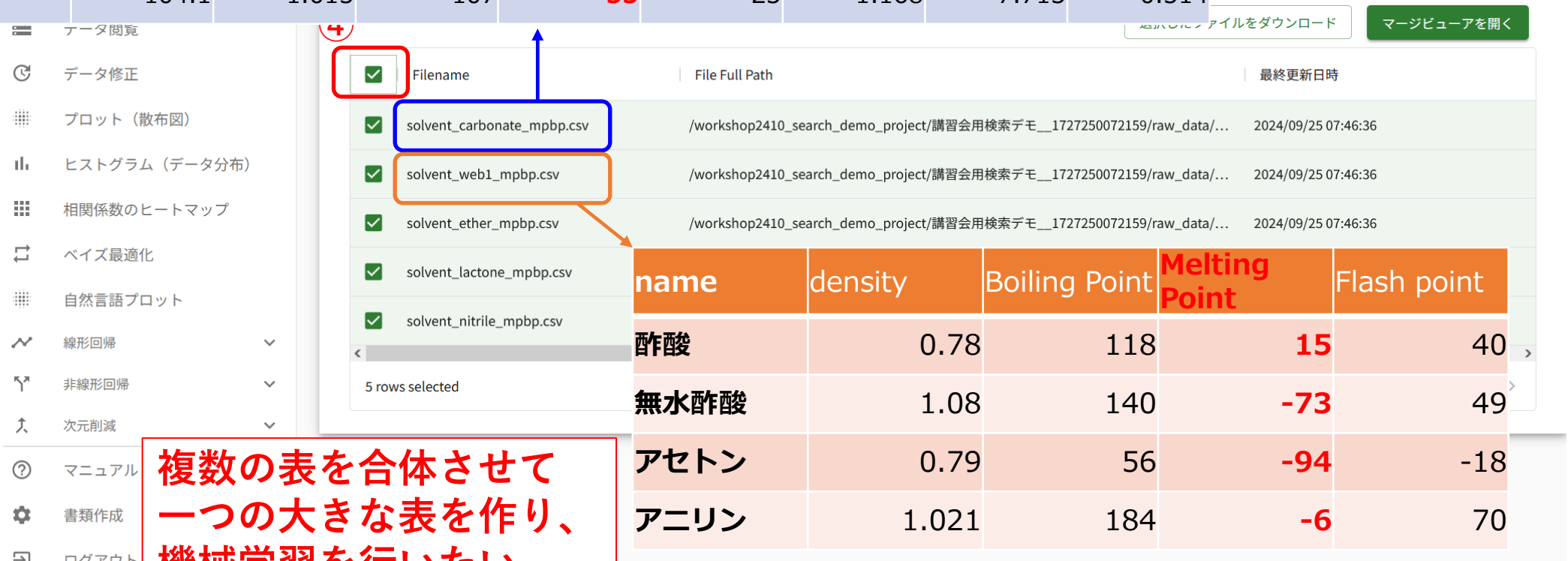
データ連携と検索



①「検索」をクリック
②にキーワードを入力
（入力しようとすると検索候補が自動表示）
⇒例えばmelting pointを選択
③「検索」をクリック
④melting pointの項目を含むファイルが表示される。ひとまず全ファイルを選択☑

name	Molecular weight	density	Boiling Point	Melting Point	Flash point	LUMO	HOMO	Dipole moment
PC	102.09	1.21	242	-49	132	0.946	-7.930	5.255
EC	88.06	1.32	244	38	152	0.919	-8.017	5.07
VC	86.05	1.255	162	22	72	-0.137	-6.973	4.365
FEC	106.05	1.497	210	17.3	122	0.493	-8.468	4.487
DMC	90.08	1.07	90	4	17	1.115	-7.774	0.342
DEC	118.13	0.975	126	-43	25	1.217	-7.654	0.613
EMC	104.1	1.015	107	-55	23	1.168	-7.713	0.514

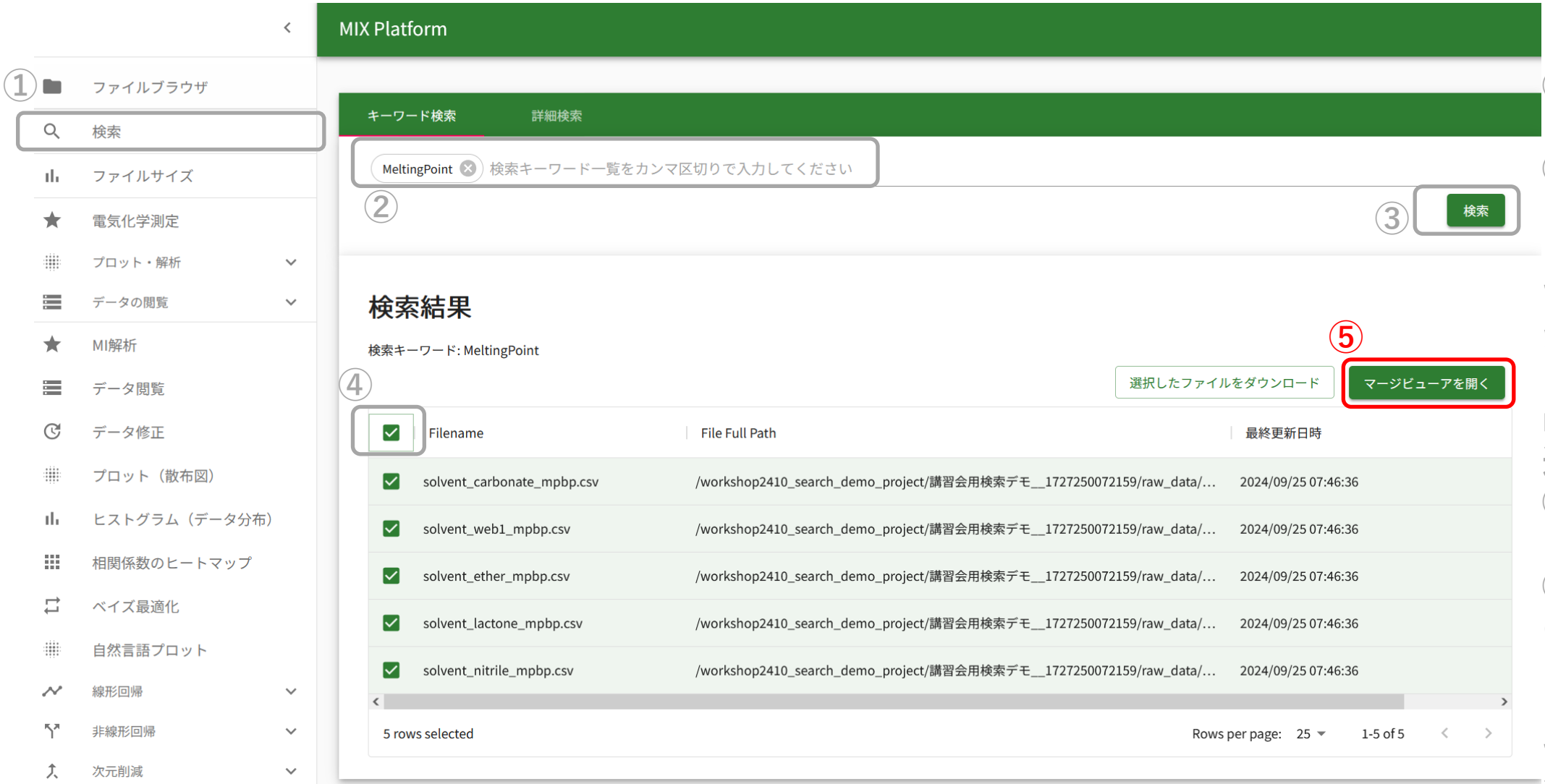
①「検索」をクリック
②にキーワードを入力
（入力しようとすると検索候補が自動表示）
⇒例えばmelting pointを選択
③「検索」をクリック
④melting pointの項目を含むファイルが表示される。ひとまず全ファイルを選択☑



