






# オープンサイエンスに向けた データ創出・活用 NIFS談話会 (第4回)



東京大学 物性研究所  
THE INSTITUTE FOR SOLID STATE PHYSICS  
THE UNIVERSITY OF TOKYO

2024/3/15(金) 15:30-17:00 東京大学 物性研究所 吉見 一慶

# 自己紹介

年月	経歴	
2009/03	東京大学理学系研究科 物理学専攻 博士課程修了 (東京大学物性研究所 加藤研究室) 有機伝導体での電荷揺らぎに関する理論研究	
2009/04- 2010/03	東京医科歯科大学 教養部 特任助教 (越野研究室) 量子光学に関する研究	
2010/04- 2012/03	東京大学理学系研究科 物理学専攻 特任研究員 有機伝導体での第一原理計算+厳密対角化を利用した研究	
2012/04- 2015/03	株式会社 構造計画研究所 入社 電波伝搬に関するシミュレーションコンサルティング	
2015/04-	東京大学物性研究所 特任研究員 (PI) ソフトウェア開発・高度化プロジェクト 2016/01-2020/03の期間は 計算物質科学人材育成コンソーシアム (PCoMS) 次世代研究員を兼任	



# Intrdocution (1)

## オープンサイエンスとは？～知識の共有

- **定義と目的**

- ✓ 研究の透明性とアクセシビリティを高めることを目指す運動
- ✓ 知識の民主化と、科学的発見への広範なアクセスを促進

- **主要要素**

- ✓ オープンアクセス: 研究論文やデータの無料公開
- ✓ オープンデータ: 研究データの共有と再利用
- ✓ オープンソース: ソフトウェアとツールの公開

- **メリット**

- ✓ 研究の再現性と検証の容易さ
- ✓ 学術コミュニティと一般社会の間の架け橋となる

# Introduction (2)

## インターネットの進化とオープンサイエンスのシナジー

- **情報共有の加速**

- ✓ インターネットの普及により、研究成果の共有が瞬時に
- ✓ オンラインジャーナル、プレプリントサーバー、学術ネットワークの利用拡大

- **アクセスの民主化**

- ✓ オープンアクセスジャーナルやオンラインデータベースにより、誰もが最新の研究にアクセス可能に。
- ✓ オープンソースソフトウェアとツールにより研究のハードルが低下。

- **メリット**

- ✓ 国内外を問わずに共同研究を促進。
- ✓ オンラインプラットフォームを通じたアイデアとデータの交換が可能に。

# Introduction (3)

## 科学を変革する情報通信技術(ICT)

- **第四のパラダイムとデータ駆動科学**

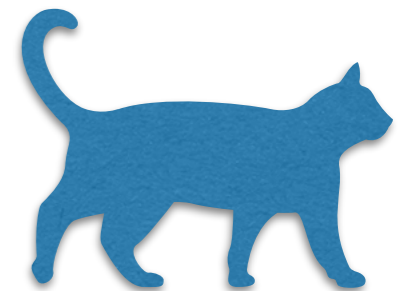
- ✓ 2009年 4th paradigm (by Jim Gray) : スーパーコンピュータによる計算科学の次のステージ
- ✓ Data-Intensive Scientific Discovery (いわゆるデータ駆動科学)

- **画像認識コンペ 「Googleの猫」 (2012年)**

- ✓ 一千万枚の画像を用いてAIを学習
- ✓ (猫を教えてないのに)猫を見分けられるようになった
- ✓ 独自にグルーピングができるようになった

- **生成AIの台頭 (2023年 ChatGPT etc.)**

- ✓ 複雑な会話、テキスト生成、知識の概説など、多岐にわたる能力を持つ
- ✓ 教育、ビジネス、エンターテインメントなど、多様な分野での活用が進み、コミュニケーションや情報検索の方法を革新



ChatGPT



# Introduction (4)

持続可能な未来のためのオープンサイエンス: 共に作り、成長するには？

- **データの創出、品質、および環境提供**

- ✓ 機械学習の精度向上のために質の高いデータを如何に創出するか。
- ✓ 高品質データの創出方法とその重要性
  - ✓ オープンソフトウェアとスーパーコンピュータを利用した効率的なデータ生成
  - ✓ 研究データの共有とアクセスを支援するインフラの提供

- **知の共有、人材育成、および多様性の促進**

- ✓ 最新のツールや技術の習得を支援する教育プログラム
- ✓ 多様なバックグラウンドを持つ参加者によるコミュニティの構築

- **持続可能性及び相互運用性**

- ✓ リソースの長期的な持続可能性と拡張性
- ✓ データとツールの相互運用性

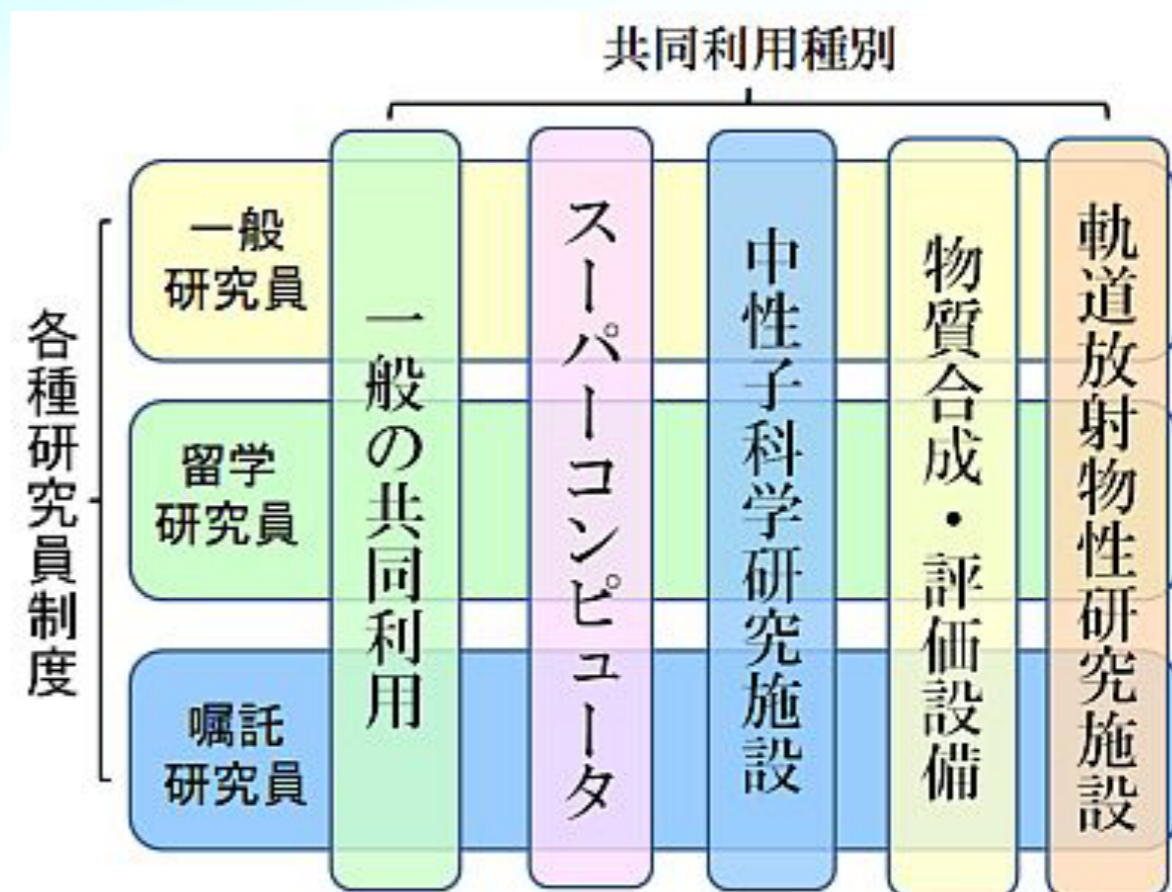
# 東京大学物性研究所の紹介

目的：物性科学の研究推進

～物性分野における世界最高水準の基礎研究の先導

1957年：東京大学附置の全国共同利用研究所として設立

2010年：物性科学研究拠点として文部科学省から認可



全国の大学・公的研究機関に所属する物性研究者は無償で利用可能

# 本日本話しすること

- オープンサイエンスに関連した物性研究所での取り組み
- オープンソースソフトウェアに関する取り組みの紹介
- オープンデータに関する取り組みの紹介
- 人材育成に関連した取り組みの紹介



# 本日本話しすること

- オープンサイエンスに関連した物性研究所での取り組み
  - オープンソースソフトウェアに関する取り組みの紹介
  - オープンデータに関する取り組みの紹介
  - 人材育成に関連した取り組みの紹介

# 物質科学とデータ駆動科学

## ◆ 物質科学の目的

- ◆ 欲しい物性の予言、新しい物質の設計、合成方法の最適化 etc.

## ◆ 既存の方法とデータ駆動科学を用いた方法の違い

- ◆ 「試行錯誤」による研究から「データと計算科学」による研究へ



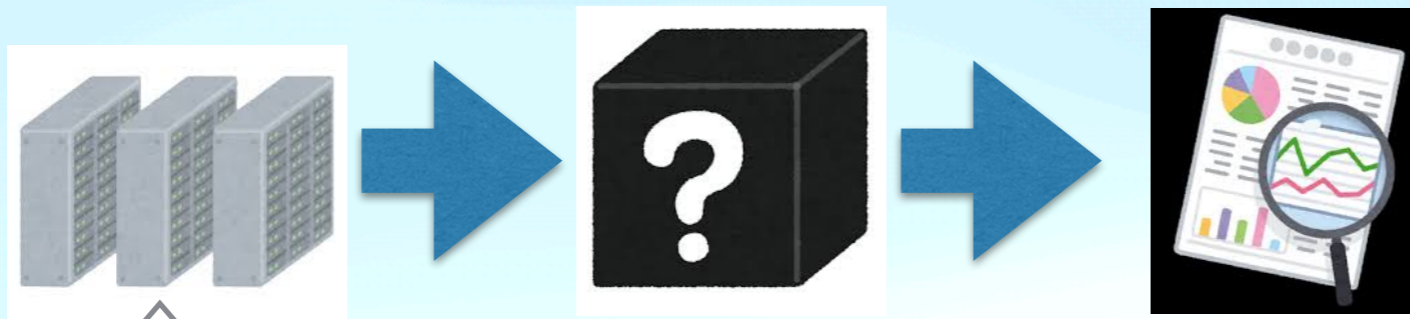
## ◆ 機械学習を用いた実際の応用例

- ◆ 理論計算・解析部分の高速化
- ◆ データ同化 (実験と理論の架け橋)
- ◆ 欲しい物質の探索 etc.

# より高精度な機械学習を実現するには？ (1)

- ◆ **機械学習: 学習データの質と量が重要**

- ◆ **1st step: 既存のデータベースを用いてモデルを学習させる**



- ◆ **マテリアルインフォマティクスに関連した有名なプロジェクト**
  - ◆ **Materials Project[1], AFLOW[2], OQMD[3], JARVIS[4], NOMAD[5] etc.**

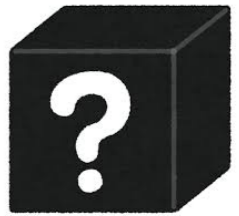
[1] <https://next-gen.materialsproject.org> , [2] <https://www.aflowlib.org/>, [3] <https://oqmd.org/>  
[4] <https://jarvis.nist.gov>, [5] <https://nomad-lab.eu/nomad-lab>

- ◆ **次のstep: 目的に応じた高品質なデータを用い、モデルをファインチューニングする。**



# より高精度な機械学習を実現するには？ (2)

- ◆ 目的に応じたデータセットを得る際のコスト
  - ◆ 実験 > シミュレーション
- ◆ 計算物質科学
  - ◆ シミュレーション：物理法則に基づいて行われる
    - ◆ メリット：未知の領域を合理的に外挿することが可能
    - ◆ デメリット：計算時に近似が含まれる場合が多い
  - ◆ データの質を担保するには？
    - ◆ 目的・物質に応じて適切な近似を選択する必要性がある
      - ◆ それらを実装したソフトウェアの開発・探索が重要！
  - ◆ データの量を担保するには？
    - ◆ ハイスループット計算が効果的
      - ◆ 特にスーパーコンピュータを用いたデータ創出は強力！



# より高精度な機械学習を実現するには？ (3)

## ソフトウェアをうまく使うまでの障壁

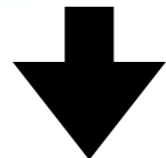
Q1. やりたいことができるソフトウェアをどうやって探せばよい？

(世の中にはたくさんのソフトウェアがあって探すのが大変...)



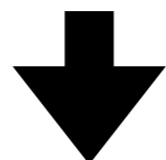
Q2. ソフトウェアを簡単に試せるか？

(ソフトウェアを使うのに複雑なインストール作業が必要だと心が折れる...)



Q3. ソフトウェアを使いこなすための計算環境はあるか？

(計算はできるけどノートPCだと30年かかる...)



Q4. ソフトウェアは簡単に使えるのか？

(使いこなすために高度な専門知識が必要だと辛い...)

# 本日本話しすること

- オープンサイエンスに関連した物性研究所での取り組み
- オープンソースソフトウェアに関する取り組みの紹介
  - ✓ ソフトウェアを見つける
  - ✓ ソフトウェアを試す
  - ✓ ソフトウェアを開発する
- オープンデータに関する取り組みの紹介
- 人材育成に関連した取り組みの紹介



# 物性研究所でのソフトウェア・データに関連した取り組み

2013年～

ポータルサイトの運用

## MateriApps

物質科学シミュレーション  
ポータルサイト

アプリを探す/試す

2010

2015年～

ソフトウェア開発・高度化

## PASUMS

Project for advancement of  
software usability in materials science

大型計算機向けに使いやすい  
ソフトウェアの整備・開発

<<スパコンの活用>>

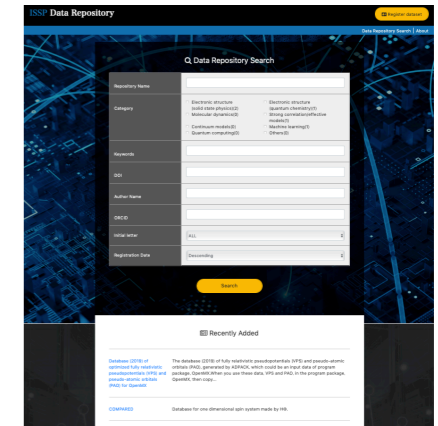
高精度データ解析

ハイスループット計算

2015

2021年～

データリポジトリ運用



物質科学データ向け  
リポジトリの開発

物質科学データ  
の蓄積・活用

2020

# 物性研究所でのソフトウェア・データに関連した取組

2013年～

ポータルサイトの運用

**MateriApps**

物質科学シミュレーション  
ポータルサイト

アプリを探す/試す

2010

2015年～

ソフトウェア開発・高度化

**PASUMS**

Project for advancement of  
software usability in materials science

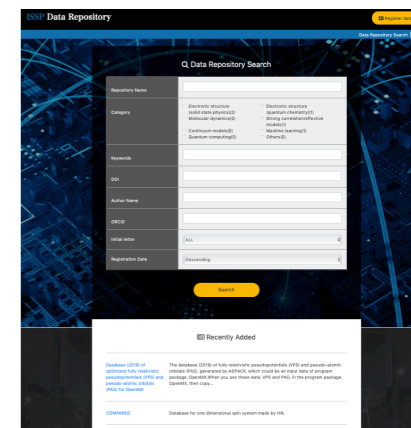
大型計算機向けに使いやすい  
ソフトウェアの整備・開発

<<スパコンの活用>>  
高精度データ解析  
ハイスループット計算

2015

2021年～

データリポジトリ運用



物質科学データ向け  
リポジトリの開発

物質科学データ  
の蓄積・活用

2020

# “MateriApps”とは？

## 物質科学シミュレーションのポータルサイト

<https://ma.issp.u-tokyo.ac.jp>

2019年に文部科学大臣賞受賞！



使いたいアプリ発見  
への障壁を下げる！

- アプリの検索

- アプリへの質問回答

月間 **6,400 ユーザ** , **17,700PV** (2022年度平均)