name	density	Boiling Point	Melting Point	Flash point
酢酸	0.78	118	15	40
無水酢酸	1.08	140	-73	49
アセトン	0.79	56	-94	-18
アニリン	1.021	184	-6	70

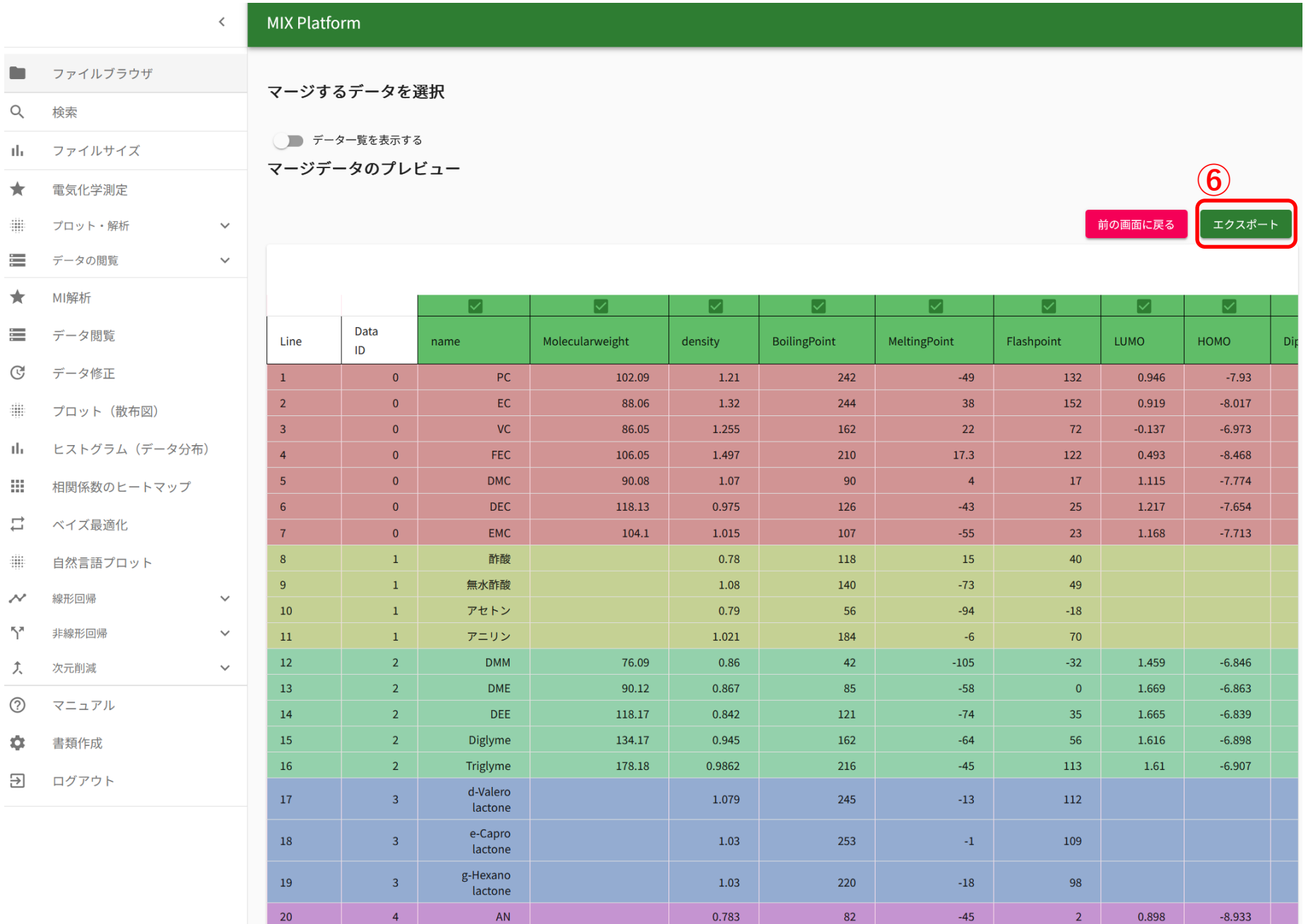
複数の表を合体させて一つの大きな表を作り、機械学習を行いたい。

データ連携と検索



①「検索」をクリック
②にキーワードを入力
（入力しようとすると検索候補が自動表示）
⇒例えばmelting pointを選択
③「検索」をクリック
④melting pointの項目を含むファイルが表示される。ひとまず全ファイルを選択☑
⑤「マージビューアを開く」をクリック