現メインメンバー：

ISSP (CCMS)：井戸康太、加藤岳生、川島直輝、古宇田光、福田将大、山崎淳、

ISSP以外：藤堂眞治(東大理)、笠松秀輔(山形大)、吉澤香奈子(RIST)



# アプリを見つける (1)

## カテゴリから探す

電子状態計算  
(固体物理分野) 83件

電子状態計算  
(量子化学分野) 39件

分子動力学 31件

可視化・モデリング 40件

強相関系・有効模型計算 64件

データ解析・補助ツール 59件

連続体シミュレーション 13件

データベース 13件

統合環境 4件

機械学習 21件

量子計算 8件

アプリ掲載件数 317件 (2024/03/13現在)

# アプリを見つける (2)

## 特徴

MateriApps LIVE!  
登録アプリ

公開度：2 ★☆☆以上

ドキュメント充実度：  
2 ★☆☆以上

講習会開催アプリ

開発者の声 掲載アプリ

レビューあり

## 計算手法・アルゴリズム

解説 をクリックするとキーワードの解説が表示されます

解説

X線分光解析

解説

エネルギー表示法

解説

拡散モンテカルロ法

解説

擬ポテンシャル法

解説

クラスター展開法

解説

KKR法

解説

原子局在基底・ガウシアン基底

解説

厳密対角化

解説

古典分子動力学法

解説

古典モンテカルロ法・  
量子モンテカルロ法

解説

コヒーレントポテンシャル近似  
(CPA)

解説

GW法

解説

解説

解説

## 対象

化学反応

光学スペクトル

生体高分子・ソフトマター

電子輸送

トポロジカル物質

熱力学量・熱力学相図

バンド計算・結晶構造

分子構造

フォノン・原子振動

磁化・電気分極

各種キーワードの  
解説ページもあり。

詳細検索から、  
特徴、対象、計算手法・アルゴリズム  
による絞り込み検索が可能。

# アプリの検索例

## HΦ

公開度：3 ★★★ ドキュメント充実度：3 ★★★

並列計算機に対応した数値厳密対角化法による有効モデルソルバーパッケージ。広汎な多体量子系の有効モデル(多軌道ハバード模型、ハイゼンベルグ模型、近藤格子模型など)の基底状態及び低励起状態の波動関数を並列計算によって求める。ランチョス法による基底状態計算、熱的純粋量子状態を利用した比熱・帯磁率の温度依存性計算が可能。さらに、シフト型クリロフ部分空間ライブラリ [Kw](#) を用いて動的グリーン関数の計算が可能である。ver. 3.0からは実時間発展の機能も追加された。



01 / 06

01 基本情報

02 事例・論文

03 講習会

04 レビュー

05 開発者の声

06 このアプリ情報についてのお問い合わせ

### 基本情報

#### 公式サイト

<http://issp-center-dev.github.io/HPhi/index.html>

マニュアルページ

フォーラム

ダウンロードページ

#### ライセンス

開発者の紹介,  
簡単なTutorial,  
講習会情報,  
事例紹介なども掲載

# アプリに関する質問の受付

Concierge

## アプリコンシェルジュ

アプリコンシェルジュに質問する

MateriAppsの中の人「アプリコンシェルジュ」となり、アプリに関するよくある質問にご回答するほか、「こんな計算にはこのアプリがオススメ」などアプリをご紹介します。

「アプリコンシェルジュ」に質問したい、このアプリを解説してほしい、などのご要望はこちらからどうぞ。

### アプリのライセンスの違いがわからないのですが。

アプリの利用にあたっては、ライセンス形態 (利用許諾) は重要な要素です。ライセンスは大きく分けて、有償か無償かの2つに分けられます。有償アプリはサポート・ドキュメント・GUIが充実しており、予算があれば十分に検討の余地… [続きを読む](#)

### 第一原理計算で求められる物理量ってなんですか？

第一原理計算では、電子状態に関する情報 (電子密度・バンド分散・状態密度・フェルミ面) が得られます。電子密度やフェルミ面などの可視化には、専用の可視化アプリを使うと

#### 👑 App Ranking

👑 VESTA (1→)

🥈 QUANTUM ESPRESSO (2→)

🥉 Materials Project (4👉)

VASP (7👉)

Bader Charge Analysis (11👉)

**MateriApps チームメンバー が質問に回答**



# (補) アプリの登録をするには？

## Webページの一番下から依頼



お問い合わせ

アプリ掲載依頼

※は必須項目

お名前※

所属

メールアドレス※

メールアドレス (確認) ※

アプリ名

アプリのURL

簡単な説明※

その他

App Ranking

- 1 VESTA (1→)
- 2 QUANTUM ESPRESSO (2→)
- 3 Materials Project (4)
- XenonPy (3)
- VASP (5→)
- Crystallography Open Database (7)
- Bader Charge Analysis (6)
- HΦ (9)
- ABINIT (10)
- BoltzTraP2 (8)

<< 登録までの基本的な流れ >>

基本情報+簡単な説明を記載

→ 掲載可否の判断後に担当者から連絡

→ 詳細情報記載後にアプリ掲載

→ MateriAppsのNEWSにも掲載

# データ活用・創出型マテリアル研究開発プロジェクト (2022年～)



データ連携部会(DxMT)の広報業務をNIMSより外部委託  
→ データ駆動型研究向けポータルサイトMatDaCsの開発  
データベースや解析ツールの情報を紹介

## ISSP 主要メンバー



尾崎 泰助



三澤 貴宏



吉見 一慶

## データ連携部会の概要

データ連携部会は横串機能として、5つの拠点に共通する技術課題の開発と共有、広報、人材育成によって、事業を推進するエンジンとなります。

社会像	カーボンニュートラル	Society5.0	国土強靱化	Well-Being社会
対象	蓄電・水電解触媒	磁石	半導体	バイオアダプティブ
代表機関	東京大学	NIMS データ連携部会 中核機関	東京工業大学	東北大学
代表研究者	杉山 正和 先端科学技術 研究センター 教授	大久保 忠勝 磁性・スピントロニクス材料研究拠点 副拠点長	神谷 利夫 教授 元素戦略研究 センター センター長	吉見 享祐 工学研究科 教授
連携機関	・NIMS ・東京理科大 ・横浜国立大 ・早稲田大 ・大阪公立大 ・名工大 ・東北大 ・理研	・北海道大 ・山梨大学 ・東京農工大 ・AIIST ・東北学院大 ・東北大学 ・名古屋大学 ・JASRI ・大阪大学 ・トヨタ自動車 ・TMI ・東京大学	・NIMS ・KEK ・JFCC	・九州大学 ・大阪大学 ・NIMS ・東京大学 ・横浜国立大学 ・JAEA ・名古屋大学
先端設備	放射光 (Spring-8, SACLA) ・中性子 (J-PARC, QST) ・電子顕微鏡等の活用			
計算科学	スパコン「富岳」の活用			
MDX	データ中核拠点・マテリアル先端リサーチインフラとの連携			



## MatDaCsのトップページ



<https://dxmt.mext.go.jp> より抜粋

<https://mat-dacs.dxmt.mext.go.jp/>

# 本日本話しすること

- オープンサイエンスに関連した物性研究所での取り組み
- オープンソースソフトウェアに関する取り組みの紹介
  - ✓ ソフトウェアを見つける
  - ✓ ソフトウェアを試す
  - ✓ ソフトウェアを開発する
- オープンデータに関する取り組みの紹介
- 人材育成に関連した取り組みの紹介

# アプリを試す (1) ~ MateriAppsLIVE!

- MateriApps Live!の特徴

- Virtual Boxで起動可能 (Debian Linux system)
- インストールなしで計算物質のアプリが実行可能

- Version 4.1 (2023/04/09にリリース, **Docker**に対応)

- **プリインストール**されている計算物質アプリと**可視化ツール**

- abinit, AkaiKKR, ALPS, CP2K, ERmod, DSQSS, Gromacs, HΦ, LAMMPS, mVMC, OCTA, OpenMX, Quantum Espresso, SMASH, TeNeS, xTAPP...

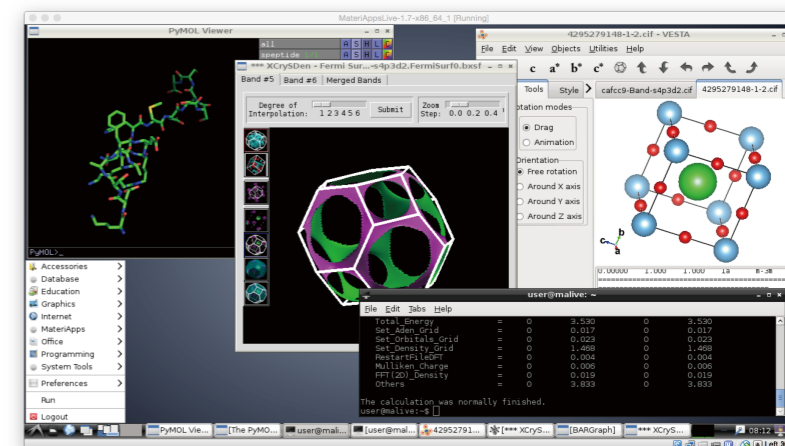
- ParaView, Tapioca, VESTA, VMD, XCrysDen ...

- **CCMSハンズオン講習会**での活用

- 月に一度の頻度で物性研究所で開催
- 授業などでも活用



2022/03/02



"MateriApps LIVE! and MateriApps Installer: Environment for starting and scaling up materials science simulations"  
Yuichi Motoyama, Kazuyoshi Yoshimi, Takeo Kato, Synge ToDo, SoftwareX, 20, 101210 (2022).



# アプリを試す (2) ~ MateriApps Installer

## 物質科学アプリのインストールスクリプト集

- 多くのプラットフォームに対応
  - スーパーコンピュータ・ワークステーション
  - 自前のデスクトップPC・ノートPC
  - 共同利用スパコン（物性研）
  - 標準的なクラスター計算機
  - Mac OS X, Linux, wsl2

MateriApps Installer  
の主開発者



藤堂眞治 (東大理)

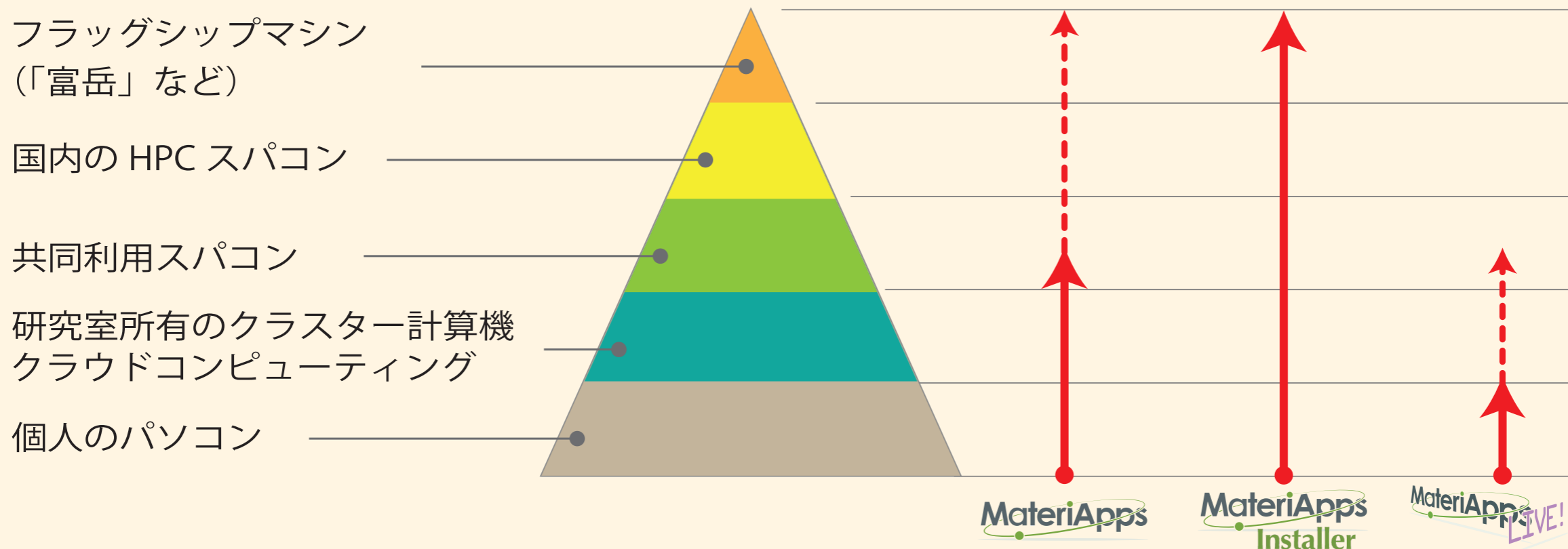
後述するソフトウェア開発・高度化プロジェクトに2020年度課題としても採択  
課題名：MateriApps Installer – オープンソースソフトウェアインストールツール  
提案者：藤堂 眞治 (東京大学 大学院理学系研究科)  
マニュアル整備などを実施し、ver.1.0としてリリース。

GitHub: <https://github.com/wistaria/MateriAppsInstaller>

公式ページ: <https://www.pasums.issp.u-tokyo.ac.jp/mainstaller/>

”MateriApps LIVE! and MateriAppsInstaller: Environment for starting and scaling up materials science simulations”  
Yuichi Motoyama, Kazuyoshi Yoshimi, Takeo Kato, Synge ToDo, SoftwareX, 20, 101210 (2022).