データ連携と検索



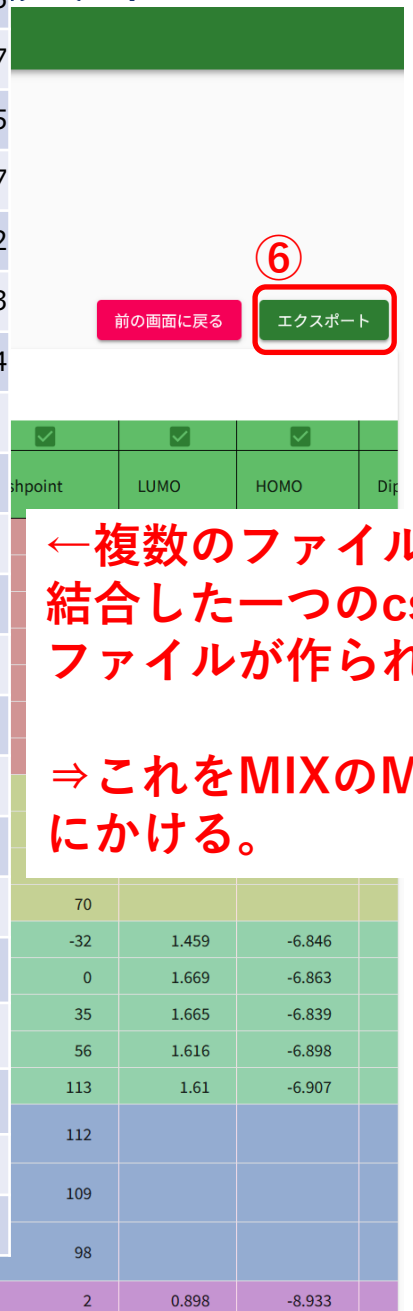
Line	Data ID	name	Molecularweight	density	BoilingPoint	MeltingPoint	Flashpoint	LUMO	HOMO	Dip
1	0	PC	102.09	1.21	242	-49	132	0.946	-7.93	
2	0	EC	88.06	1.32	244	38	152	0.919	-8.017	
3	0	VC	86.05	1.255	162	22	72	-0.137	-6.973	
4	0	FEC	106.05	1.497	210	17.3	122	0.493	-8.468	
5	0	DMC	90.08	1.07	90	4	17	1.115	-7.774	
6	0	DEC	118.13	0.975	126	-43	25	1.217	-7.654	
7	0	EMC	104.1	1.015	107	-55	23	1.168	-7.713	
8	1	酢酸		0.78	118	15	40			
9	1	無水酢酸		1.08	140	-73	49			
10	1	アセトン		0.79	56	-94	-18			
11	1	アニリン		1.021	184	-6	70			
12	2	DMM	76.09	0.86	42	-105	-32	1.459	-6.846	
13	2	DME	90.12	0.867	85	-58	0	1.669	-6.863	
14	2	DEE	118.17	0.842	121	-74	35	1.665	-6.839	
15	2	Diglyme	134.17	0.945	162	-64	56	1.616	-6.898	
16	2	Triglyme	178.18	0.9862	216	-45	113	1.61	-6.907	
17	3	d-Valero lactone		1.079	245	-13	112			
18	3	e-Capro lactone		1.03	253	-1	109			
19	3	g-Hexano lactone		1.03	220	-18	98			
20	4	AN		0.783	82	-45	2	0.898	-8.933	

⑥「エクスポート」をクリック
⇒データ連携されたcsvファイルが、ローカルPCにダウンロードされる。

←複数のファイルが、共通の項目名を持つものに対して、列を揃えて出力される。

検索

name	Molecular weight	density	Boiling Point	Melting Point	Flash point	LUMO	HOMO	Dipole moment
PC	102.09	1.21	242	-49	132	0.946	-7.93	5.255
EC	88.06	1.32	244	38	152	0.919	-8.017	5.07
VC	86.05	1.255	162	22	72	-0.137	-6.973	4.365
FEC	106.05	1.497	210	17.3	122	0.493	-8.468	4.487
DMC	90.08	1.07	90	4	17	1.115	-7.774	0.342
DEC	118.13	0.975	126	-43	25	1.217	-7.654	0.613
EMC	104.1	1.015	107	-55	23	1.168	-7.713	0.514
酢酸		0.78	118	15	40			
無水酢酸		1.08	140	-73	49			
アセトン		0.79	56	-94	-18			
アニリン		1.021	184	-6	70			
DMM	76.09	0.86	42	-105	-32	1.459	-6.846	
DME	90.12	0.867	85	-58	0	1.669	-6.863	
DEE	118.17	0.842	121	-74	35	1.665	-6.839	
Diglyme	134.17	0.945	162	-64	56	1.616	-6.898	
Triglyme	178.18	0.9862	216	-45	113	1.61	-6.907	
d-Valero lactone		1.079	245	-13	112			
e-Capro lactone		1.03	253	-1	109			
g-Hexano lactone		1.03	220	-18	98			
AN		0.783	82	-45	2	0.898	-8.933	
PN		0.772	97	-93	6	0.587	-8.802	