# アプリを試す (3) ~ 全体の位置付け



## アプリのインストールスクリプトの整備

- 開発ツール: GCC, CMake, Git, Python, Boost, ...
- 計算物質科学アプリ: ALPS, HΦ, mVMC, LAMMPS, OpenMX, QE, ...
- カバーしているOS  
GNU/Linux(Ubuntu, Debian, RedHat, CentOS etc.), macOS etc.
- MA Installerを用いてプリインストールされている(いた)スパコン  
「京」コンピュータ, ISSP supercomputer (システム B, C)

# 本日本話しすること

- オープンサイエンスに関連した物性研究所での取り組み
- オープンソースソフトウェアに関する取り組みの紹介
  - ✓ ソフトウェアを見つける
  - ✓ ソフトウェアを試す
  - ✓ ソフトウェアを開発する
- オープンデータに関する取り組みの紹介
- 人材育成に関連した取り組みの紹介

# 物性研究所でのソフトウェア・データに関連した取組

2013年～

ポータルサイトの運用

**MateriApps**

物質科学シミュレーション  
ポータルサイト

アプリを探す/試す

2010

2015年～

ソフトウェア開発・高度化

**PASUMS**

Project for advancement of  
software usability in materials science

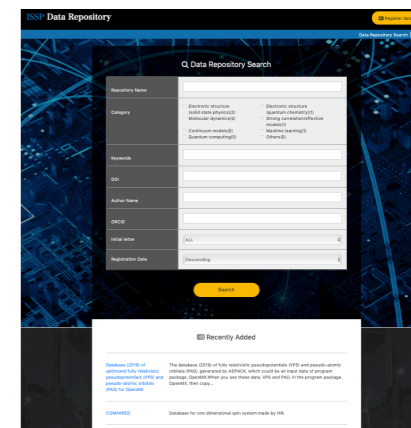
大型計算機向けに使いやすい  
ソフトウェアの整備・開発

<<スパコンの活用>>  
高精度データ解析  
ハイスループット計算

2015

2021年～

データリポジトリ運用



物質科学データ向け  
リポジトリの開発

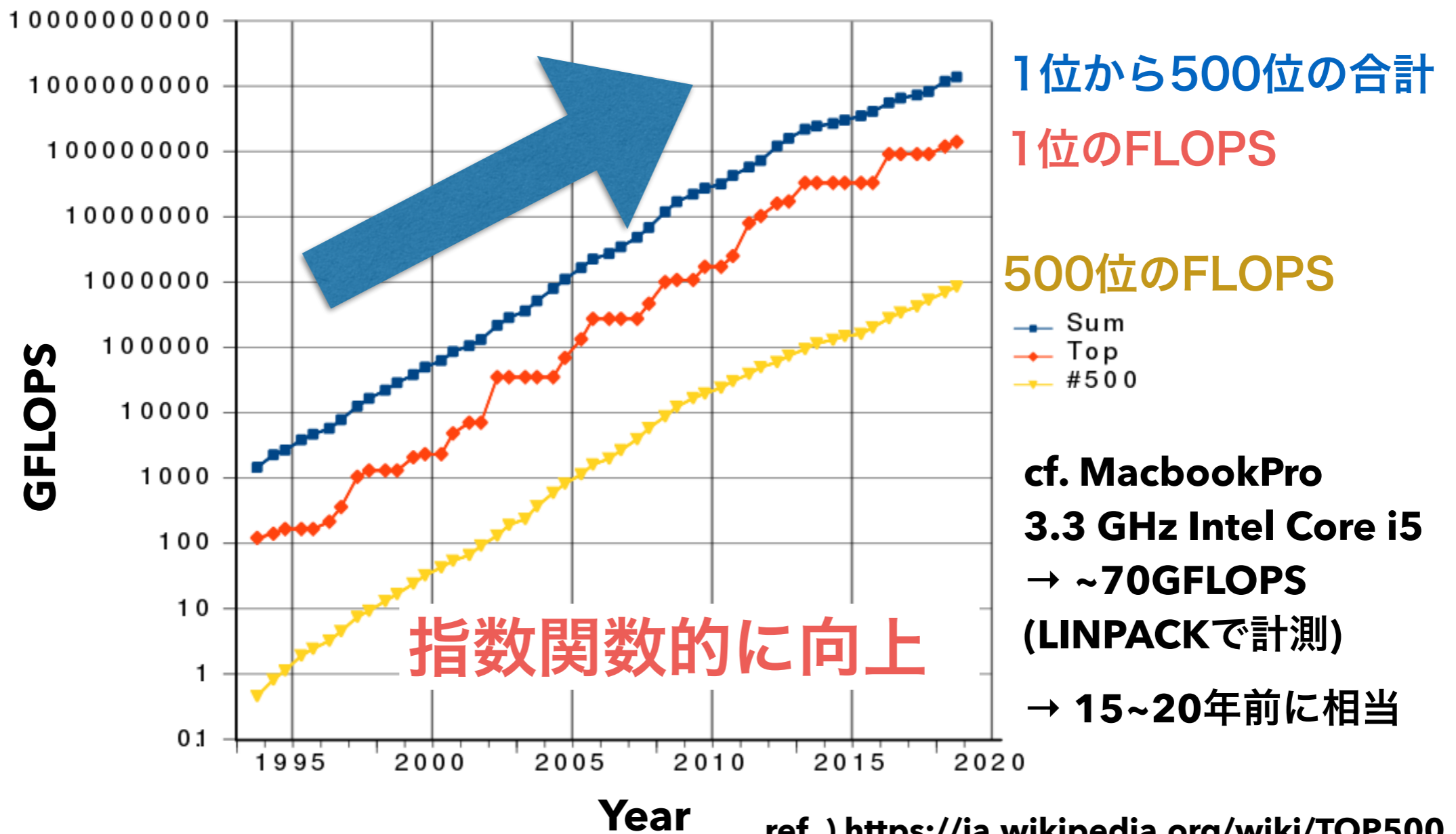
物質科学データ  
の蓄積・活用

2020



# スーパーコンピュータの性能

Top500にランキングされているスーパーコンピュータの性能



ref. ) <https://ja.wikipedia.org/wiki/TOP500>

# スーパーコンピュータ全国共同利用

## <<スパコンの導入>>

1995年 初代スパコン  
2000年 第二期のシステム  
...

(5年サイクルで更新)

2015年 sekirei (システムB)

2018年 enaga (システムC)

2020年 ohtaka (システムB)

## 安定的な計算環境の提供

2000年以降、概ね**95%**を越える稼働率

500人規模の全国のアカデミック利用者

関連論文は毎年500編前後の出版

物性研スパコン パンフレット(2016年2月)より

<https://www.issp.u-tokyo.ac.jp/supercom/about-us>



最先端の大規模計算機を駆使した  
先端的な研究

その一方で...

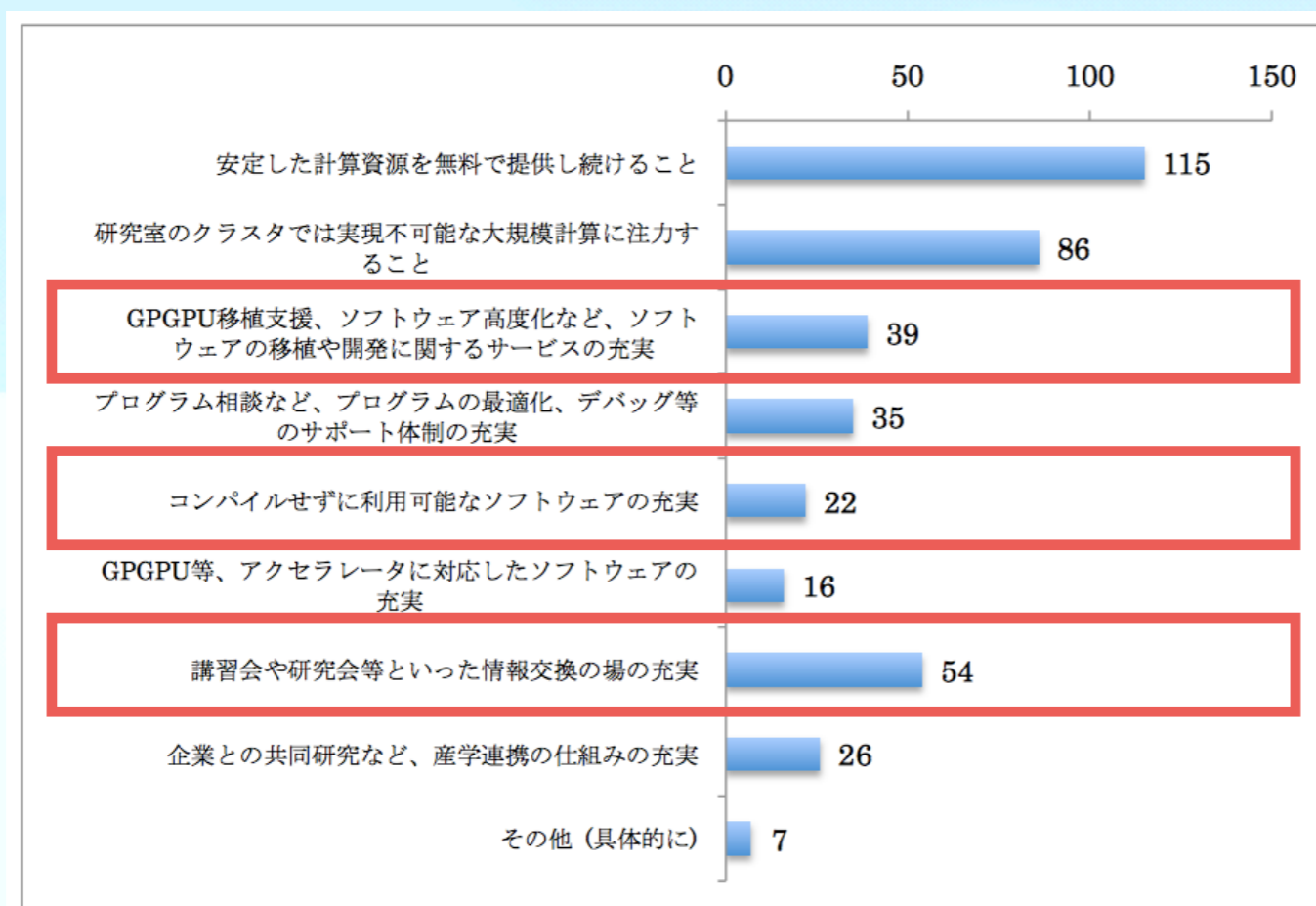
ハードウェアの進化に伴った「並列計算の高度化・複雑化」

→ スパコンの利用にあたり、利用者に高いハードルを与えてしまう可能性。

# ウェブアンケート結果

## • 今後の物性研スパコンに望むことは何ですか？

物性研スパコンシステムに関するウェブアンケート調査(2015年度) より



- GPGPU移植支援、ソフトウェア高度化など、ソフトウェアの移植や開発に関するサービスの充実 (39/138)
- コンパイルせずに利用可能なソフトウェアの充実 (22/138)
- 講習会や研究会等といった情報交換の場の充実 (54/138)

# ソフトウェア開発高度化プロジェクト (1)

よりユーザーフレンドリーな  
スーパーコンピュータの提供を目指して

**PASUMS**

Project for advancement of  
software usability in materials science

毎年12月に公募 (2015年～)

主要メンバー (2024/3/15現在)

- 1,2件の課題を選定

- OSSとして年度末に公開

年間スケジュールイメージ



吉見一慶  
(2015/4～)



本山裕一  
(2018/4～)



青山龍美  
(2022/5～)

April



Kick-off meeting

May ~ Feb.



GitHub

Software development



Regular meeting/month

March ~



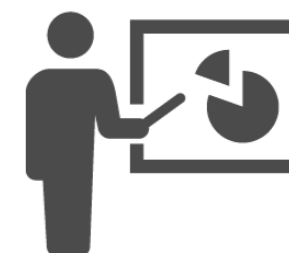
Website

MateriApps

Portal site



Preinstall



Lecture



# ソフトウェア開発高度化プロジェクト (2)

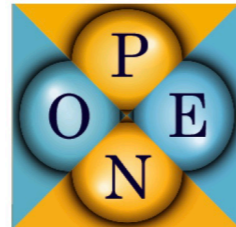
## 高度化したソフトウェアの一覧

### First principles calculations and related applications



abICS

abICS



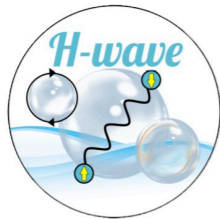
OpenMX



RESPACK

ESM-RISM

### Quantum lattice model solvers



H-wave



HΦ



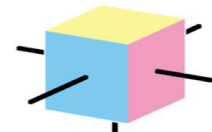
mVMC



DSQSS



DCore



TeNeS

### Mathematical libraries and others



PHYSBO



Kw



2dmat



MateriApps

開発したソフトウェアに関する論文も執筆・出版 URL: <http://www.pasums.issp.u-tokyo.ac.jp>  
物性分野の研究者と幅広い連携が可能

# 開発したソフトウェアに関するニュース

## ソフトウェア開発・高度化ホームページ

URL: <https://www.pasums.issp.u-tokyo.ac.jp>

物性研スパコンホームページの  
「ソフトウェア開発・高度化プロジェクト」  
から遷移可能。

### (1) プロジェクト活動に関するニュース

#### News

- 公募関連**  
2020/11/9/  
物性研2021年度ソフトウェア開発・高度化プロジェクト公募のご案内
- 論文**  
2020/07/22/  
DSQSSの論文がarXivに投稿されました。
- 論文**  
2020/07/3/  
DCoreの論文がarXivに投稿されました。
- 其他のお知らせ**  
2020/04/20/  
TeNeSを物性研スパコンにインストールしました。
- 其他のお知らせ**  
2020/04/1/  
2020年度の採択課題が決定しました。
- 其他のお知らせ**  
2020/02/17/  
TeNeSのホームページを公開しました。また、MateriAppsページへの掲載.....

プロジェクト公募情報、該当年度課題の情報、  
スパコンへのプリインストール情報 etc.

### (2) 開発したソフトウェアに関するニュース

#### 高度化ソフトウェアの最新NEWS

- PHYSBO** **リリース情報**  
2020/12/11/  
ver. 0.2.0が公開されました。
- MateriApps Installer** **リリース情報**  
2020/12/9/  
ver.0.1.0がリリースされました。
- PHYSBO** **リリース情報**  
2020/11/18/  
ver. 0.1.0が公開されました。
- TeNeS** **リリース情報**  
2020/11/9/  
ver.1.1.1をリリースしました。

バージョンアップ情報、  
ソフトウェア講習会の開催 (CCMS) etc.

# 本日本話しすること

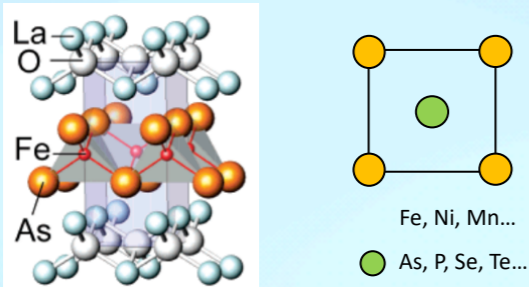
- オープンサイエンスに関連した物性研究所での取り組み
- オープンソースソフトウェアに関する取り組みの紹介
  - ✓ ソフトウェアを見つける
  - ✓ ソフトウェアを試す
  - ✓ ソフトウェアを開発する
    - ✓ 事例の紹介
- オープンデータに関する取り組みの紹介
- 人材育成に関連した取り組みの紹介



# 研究事例紹介～鉄系超伝導体の記述子探索

物性研スパコン上にシームレスな高精度解析を行う環境が実現！

## 結晶構造



ref.) Y. Kamihara, et al.,  
J. Am. Chem. Soc. **130**, 3296 (2008)

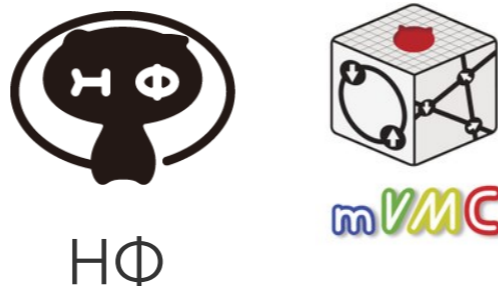
## 第一原理計算



電子の絡み合いを  
きちんと取り入れた  
モデル(有効モデル)の導出



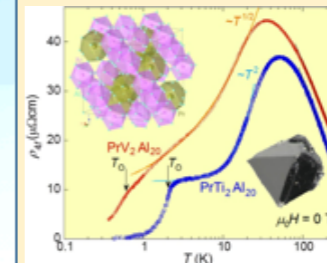
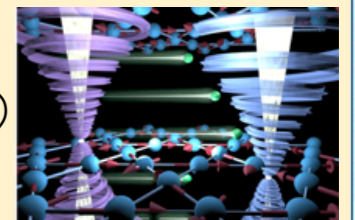
## 有効モデルの解析



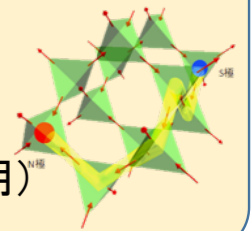
高精度な計算が可能！

## 新しい物理概念に基づく物質設計

磁気メモリ  
(トポロジカル物質)

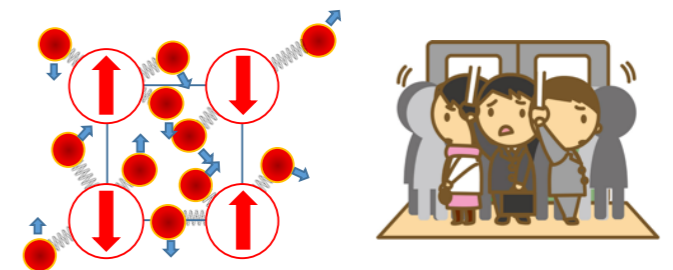


スピントロニクス  
(磁気モノポール利用)<sup>TM</sup>



物性研HPより

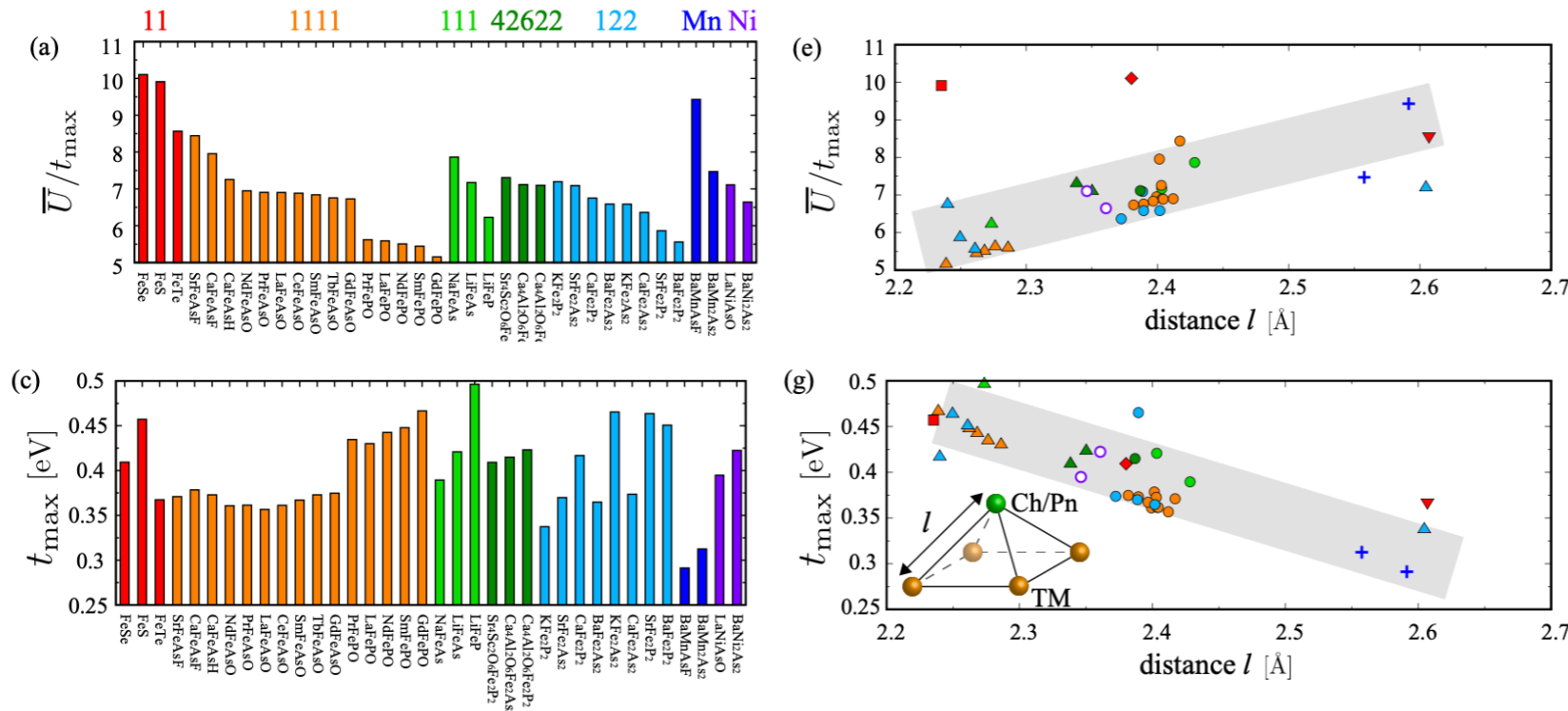
電子がひしめき合う強相関電子系ゆえの機能性





# 研究事例紹介～鉄系超伝導体の記述子探索

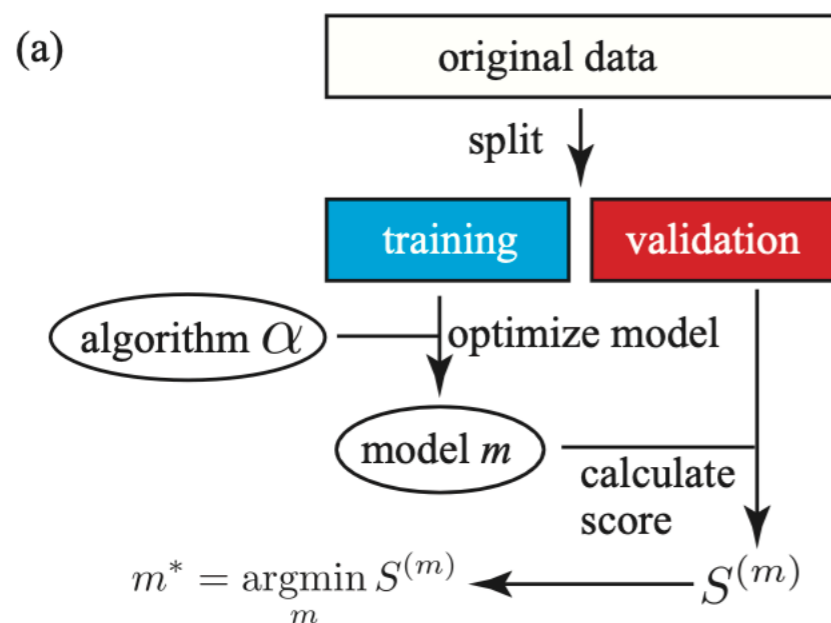
## 1. 29物質について第一原理計算→有効模型



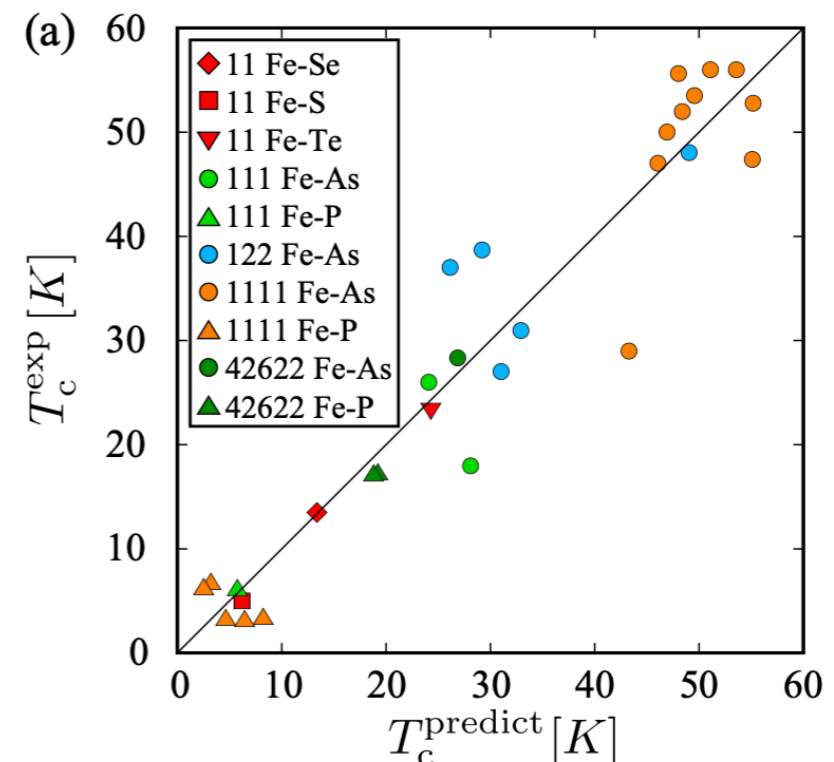
記述子を36個作成

平均値、最低値、最大値、  
対角成分、非対角成分 etc.

## 2. 記述子の組み合わせを交差検証で決定



## 3. 超伝導転移温度の推定



# ソフトウェア開発・高度化プロジェクト

2020年度課題：二次元物質構造解析むけ実験データ解析の高度化

2021年度課題：二次元物質構造解析むけ実験データ解析

の統合プラットフォーム

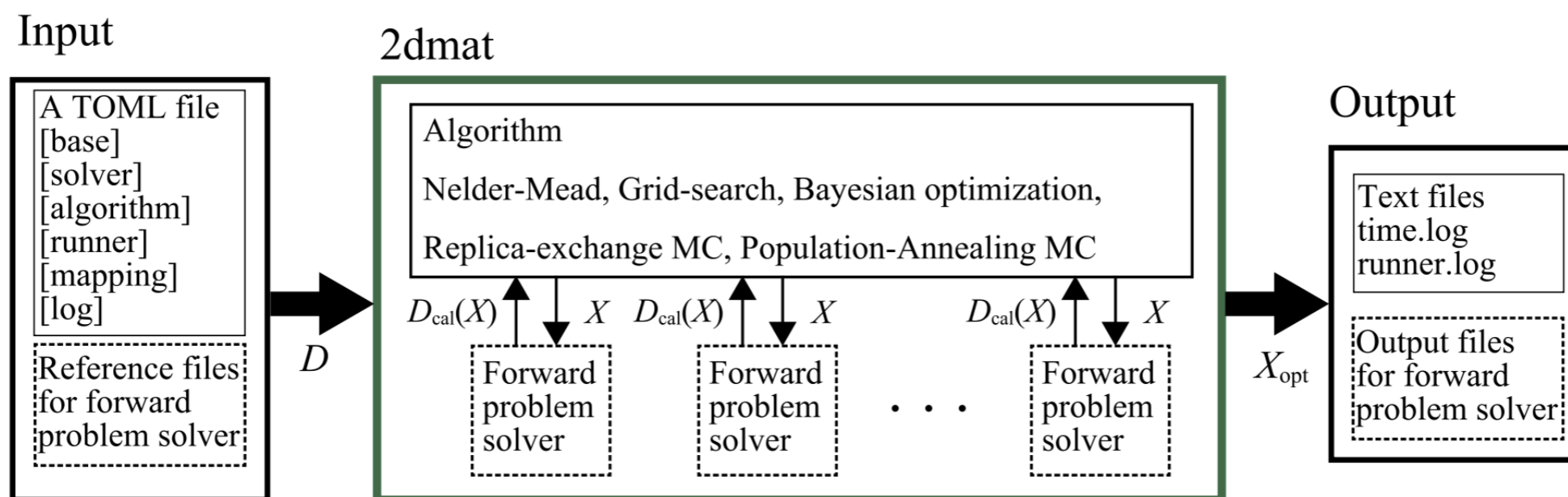


提案者：星 健夫 (提案当時 鳥取大学 大学院工学研究科)

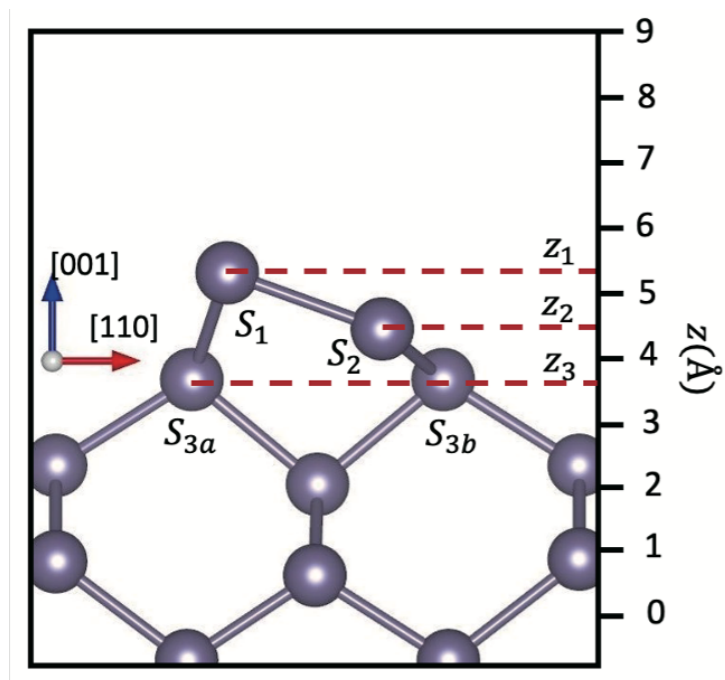
<https://www.pasums.issp.u-tokyo.ac.jp/2dmat/>

順問題ソルバーに対して探索アルゴリズムを適用し  
最適解を探すためのフレームワーク。

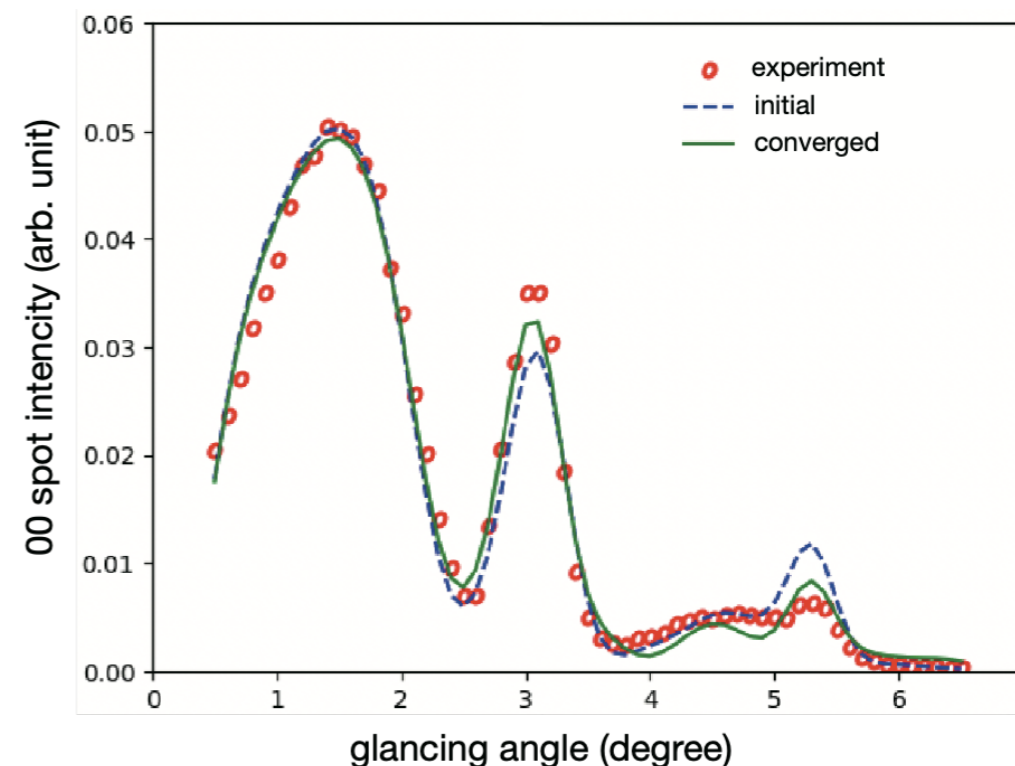
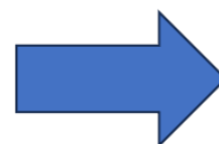
2DMAT