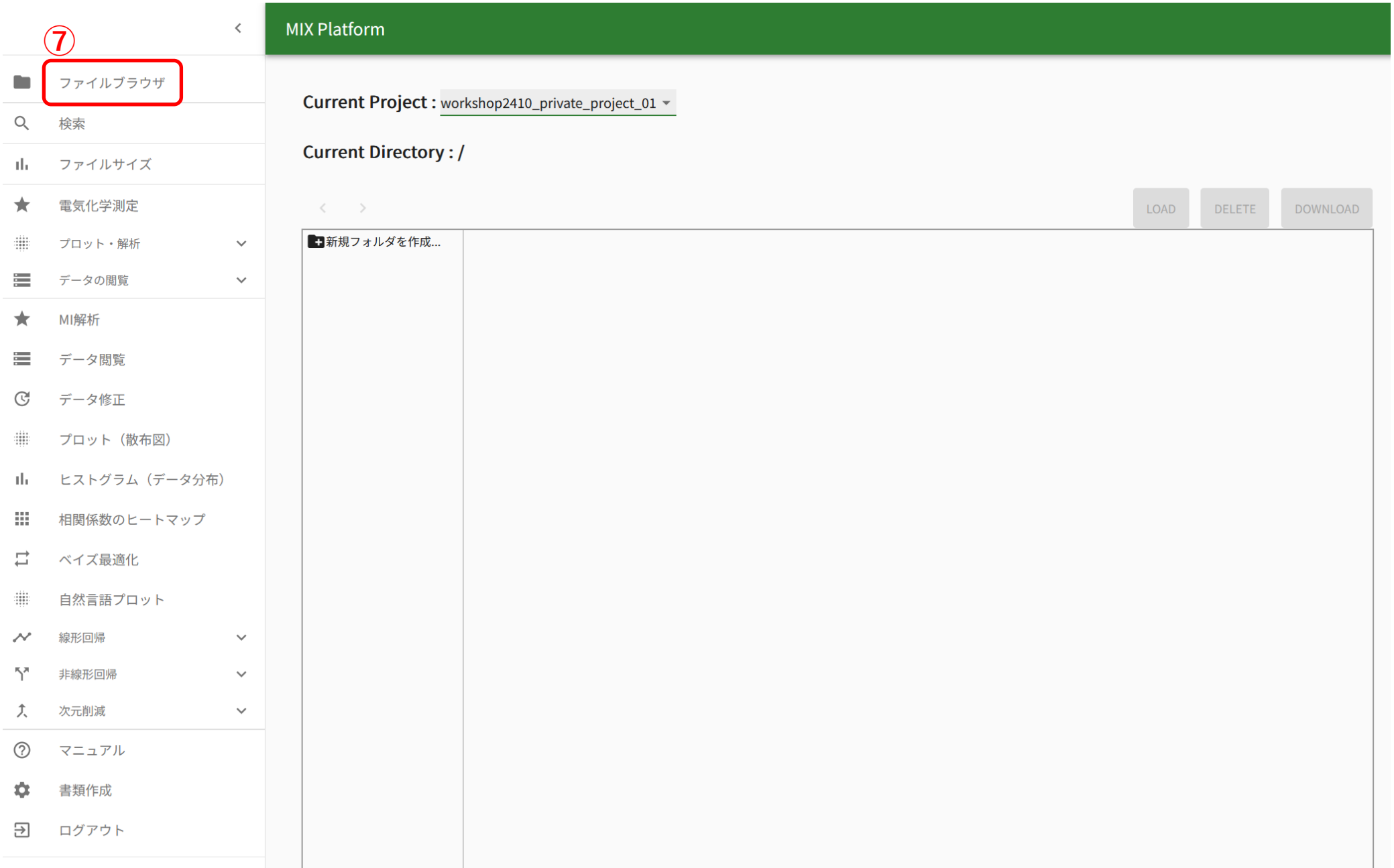


←複数のファイルが結合した一つのcsvファイルが作られる

⇒これをMIXのMI解析にかける。

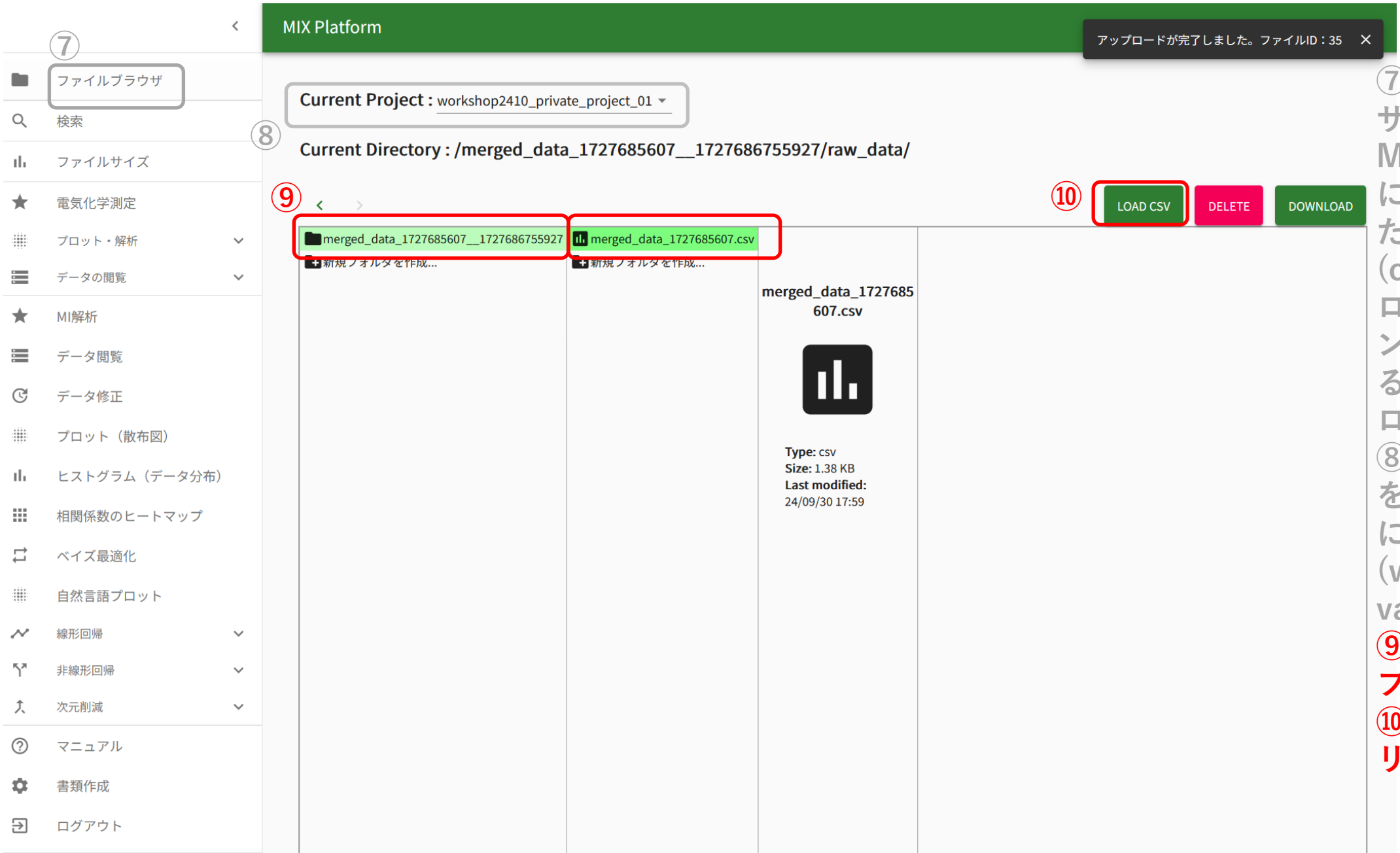
⑥「エクスポート」をクリック
⇒データ連携されたcsvファイルが、ローカルPCにダウンロードされる。