# 研究事例紹介~モンテカルロ法を活用したパラメータ推定



TRHEPD  
での実験



Ge(001)-c(4x2)の表面付近の構造

一波条件(z座標決定むけ計測条件)の解析→ TRHEPDの実験結果(赤丸)とシミュレーション(\*)を組み合わせ、尤もらしいz座標のペアを決める必要性がある。

実験値とシミュレーション値とのずれの2乗を最小にする

$$F = F(X) \equiv d(D_{\text{cal}}(X), D), \quad d(u, v) \equiv \left( \sum_{i=1}^m (u_i - v_i)^2 \right)^{1/2},$$

計算データ, 実験データ

真の解を以下に設定  
( $z_1, z_2, z_3$ ) = (5.231, 4.731, 3.596)

(\*) sim-trhepd-rheed

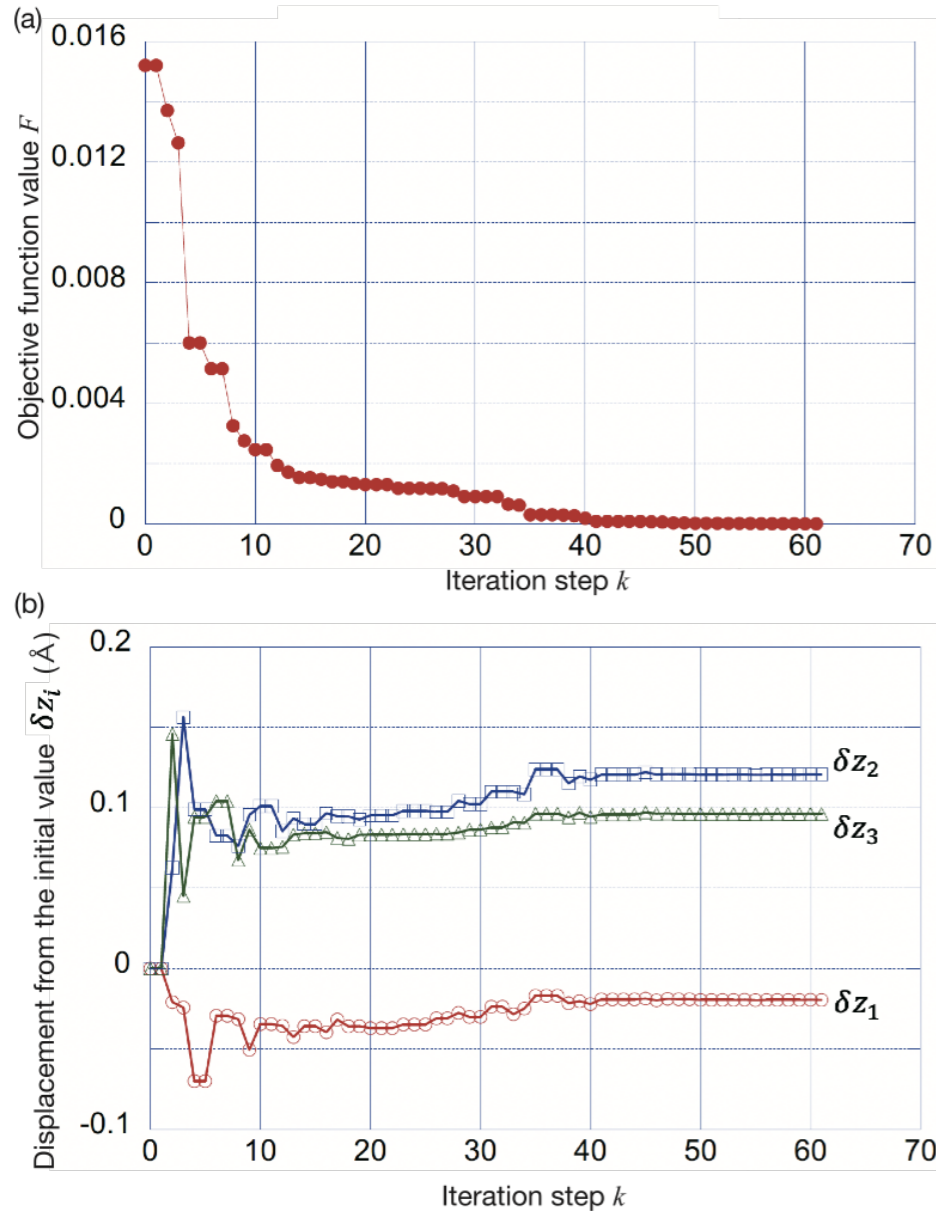
URL: <https://github.com/sim-trhepd-rheed/sim-trhepd-rheed>

ref.) T. Hanada et al., Computer Physics Communications 277, 108371 (2022).



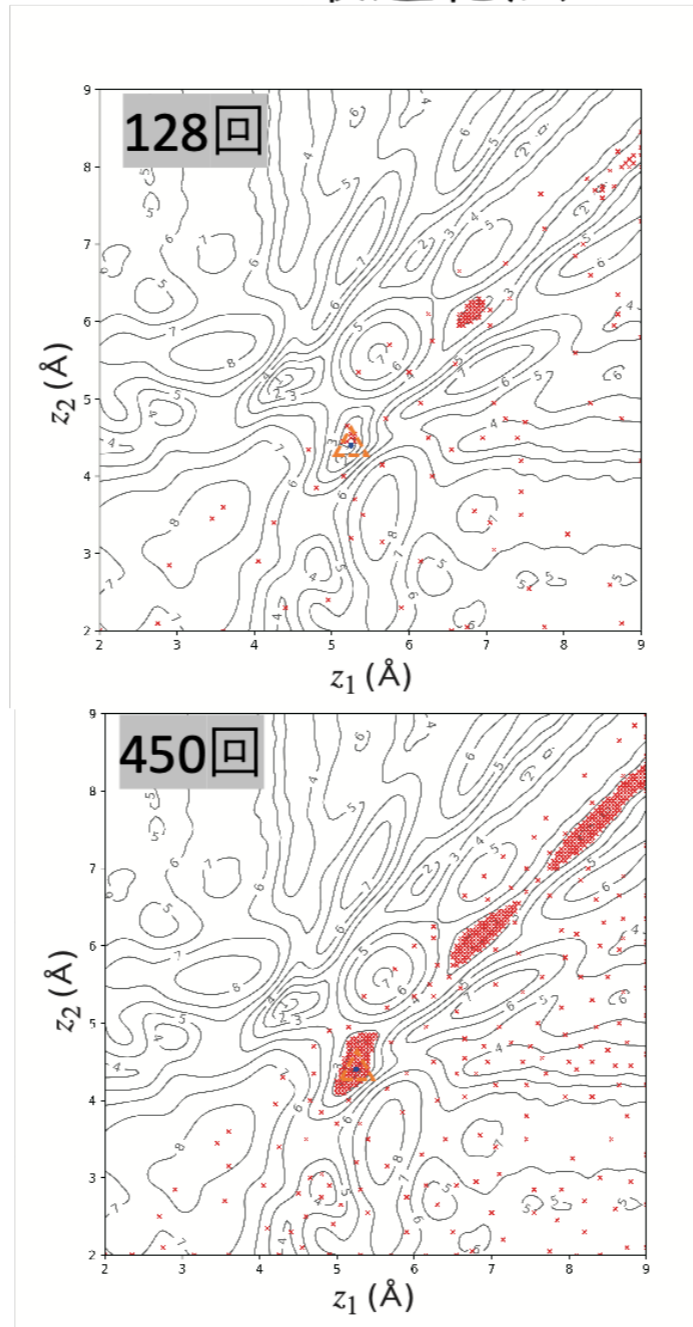
# 研究事例紹介~モンテカルロ法を活用したパラメータ推定

## 1. Nelder-Mead法



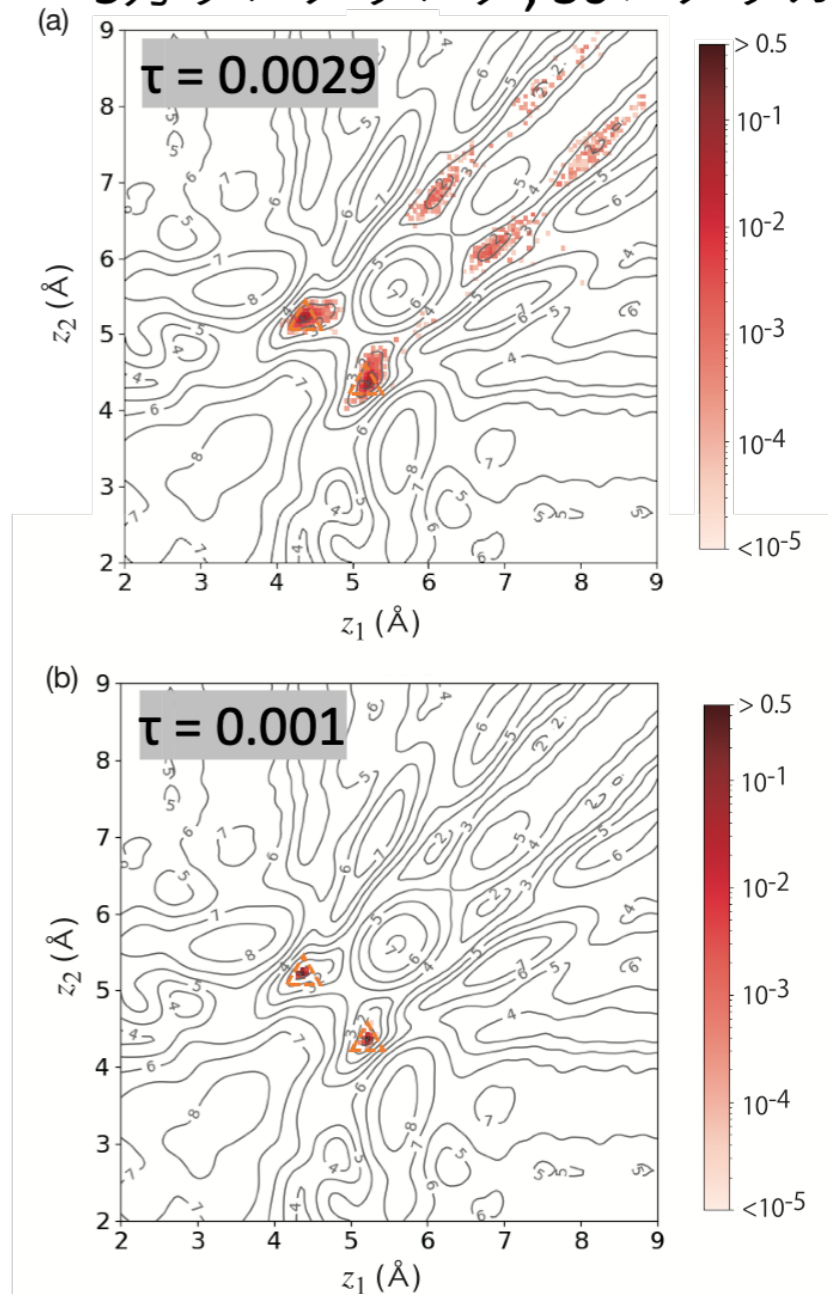
初期解  
 $(z_1, z_2, z_3) = (5.25, 4.25, 3.5)$   
 $\rightarrow z_1, z_3$ についてはほぼ正解に辿り着く

## 2. ベイズ最適化法



450回行くと正解付近  
 を含む怪しいところ  
 まで探索済

## 3. レプリカ交換モンテカルロ法 5万サンプリング, 36レプリカ



温度 $\tau$ が大きい箇所では正解  
 以外の箇所、小さい箇所では  
 正解付近の箇所が探索済



# ソフトウェア開発・高度化プロジェクト

2019年度課題：拡張アンサンブル法・第一原理計算結合フレームワークの高度化

2022年度課題：第一原理計算基統計熱力学計算フレームワークの高度化

提案者：笠松 秀輔 (山形大学学術研究院 (理学部主担当))