データ連携と検索



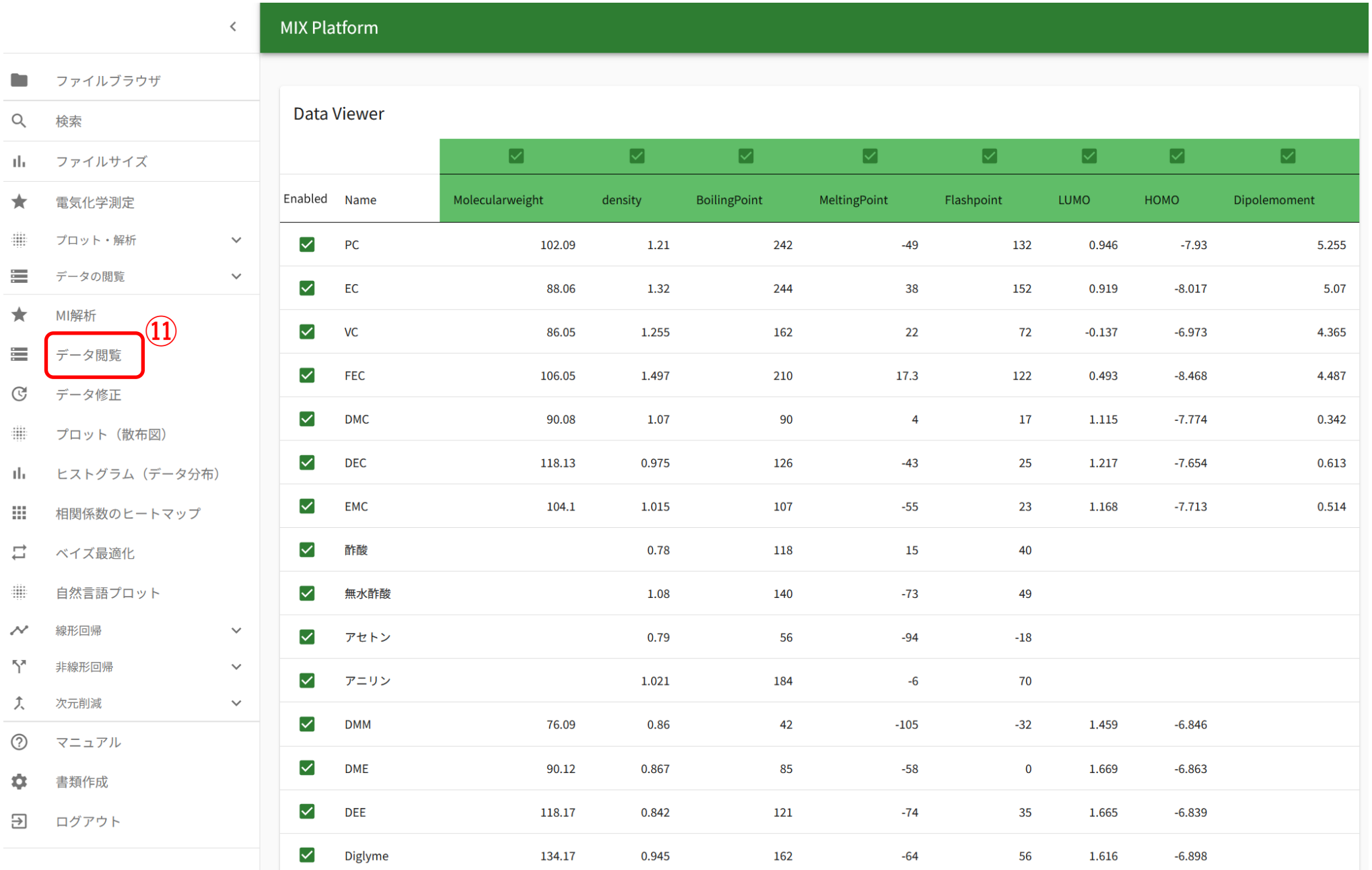
⑦「ファイルブラウザ」をクリック。MI解析をするために、先ほど合体させたExcelファイル(csvファイルとしてローカルPCにダウンロードされている)をMIXにアップロードする。

データ連携と検索



⑦「ファイルブラウザ」をクリック。
MI解析をするために、先ほど合体させたExcelファイル(csvファイルとしてローカルPCにダウンロードされている)をMIXにアップロードする。
⑧Current Projectを必ずprivateの方にする。(workshop2410_private_project_**)
⑨アップロードしたファイルを選択。
⑩LOAD CSVをクリック