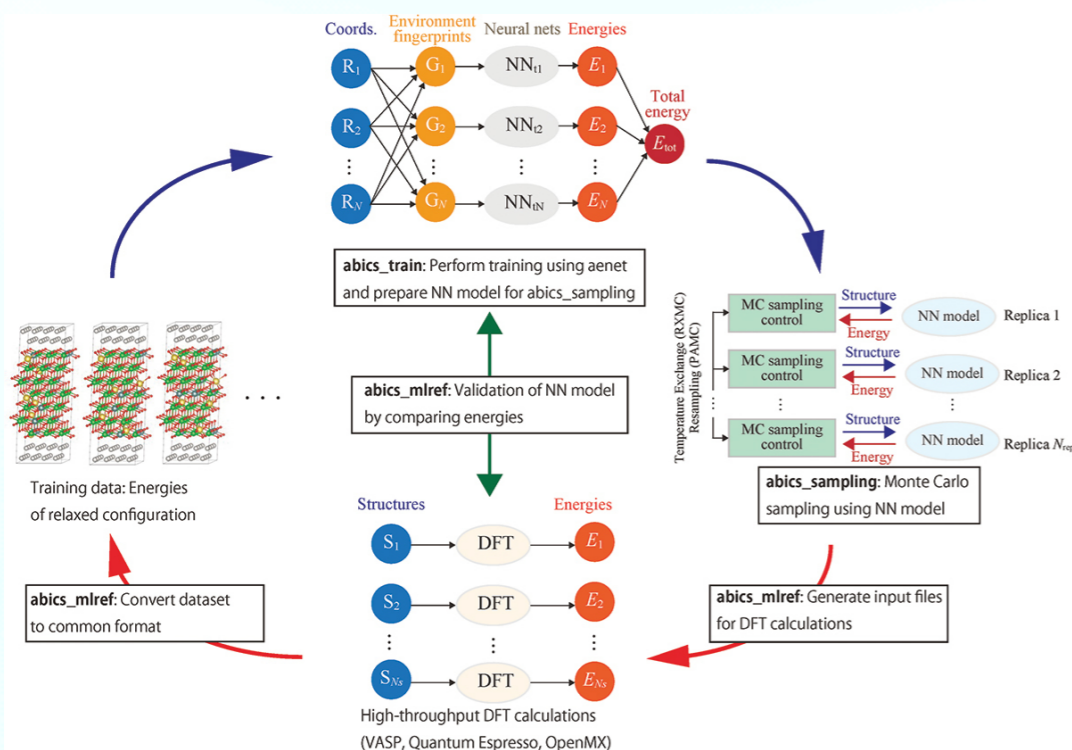
abICS



<https://www.pasums.issp.u-tokyo.ac.jp/abics>

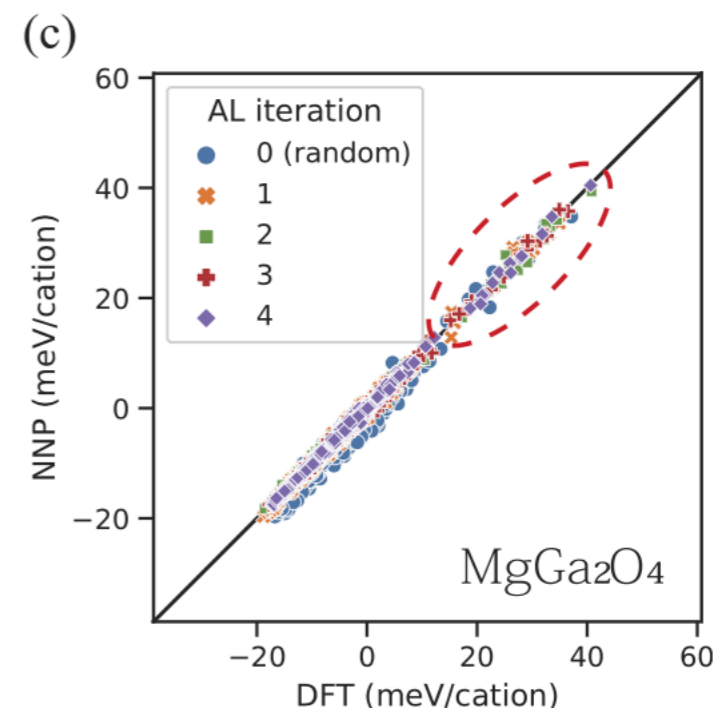
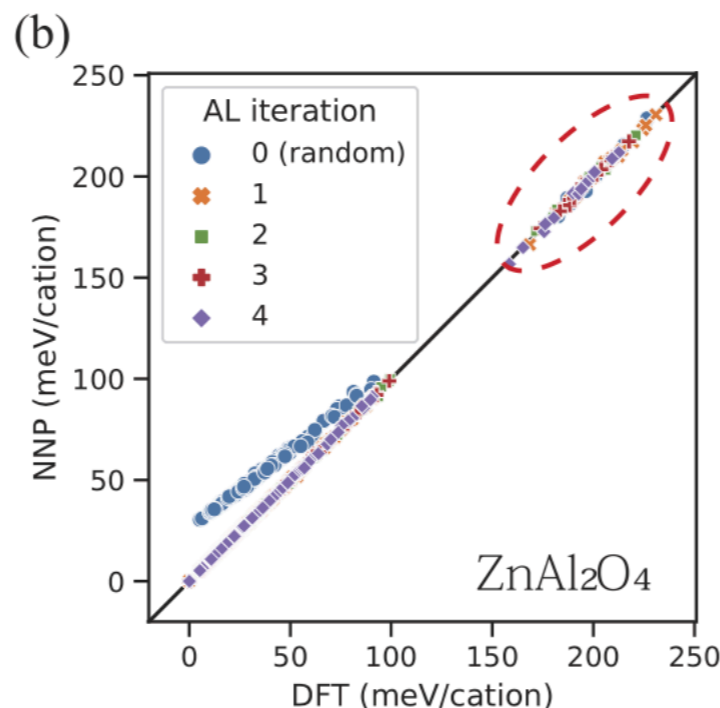
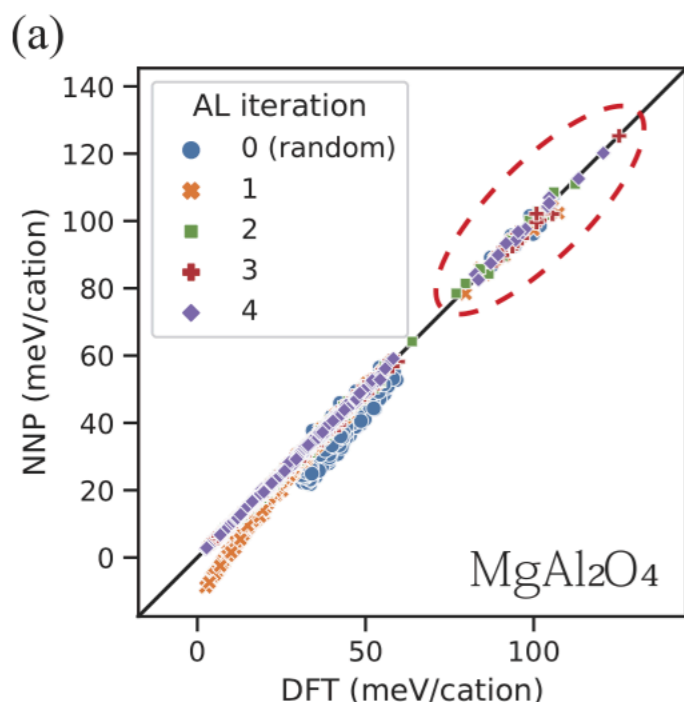
合金の配置をモンテカルロ法を用いて決定する。  
エネルギー計算についてはニューラルネットワークaenetを使用。  
<http://ann.atomistic.net/>

1. 原子位置を出力  
(初めはrandom, 2回目以降はMC計算の結果)
2. 第一原理計算でエネルギー計算
3. aenetを用いてニューラルネットワークを作成  
(原子配置からエネルギーを予測)
4. ニューラルネットワークを利用してMC計算し、  
最適な配置を予測  
(途中のステップで原子配置をいくつか出力)

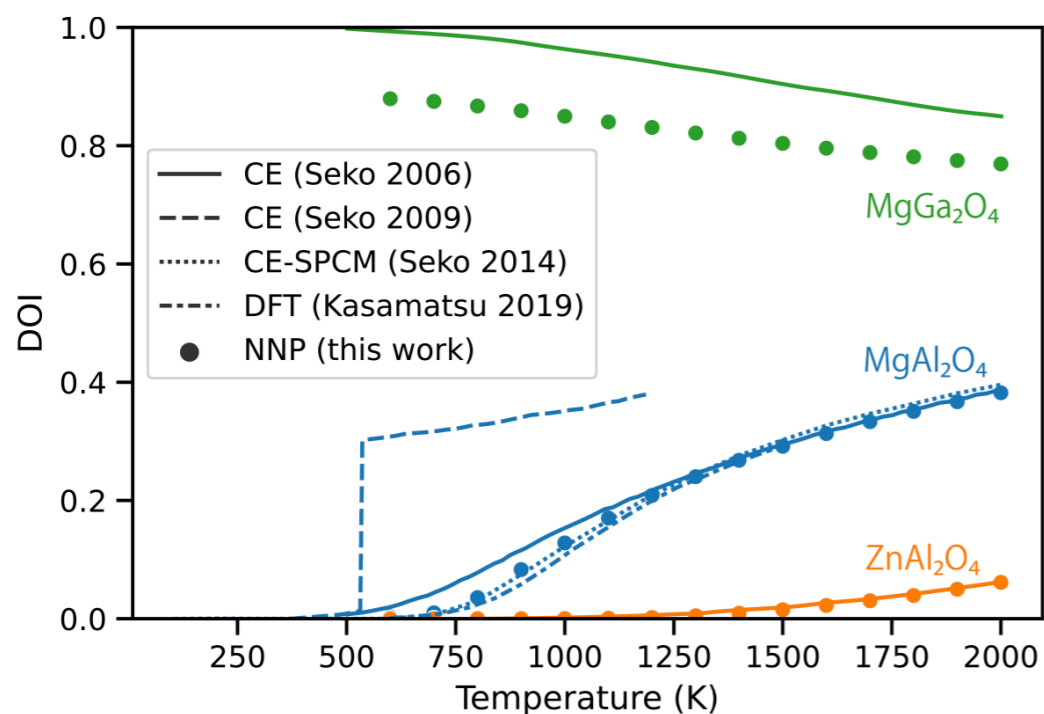


# 研究事例紹介~深層学習+モンテカルロ法を活用した構造推定

## 能動学習の計算結果



## 反転度(DOI)の結果



- 能動学習(AL) → 十分に機能している
  - AL iterationが進むごとにエネルギー差は小さくなる。
  - DFTで得られたDOIとaenetにより得られたDOIはほぼ一致している。

# 網羅計算をしやすいとするツールの作成

2023年度課題：ハイスループット計算による第一原理計算データベース作成ツールの開発

提案者：吉見 一慶 (東大物性研)、

協力者：福島鉄也(AIST)、笠松秀輔 (山形大)、是常隆(東北大)、福田将大(ISSP)、井戸康太(ISSP)

物性研スパコン上で、気軽に第一原理計算を含んだ網羅計算の環境構築。

機械学習をするためのデータベースを作成できるようにする。

結晶構造(cif)ファイルから  
第一原理計算用入力ファイル生成  
cif2x (x=qe, vasp, openmx, kkr)

物性研スパコン上で網羅計算  
moller

その他ソルバーにも対応可

大規模計算させたいソルバーの  
入力ファイル生成

各フォルダに計算結果が格納

ハンズオン講習会・研究会・ワークショップの開催  
→ 利活用可能なデータベースの作成へ

# 本日本話しすること

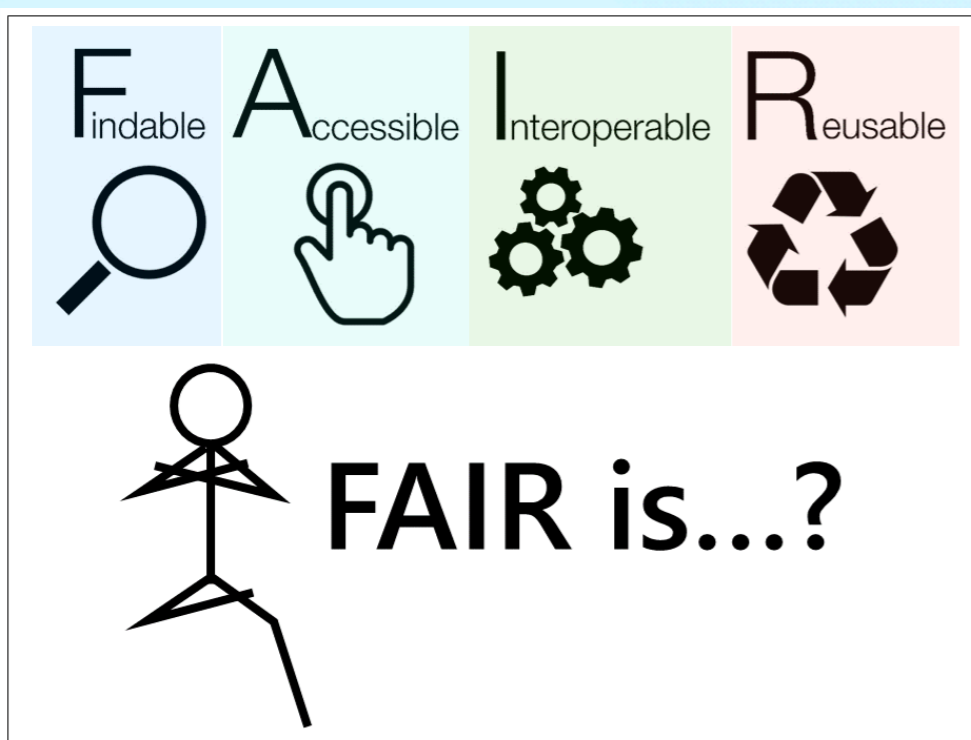
- オープンサイエンスに関連した物性研究所での取り組み
- オープンソフトウェアに関する取り組みの紹介
- オープンデータに関する取り組みの紹介
  - ✓ 物質科学分野におけるオープンデータに関する動向・関連データベースの紹介
  - ✓ データリポジトリ事業に関する紹介
- 人材育成に関連した取り組みの紹介



# オープンデータに関する世界的な動き

## 2013 G8 サミット：オープンデータ憲章

### FAIR原則



FAIR guiding principles for data resources  
© 2016 SangyaPundir licensed under  
Creative Commons

#### 1. Findable (発見可能)

データを簡単に検索できるようにするため、一意の識別子を割り当てる。

#### 2. Accessible (アクセス可能)

誰もが障害なくデータにアクセスできるようにするため、オンラインで公開し標準プロトコルを使用する。

#### 3. Interoperable (相互運用可能)

データを他のデータセットやアプリケーションと組み合わせやすくするため、共通の標準とフォーマットを使用する。

#### 4. Reusable (再利用可能)

データの将来的な再利用を容易にするため、詳細なメタデータと使用条件を提供する。

# オープンアクセスに向けたグローバルな動き

## ・アメリカのオープンアクセス政策

- ・ 2022年8月25日に、連邦資金による研究成果を即時公開する新政策を発表[1]。
  - ・ 納税者が支援する研究成果が、公開された直後からアメリカ公衆に無償で提供されるようにする。公開遅延はなく、全ての研究成果が出版後すぐに利用できるようになる。
  - ・ 全ての連邦機関は、2025年12月31日までにこの新しい公開アクセス政策を完全に実施することが求められている

## ・ユネスコ

- ・ グローバルオープンアクセスポータル(GOAP)[2]により、各国のオープンアクセス関連情報を提供し、グローバルな取り組みをサポート

[1] <https://www.whitehouse.gov/ostp/news-updates/2022/08/25/ostp-issues-guidance-to-make-federally-funded-research-freely-available-without-delay/>

[2] <https://goap.info/>

# 海外でのデータリポジトリ

## データの置き場はどうすれば良いのか？

### ・ 汎用的なデータリポジトリ

- Zenodo[1]: あらゆる学術分野のデータを扱い、研究成果の引用とアクセスをDOIを通じて提供。
- Figshare[2]: 視覚的なデータ共有に特化し、直感的なインターフェースを提供。

### ・ 物質科学に特化したデータリポジトリ

- Materials Project[3], Materials Data Facility[4], AFLOW[5], JARVIS[6], NOMAD[7] etc.

[1] <https://zenodo.org/>, [2] <https://figshare.com/>

[3] <https://next-gen.materialsproject.org> , [4] <https://materialsdatafacility.org/>

[5] <https://www.aflowlib.org/>, [6] <https://jarvis.nist.gov>,

[7] <https://nomad-lab.eu/nomad-lab>



# 日本でのオープンアクセスに向けた取り組み 内閣府によるガイドラインの作成

- ・ 2019年3月 研究データリポジトリ整備・運用ガイドライン[1]
- ・ 2019年4月 国立研究開発法人における データポリシー策定のためのガイドライン [2]
- ・ 2021年3月 第6期科学技術・イノベーション基本計画 [3]
- ・ 2024年2月 学術論文等の即時オープンアクセスの実現に向けた基本方針[4]

## (1) 公的資金による学術論文等の即時オープンアクセスの実施

- ・ 公的資金<sup>1</sup>のうち2025年度から新たに公募を行う即時オープンアクセスの対象となる競争的研究費を受給する者（法人を含む）に対し、該当する競争的研究費による学術論文及び根拠データの学術雑誌への掲載後、即時に機関リポジトリ等の情報基盤への掲載を義務づける<sup>2</sup>。
- ・ 即時オープンアクセスの対象となる競争的研究費制度は、学術論文を主たる成果とするものとし、関係府省が定める。
- ・ 即時オープンアクセスの対象は、査読付き学術論文（電子ジャーナルに掲載された査読済みの研究論文（著者最終稿を含む））及び根拠データ（掲載電子ジャーナルの執筆要領、出版規程等において、透明性や再現性確保の観点から必要とされ、公表が求められる研究データ）とする。

[1] <https://www8.cao.go.jp/cstp/tyousakai/kokusaiopen/guideline.pdf>

[2] <https://www8.cao.go.jp/cstp/stsonota/datapolicy/datapolicy.html>


[3] [https://www.mext.go.jp/content/20210608-mxt\\_jyohoka01-000015787\\_06.pdf](https://www.mext.go.jp/content/20210608-mxt_jyohoka01-000015787_06.pdf)

[4] [https://www8.cao.go.jp/cstp/oa\\_240216.pdf](https://www8.cao.go.jp/cstp/oa_240216.pdf)



# 日本で進められつつあるデータ管理基盤の構築事業

## GakuNin RDM (研究データ管理基盤)

研究データ管理(RDM)サービス  
GakuNin RDMの概要  <https://rdm.nii.ac.jp/>


**研究データ共有・管理機能**

学認による認証とグループ管理でデータ共有

**NII研究データ基盤や外部ツールとの連携**


研究ツール  
リポジトリ  
プライベートクラウド  
パブリッククラウド

**研究証跡の保存機能**

時刻認証事業者のタイムスタンプで  
ある時刻のファイルの存在を証明 

**機関利用のシステム管理者のための管理機能**

組織内のRDMサービスのカスタマイズ <https://rcos.nii.ac.jp/service/rdm/> より抜粋



ref. 物性アプリオープンフォーラムでも紹介 (以下のリンクに資料あり。「PASUMS gakunin」でGoogle検索)

<https://www.pasums.issp.u-tokyo.ac.jp/teams/seminar/1352>

# コミュニティベースのデータリポジトリの必要性

## • 背景と動機

- 学術的柔軟性や外交的・戦略的制約などの必要性から、既存の国際的なデータリポジトリサービス(Zenodo, figshare等)に頼るだけではなく、より適応性の高いアプローチが求められる。

物性科学コミュニティを対象とした、共有・利用が容易なデータリポジトリの構築はできないか？

## • コミュニティベースのデータベースの利点

- コミュニティ内での研究データの共有と再利用の促進。
- 研究成果の迅速な普及とコラボレーションの強化。

→ コミュニティの合意形成と技術的な実現性の検討が可能に！

# 本日本話しすること

- 物質科学分野におけるオープンソフトウェアに関する取り組みの紹介
- 物質科学分野におけるオープンデータに関する取り組みの紹介
- オープンデータに関する動向・関連データベースの紹介
- データリポジトリ事業に関する紹介
- 人材育成に関連した取り組みの紹介

# 物性研究所でのソフトウェア・データに関連した取組

2013年～

ポータルサイトの運用

**MateriApps**

物質科学シミュレーション  
ポータルサイト

アプリを探す/試す

2010

2015年～

ソフトウェア開発・高度化

**PASUMS**

Project for advancement of  
software usability in materials science

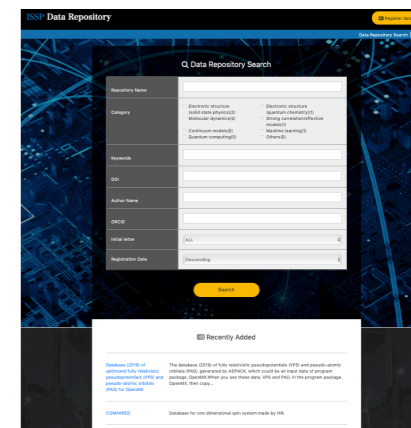
大型計算機向けに使いやすい  
ソフトウェアの整備・開発

<<スパコンの活用>>  
高精度データ解析  
ハイスループット計算

2015

2021年～

データリポジトリ運用



物質科学データ向け  
リポジトリの開発

物質科学データ  
の蓄積・活用

2020

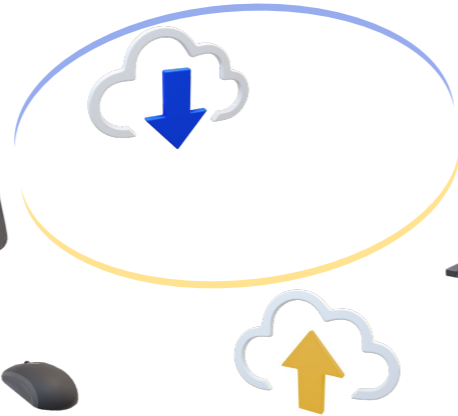


# データリポジトリ事業 (2021年～)

ポータルサイトで  
データリポジトリの検索 🔍



- ✓ 論文データや研究・計算データをアップロード
- ✓ ダウンロードしたデータの利活用
- ✓ GitLabによりGUIとCUIの両方に対応  
ex.) 事前にアップロードしてarXiv・投稿論文にURLを記載



- ☆フロントエンドサーバ
  - CPU 2cores
  - Memory 6GB
  - Storage 1TB
- ☆GitLabサーバ
  - CPU 4cores
  - Memory 8GB
- ☆データ領域 2TB  
(将来的に拡張予定)

## 運営チーム



川島直輝



吉見一慶



矢田裕行



井戸康太



福田将大



本山裕一

写真なし：福田毅哉、荒木繁行

[https://mdcl.issp.u-tokyo.ac.jp/scc/guide/application\\_proposal/issp-datarepo](https://mdcl.issp.u-tokyo.ac.jp/scc/guide/application_proposal/issp-datarepo)

**1st step : 論文データの蓄積/公表 (再現性のある研究)**

**2nd step: 研究データベース作成/アプリと連携した高速な解析**

**3rd step: データベースを用いた機械学習による物質探索/物性予測 etc.**

→ 情報基盤センター, NIMSなど各種機関と連携し有用なデータベースを共有し、ツールの機能向上・拡張などを行い、より幅広いユーザ層への展開を目指す。

# データリポジトリ事業概要

ISSP Data Repository

Register dataset

Data Repository Search | About

Q Data Repository Search

Repository Name

Category

- Electronic structure (solid state physics)(10)
- Molecular dynamics(0)
- Continuum models(0)
- Quantum computing(1)
- Electronic structure (quantum chemistry)(1)
- Strong correlation/effective models(7)
- Machine learning(4)
- Others(1)

Keywords

DOI

Author Name

ORCID

Initial letter

Registration Date

Search

Recently Added

Ab initio Hamiltonians for molecular solids alpha-ET and alpha-BETS<sub>2/3</sub> We systematically derive low-energy effective Hamiltonians for molecular solids alpha-ET<sub>2/3</sub> and alpha-BETS<sub>2/3</sub> with available room- and low-

- 2021年度から運用開始
- 登録件数 (2024/3/13現在)
  - 34件 (公開28件、非公開6件)
    - データベース：6件
    - 論文データ：17件
    - ソフトウェア関連データ：11件
- データ検索
  - DOI
  - Category
  - keywords
  - 氏名
  - 連絡先 etc.
- 最新ニュース(新規登録情報)なども掲載

# データベース事例の紹介 (1)

## 論文データのアップロード

- プレプリントサーバ (arXiv)に論文投稿
  - データも同時に登録
    - 登録する情報
      - ソフトウェアの入力ファイル
      - 入力ファイル生成の手順・ツール
      - シミュレーション結果・出力ファイル
      - 出力ファイルを加工した手順・ツール

[Submitted on 22 Sep 2021 (v1), last revised 6 Jan 2022 (this version, v3)]

### *Ab initio* derivation and exact-diagonalization analysis of low-energy effective Hamiltonians for $\beta'$ - $X[\text{Pd}(\text{dmit})_2]_2$

Kazuyoshi Yoshimi, Takao Tsumuraya, Takahiro Misawa

The molecular solids  $\beta'$ - $X[\text{Pd}(\text{dmit})_2]_2$  (where  $X$  represents a cation) are typical compounds whose electronic structures are described by single-orbital Hubbard-type Hamiltonians with geometrical frustration. Using the *ab initio* downfolding method, we derive the low-energy effective Hamiltonians for  $\beta'$ - $X[\text{Pd}(\text{dmit})_2]_2$  with available room- and low-temperature structures. We find that the amplitudes of the Coulomb interactions and the anisotropy of the hopping parameters in the effective Hamiltonians are sensitive to the changes in the lattice constants induced by lowering the temperature. The obtained effective Hamiltonians are analyzed using the exact diagonalization method with the boundary-condition average. We find that a significant reduction of the antiferromagnetic ordered moment occurs in the effective Hamiltonian of  $\beta'$ - $\text{EtMe}_3\text{Sb}[\text{Pd}(\text{dmit})_2]_2$  with the low-temperature structure. The reduction is consistent with the quantum spin liquid behavior observed in experiments. The comprehensive derivations of the effective Hamiltonians and exact-diagonalization analyses of them will clarify the microscopic origins of the exotic quantum states of matter found in  $\beta'$ - $X[\text{Pd}(\text{dmit})_2]_2$  such as the quantum spin liquid behavior.

Comments: 12 pages, 6 figures, 1 table  
Subjects: Strongly Correlated Electrons (cond-mat.str-el); Materials Science (cond-mat.mtrl-sci)  
Cite as: arXiv:2109.10542 [cond-mat.str-el]  
(or arXiv:2109.10542v3 [cond-mat.str-el] for this version)  
Journal reference: Phys. Rev. Research 3, 043224 (2021)  
Related DOI: <https://doi.org/10.1103/PhysRevResearch.3.043224>

#### Submission history

From: Kazuyoshi Yoshimi [view email]  
[v1] Wed, 22 Sep 2021 06:51:59 UTC (2,638 KB)  
[v2] Mon, 4 Oct 2021 05:36:46 UTC (2,638 KB)  
[v3] Thu, 6 Jan 2022 01:32:43 UTC (2,639 KB)

Bibliographic Tools | Code & Data | Related Papers | About arXivLabs

Code and Data Associated with this Article

arXiv Links to Code & Data (What is Links to Code & Data?)

Official Code

[https://isspns-gitlab.issp.u-tokyo.ac.jp/k-yoshimi/physrevresearch\\_vol3\\_page043224\\_year2021](https://isspns-gitlab.issp.u-tokyo.ac.jp/k-yoshimi/physrevresearch_vol3_page043224_year2021)

Community Code

Submit your implementations of this paper on [Papers With Code](#)

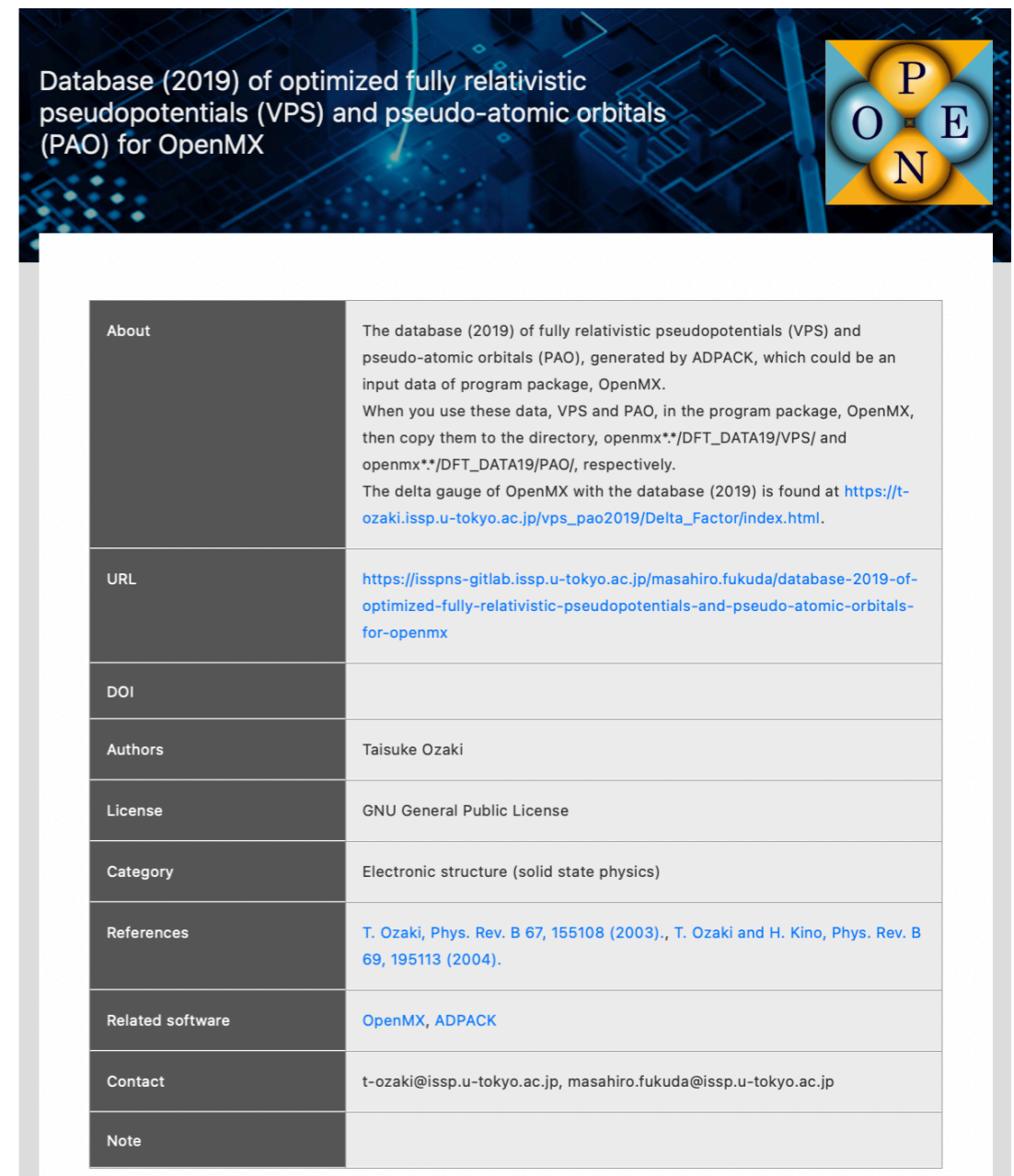
ref. ) <https://arxiv.org/abs/2109.10542>



# データベース事例の紹介 (2)

## Database (2019) of optimized fully relativistic pseudopotentials (VPS) and pseudo-atomic orbitals (PAO) for OpenMX

- 第一原理計算ソフトウェア  
OpenMXの擬ポテンシャル集  
(物性研究所 尾崎先生)
  - 格納データ
    - OpenMXでの計算に必要な擬ポテンシャルファイルを格納
  - 目的
    - OpenMXユーザーが擬ポテンシャルを一括ダウンロードすることが可能。



Database (2019) of optimized fully relativistic pseudopotentials (VPS) and pseudo-atomic orbitals (PAO) for OpenMX