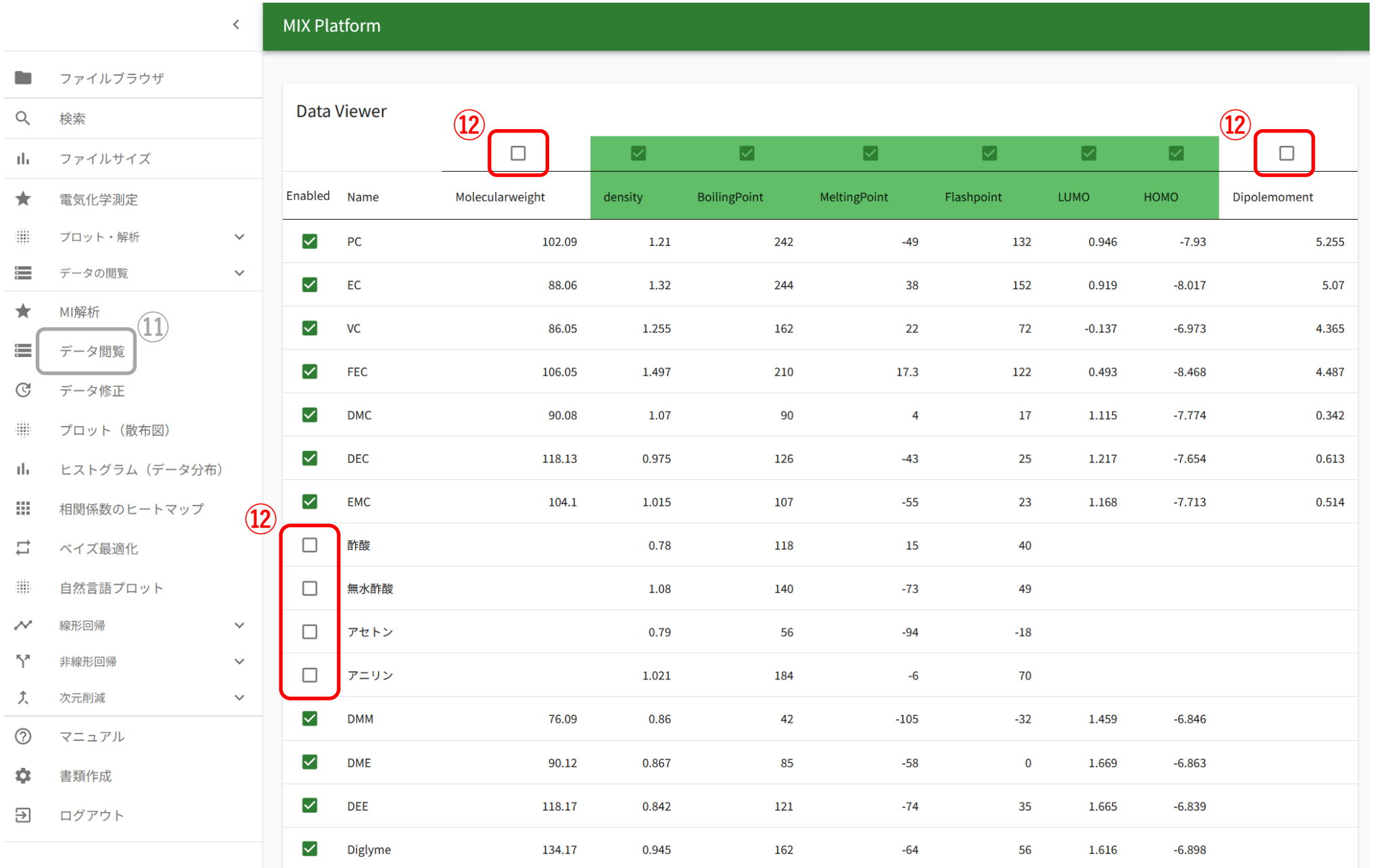
データ連携と検索



Enabled	Name	Molecularweight	density	BoilingPoint	MeltingPoint	Flashpoint	LUMO	HOMO	Dipolemoment
☑	PC	102.09	1.21	242	-49	132	0.946	-7.93	5.255
☑	EC	88.06	1.32	244	38	152	0.919	-8.017	5.07
☑	VC	86.05	1.255	162	22	72	-0.137	-6.973	4.365
☑	FEC	106.05	1.497	210	17.3	122	0.493	-8.468	4.487
☑	DMC	90.08	1.07	90	4	17	1.115	-7.774	0.342
☑	DEC	118.13	0.975	126	-43	25	1.217	-7.654	0.613
☑	EMC	104.1	1.015	107	-55	23	1.168	-7.713	0.514
☑	酢酸		0.78	118	15	40			
☑	無水酢酸		1.08	140	-73	49			
☑	アセトン		0.79	56	-94	-18			
☑	アニリン		1.021	184	-6	70			
☑	DMM	76.09	0.86	42	-105	-32	1.459	-6.846	
☑	DME	90.12	0.867	85	-58	0	1.669	-6.863	
☑	DEE	118.17	0.842	121	-74	35	1.665	-6.839	
☑	Diglyme	134.17	0.945	162	-64	56	1.616	-6.898	

⑪データ閲覧をクリック

データ連携と検索

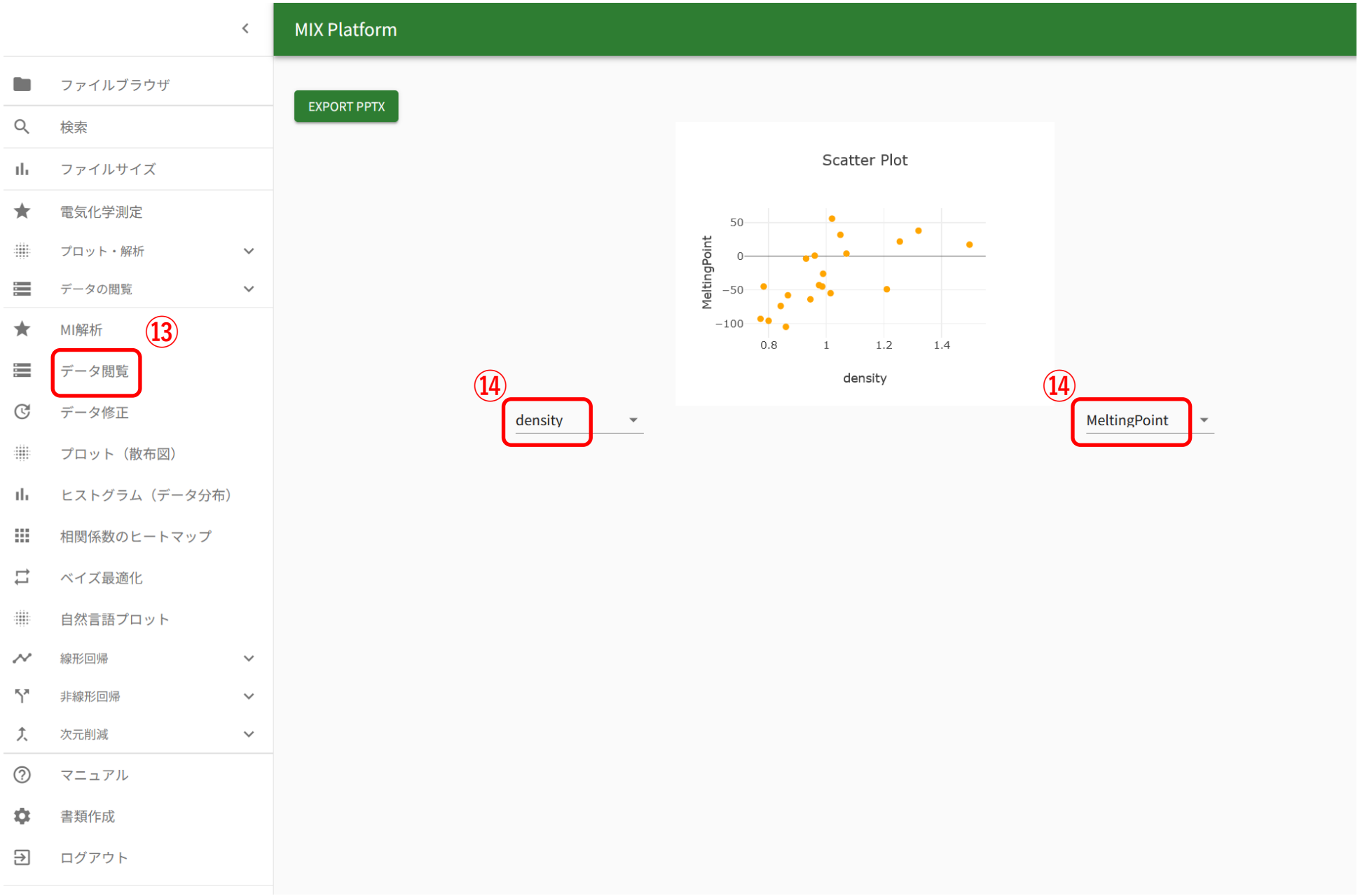


Enabled	Name	Molecularweight	density	BoilingPoint	MeltingPoint	Flashpoint	LUMO	HOMO	Dipolemoment
☑	PC	102.09	1.21	242	-49	132	0.946	-7.93	5.255
☑	EC	88.06	1.32	244	38	152	0.919	-8.017	5.07
☑	VC	86.05	1.255	162	22	72	-0.137	-6.973	4.365
☑	FEC	106.05	1.497	210	17.3	122	0.493	-8.468	4.487
☑	DMC	90.08	1.07	90	4	17	1.115	-7.774	0.342
☑	DEC	118.13	0.975	126	-43	25	1.217	-7.654	0.613
☑	EMC	104.1	1.015	107	-55	23	1.168	-7.713	0.514
☐	酢酸		0.78	118	15	40			
☐	無水酢酸		1.08	140	-73	49			
☐	アセトン		0.79	56	-94	-18			
☐	アニリン		1.021	184	-6	70			
☑	DMM	76.09	0.86	42	-105	-32	1.459	-6.846	
☑	DME	90.12	0.867	85	-58	0	1.669	-6.863	
☑	DEE	118.17	0.842	121	-74	35	1.665	-6.839	
☑	Diglyme	134.17	0.945	162	-64	56	1.616	-6.898	

⑪データ閲覧をクリック

⑫データ欠損が無いように、空白部分を消す。

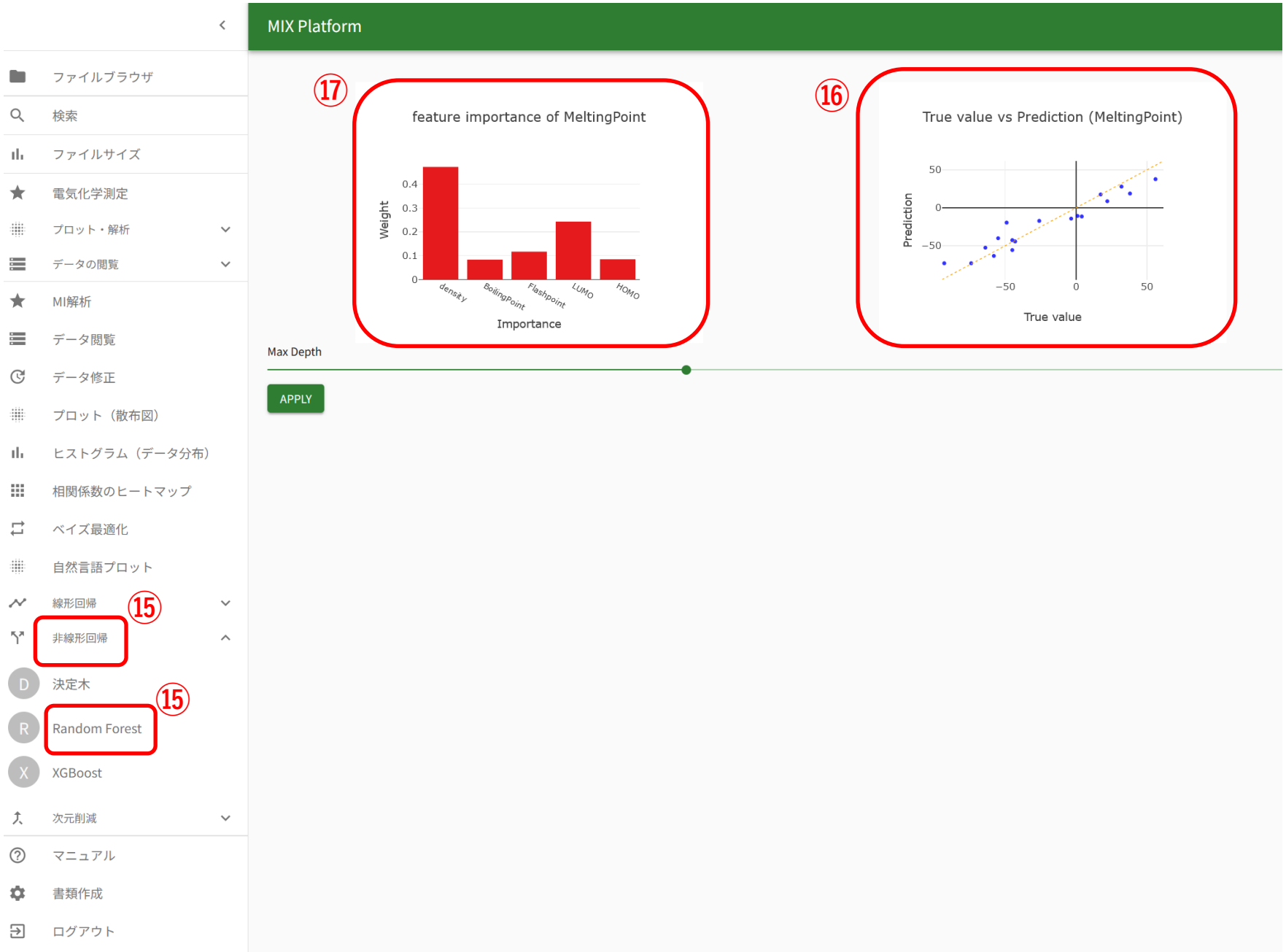
データ連携と検索



⑬MI解析のデータ閲覧をクリック

⑭例えば、
y軸をdensityに、
x軸をMeltingPointにする。
⇒MeltingPointを目的変数として予測を実行する。

データ連携と検索



⑮非線形回帰のRandom Forestをクリックして実行。

⑯比較的よくMelting Pointを予測できた。

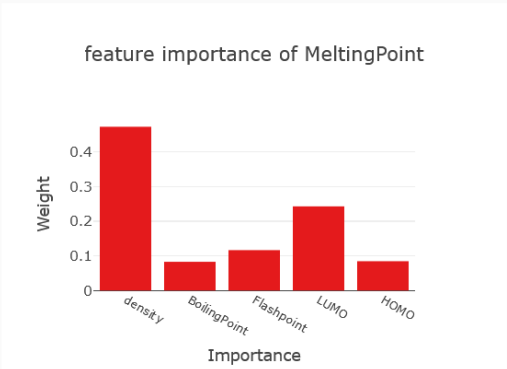
⑰Melting Pointの予測にはdensityが効く。（実験的に知られた事実）

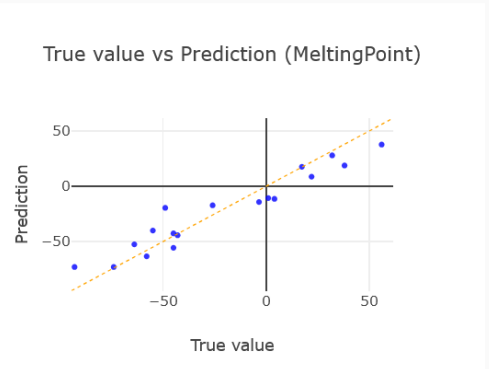
データ連携と検索

複数ファイルを連携して利用

Line	Data ID	name	Molecularweight	density	BoilingPoint	MeltingPoint	Flashpoint	LUMO	HOMO	Di
1	0	PC	102.09	1.21	242	-49	132	0.946	-7.93	
2	0	EC	88.06	1.32	244	38	152	0.919	-8.017	
3	0	VC	86.05	1.255	162	22	72	-0.137	-6.973	
4	0	FEC	106.05	1.497	210	17.3	122	0.493	-8.468	
5	0	DMC	90.08	1.07	90	4	17	1.115	-7.774	
6	0	DEC	118.13	0.975	126	-43	25	1.217	-7.654	
7	0	EMC	104.1	1.015	107	-55	23	1.16	-7.713	
8	1	酢酸		0.78	118	15	40			
9	1	無水酢酸		1.08	140	-73	49			
10	1	アセトン		0.79	56	-94	-18			
11	1	アニリン		1.021	184	-6	70			
12	2	DMM	76.09	0.86	42	-105	-32	1.459	-6.84	
13	2	DME	90.12	0.867	85	-58	0	1.669	-6.863	
14	2	DEE	118.17	0.842	121	-74	35	1.665	-6.839	
15	2	Diglyme	134.17	0.945	162	-64	56	1.616	-6.898	
16	2	Triglyme	178.18	0.9862	216	-45	113	1.61	-6.907	
17	3	d-Valero lactone		1.079	245	-13	112			
18	3	e-Capro lactone		1.03	253	-1	109			
19	3	g-Hexano lactone		1.03	220	-18	98			
20	4	AN		0.783	82	-45	2	0.898	-8.933	
21	4	PN		0.772	97	-93	6	0.587	-8.802	
22	4	Adiponitrile		0.96	295	1	93	0.312	-8.976	
23	4	Valeronitrile		0.8	140	-96	40	0.728	-8.704	
24	4	Glutaronitrile		0.989	286	-26	109	-0.136	-9.081	
25	4	Malononitrile		1.049	220	32	112	-0.51	-9.544	
26	4	Succinonitrile		1.02	267	56	158	-0.097	-9.31	
27	4	Suberonitrile		0.93	172	-3.5	161	0.497	-8.789	
28	4	Isobutyronitrile		0.77	107.5	-72	8	0.642	-8.727	

MIX Platform

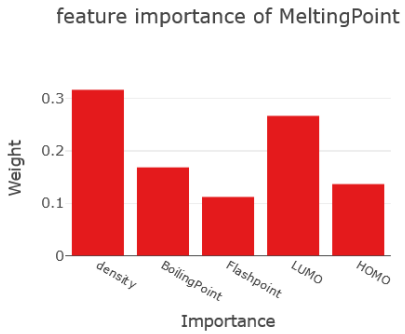


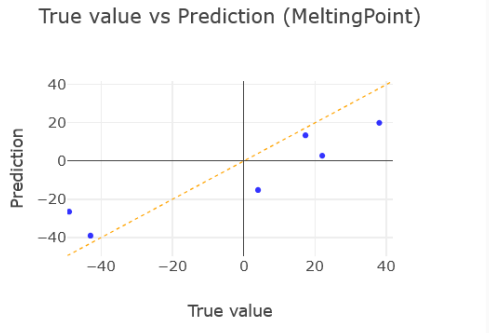


多数のデータで精度よく予測
densityという特徴量を正しく選択

1ファイルのみのデータを利用

MIX Platform





Max Depth

少数データでの予測となり、信頼度が低い
特徴量の選択も信頼性が下がる

まとめ

1．PEFCに特化したデータ解析・蓄積プラットフォーム（MIX Platform)を構築した。

2．本日は、電気化学測定データの解析と、MI解析部分に絞って説明した。

3．今後、機能のアップデートを順次実施。

4．企業、アカデミアの方のご利用をお待ちしています。
・アカデミアの方、ご連絡頂ければすぐにご利用可能です。
・企業の方、LIO後すぐにご利用可能です。

MI解析の勘所（MIノウハウ）

- どういう問題に対して、どういう手法を使えばよいか？

解析の目的	データ数	MI手法	例
1. 要因解析	少 多	回帰分析 Deep Learning	膜劣化の要因解析
2. (人に認識できない) 要因解析	中(スペクトル)	次元削減+回帰分析	膜劣化の要因解析
3. (実験が大変な) 物性予測	少 多	回帰分析 Deep Learning	触媒活性、耐久性予測
4. 実験条件最適化	少 多	ベイズ最適化 Deep Learning	触媒作成条件の最適化
5. 新物質提案（組成） 新物質提案（分子式）	少 中	ベイズ最適化 Deep Learning（SMILES-X）	新規触媒組成の提案 高分子膜材の提案