About	The database (2019) of fully relativistic pseudopotentials (VPS) and pseudo-atomic orbitals (PAO), generated by ADPACK, which could be an input data of program package, OpenMX. When you use these data, VPS and PAO, in the program package, OpenMX, then copy them to the directory, <code>openmx*/DFT_DATA19/VPS/</code> and <code>openmx*/DFT_DATA19/PAO/</code> , respectively. The delta gauge of OpenMX with the database (2019) is found at <a href="https://t-ozaki.issp.u-tokyo.ac.jp/vps_pao2019/Delta_Factor/index.html">https://t-ozaki.issp.u-tokyo.ac.jp/vps_pao2019/Delta_Factor/index.html</a> .
URL	<a href="https://isspns-gitlab.issp.u-tokyo.ac.jp/masahiro.fukuda/database-2019-of-optimized-fully-relativistic-pseudopotentials-and-pseudo-atomic-orbitals-for-openmx">https://isspns-gitlab.issp.u-tokyo.ac.jp/masahiro.fukuda/database-2019-of-optimized-fully-relativistic-pseudopotentials-and-pseudo-atomic-orbitals-for-openmx</a>
DOI	
Authors	Taisuke Ozaki
License	GNU General Public License
Category	Electronic structure (solid state physics)
References	<a href="#">T. Ozaki, Phys. Rev. B 67, 155108 (2003).</a> , <a href="#">T. Ozaki and H. Kino, Phys. Rev. B 69, 195113 (2004).</a>
Related software	<a href="#">OpenMX</a> , <a href="#">ADPACK</a>
Contact	<a href="mailto:t-ozaki@issp.u-tokyo.ac.jp">t-ozaki@issp.u-tokyo.ac.jp</a> , <a href="mailto:masahiro.fukuda@issp.u-tokyo.ac.jp">masahiro.fukuda@issp.u-tokyo.ac.jp</a>
Note	

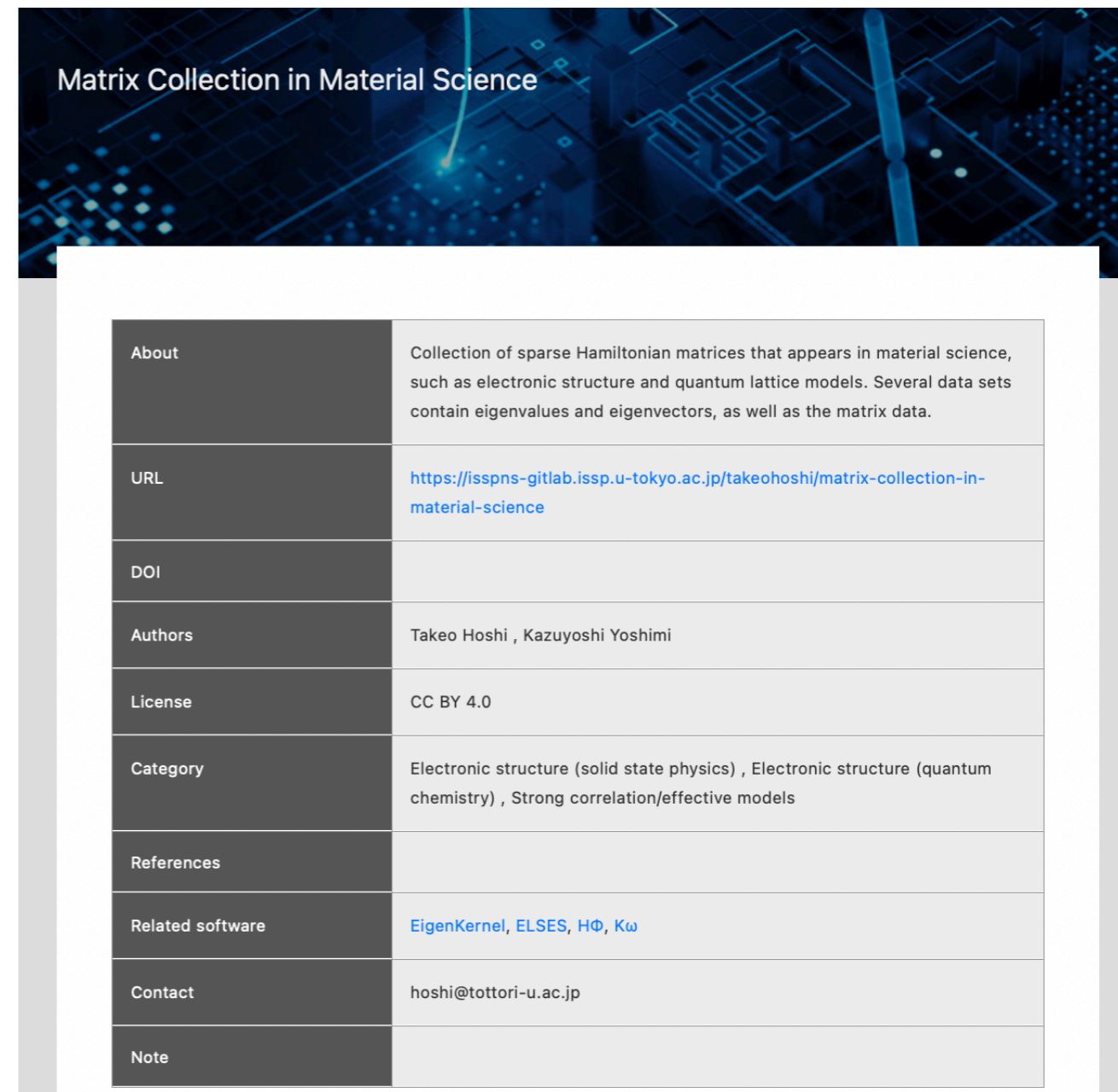
<https://isspns-gitlab.issp.u-tokyo.ac.jp/masahiro.fukuda/database-2019-of-optimized-fully-relativistic-pseudopotentials-and-pseudo-atomic-orbitals-for-openmx>



# データベース事例の紹介 (3)

## Matrix Collection in Material Science

- 疎行列を集めたデータベース集 (鳥取大 星先生)
  - 格納データ
    - 電子状態や量子格子模型のハミルトニアン行列がMatrix Market形式で格納。elsesとHΦを用いて生成。
  - 目的
    - 実用的な幾つかの模型(収束しにくい)に適用して、行列ソルバーの性能が改善されているかどうかなど検証する。




About	Collection of sparse Hamiltonian matrices that appears in material science, such as electronic structure and quantum lattice models. Several data sets contain eigenvalues and eigenvectors, as well as the matrix data.
URL	<a href="https://isspns-gitlab.issp.u-tokyo.ac.jp/takeohoshi/matrix-collection-in-material-science">https://isspns-gitlab.issp.u-tokyo.ac.jp/takeohoshi/matrix-collection-in-material-science</a>
DOI	
Authors	Takeo Hoshi , Kazuyoshi Yoshimi
License	CC BY 4.0
Category	Electronic structure (solid state physics) , Electronic structure (quantum chemistry) , Strong correlation/effective models
References	
Related software	<a href="#">EigenKernel</a> , <a href="#">ELSES</a> , <a href="#">HΦ</a> , <a href="#">Kw</a>
Contact	<a href="mailto:hoshi@tottori-u.ac.jp">hoshi@tottori-u.ac.jp</a>
Note	

<https://isspns-gitlab.issp.u-tokyo.ac.jp/takeohoshi/matrix-collection-in-material-science>

# データベース事例の紹介 (4)

## abICS Gallery

- ソフトウェアabICSのデータ集  
(山形大 笠松先生)
  - 格納データ
    - abICS関連論文の計算結果を再現するためのデータ一式 (aenetを利用して作成したニューラルネットワークを含む)
    - サンプルデータ一式
- 目的
  - 講習会開催時の演習問題、チュートリアルなどに利用。



About	abICS is a software framework for training a machine learning model to reproduce first-principles energies and then using the model to perform configurational sampling in disordered systems. Specific emphasis is placed on multi-component solid state systems such as metal and oxide alloys. The current version of abics can use neural network models implemented in aenet to be used as the machine learning model. As of this moment, abICS can also generate Quantum Espresso, VASP, and OpenMX input files for obtaining the reference training data for the machine learning model. This repository provides the sample data including input and output data for abICS.
URL	<a href="https://isspns-gitlab.issp.u-tokyo.ac.jp/abics-dev/abics-gallery">https://isspns-gitlab.issp.u-tokyo.ac.jp/abics-dev/abics-gallery</a>
DOI	
Authors	<a href="#">Shusuke Kasamatsu</a> , <a href="#">Kazuyoshi Yoshimi</a> , <a href="#">Yuichi Motoyama</a>
License	Documents: CC BY 4.0, Codes: GNU General Public License v3.0 or later
Category	Electronic structure (solid state physics) , Machine learning
References	
Related software	<a href="#">abICS</a>
Contact	<a href="mailto:abICS-dev@issp.u-tokyo.ac.jp">abICS-dev@issp.u-tokyo.ac.jp</a>
Note	

<https://isspns-gitlab.issp.u-tokyo.ac.jp/abics-dev/abics-gallery>

# 残されている課題

## エコシステム化する必要性

### ・データキュレーションなどのシステム導入

- ・いわゆる「失敗データ」も機械学習などには重要。データキュレーションシステムを導入して収集を自動化すると良さそう。ただし、データ量は膨大になってしまうので、どう管理するかが重要。
- ・システム導入について技術を身につけるための場が必要そう(話を受動的に聞く講習会ではなく、能動的に参加することが可能なハンズオン形式でやった方が良さそう)。

### ・運用コストに関する問題

- ・試験的なりポジトリ運用とはいっても運営に金額が非常にかかるため、各機関が運営するというのは無理がある。
- ・本来はハードウェアなどは一つの機関が中心になって管理して、そこを各機関が利用するという形の方が良いかもしれない → mdxの利用も検討中。



# 本日本話しすること

- 物質科学分野におけるオープンソフトウェアに関する取り組みの紹介
- 物質科学分野におけるオープンデータに関する取り組みの紹介
- 人材育成に関連した取り組みの紹介



# コミュニティ育成の重要性

知の共有、人材育成、および多様性の促進

最新のツールや技術の習得を支援する教育プログラム

✓ CCMSハンズオン講習会

多様なバックグラウンドを持つ参加者によるコミュニティの構築

✓ ハッカソン：技術者が集い、限定された時間で開発に挑戦

✓ 計算物質科学イノベーションキャンプ

✓ 2018年1月[1]：マテリアルズインフォマティクス関連

✓ 2018年10月[2]：数値計算ライブラリの活用

✓ 2019年10月[3]：データ科学の数値計算への応用

✓ **MateriAI2023** [4]：計算物質科学分野におけるAI技術の活用

[1] <https://ccms.issp.u-tokyo.ac.jp/event/489>, [2] <https://ccms.issp.u-tokyo.ac.jp/event/670>

[3] <https://ccms.issp.u-tokyo.ac.jp/event/2113>, [4] <https://ccms.issp.u-tokyo.ac.jp/event/6320>

# コミュニティ育成の重要性

知の共有、人材育成、および多様性の促進

最新のツールや技術の習得を支援する教育プログラム

✓ **CCMSハンズオン講習会**

多様なバックグラウンドを持つ参加者によるコミュニティの構築

- ✓ ハッカソン：技術者が集い、限定された時間で開発に挑戦
- ✓ 計算物質科学イノベーションキャンプ
  - ✓ 2018年1月[1]：マテリアルズインフォマティクス関連
  - ✓ 2018年10月[2]：数値計算ライブラリの活用
  - ✓ 2019年10月[3]：データ科学の数値計算への応用
- ✓ **MateriAI2023** [4]：計算物質科学分野におけるAI技術の活用

[1] <https://ccms.issp.u-tokyo.ac.jp/event/489>, [2] <https://ccms.issp.u-tokyo.ac.jp/event/670>

[3] <https://ccms.issp.u-tokyo.ac.jp/event/2113>, [4] <https://ccms.issp.u-tokyo.ac.jp/event/6320>

# CCMSハンズオン講習会 (1)

## 概要とサポート内容



井戸康太 助教

### • 概要

- 計算物質科学ソフトウェアに関する無料で参加できる実践的な講習会
- 講師：ソフトウェア開発者、サポートメンバー：CCMSスタッフ
- 参加者：毎回10~20名程度 (学生、PD、企業研究者含む)
- **コミュニティ支援の一環として、講義場所の提供と下記のサポートを実施**
  - Before：Webページ作成・メーリングリストへの告知+講習会当日のコミュニケーション環境 (Slack, Webex etc.) の準備
  - 当日：開発環境(MateriAppsLIVE!/物性研スパコン)の使用方法の説明+開発環境に関する質疑応答対応
  - After: アンケートを実施し、その結果を講師陣に報告



# CCMSハンズオン講習会 (2)

## スケジュールイメージ

### ・講習会開催までの流れ

- ・ 1ヶ月前：メーリングリストなどでの告知 (開催1週間前に締め切り)
- ・ 1週間前：参加者に事前準備の連絡＋スーパーコンピュータ利用のための手続き
- ・ 前日：Webexの接続テスト＋事前環境準備の対応

### ・当日スケジュールイメージ

- ・ 13:30～14:30：ソフトウェア概要の説明
- ・ 14:30～15:00：MateriApps LIVE!とスーパーコンピュータの使用方法の紹介
- ・ 15:00～16:00：チュートリアルなどを実演しながらのレクチャー
- ・ 16:00～17:30：自由演習



# CCMSハンズオン講習会 (3)

## 2023年度に開催した講習会一覧

### 2023年

- ・ 7月 AkaiKKR講習会



- ・ 9月 Quantum ESPRESSO & Wannire90講習会



- ・ 11月 TeNeS講習会



### 2024年

- ・ 1月 RESPACK講習会



- ・ 2月 SALMON講習会



- ・ 3月 2DMAT講習会

**2DMAT** ← 絶賛募集中！ (3/28開催予定)

# コミュニティ育成の重要性

知の共有、人材育成、および多様性の促進

最新のツールや技術の習得を支援する教育プログラム

✓ CCMSハンズオン講習会

多様なバックグラウンドを持つ参加者によるコミュニティの構築

✓ ハッカソン：技術者が集い、限定された時間で開発に挑戦

✓ 計算物質科学イノベーションキャンプ

✓ 2018年1月[1]：マテリアルズインフォマティクス関連

✓ 2018年10月[2]：数値計算ライブラリの活用

✓ 2019年10月[3]：データ科学の数値計算への応用

✓ **MateriAI2023** [4]：計算物質科学分野におけるAI技術の活用

[1] <https://ccms.issp.u-tokyo.ac.jp/event/489>, [2] <https://ccms.issp.u-tokyo.ac.jp/event/670>

[3] <https://ccms.issp.u-tokyo.ac.jp/event/2113>, [4] <https://ccms.issp.u-tokyo.ac.jp/event/6320>

# MateriAI2023 (1)

## MateriAI2023開催の背景

- **2019年 社会連携研究部門-データ統合型材料物性研究部門の設立**

(第Ⅰ期：2019年～2023年、第Ⅱ期：2024年～)

**TOYOTA**

- 目的：実験と数値計算をデータ科学的手法によって統合し、電子相関の理解に基づいて、革新的な機能を持つ材料の物性予測・探索手法を開発する。
- 対象：永久磁石、軟磁石、スピントロニクス材料、超伝導材料など

- **ハッカソン開催に至った動機**

- 最先端のデータ科学技術を学び、その技術を実践・共有する機会を設けることができれば、社会連携部門・産業・学術の垣根を超えた人財の育成に繋げることができるのではないかと？



# MateriAI2023 (2)

## ハッカソンの概要



東京大学 物性研究所  
THE INSTITUTE FOR SOLID STATE PHYSICS  
THE UNIVERSITY OF TOKYO

TOYOTA

- 日時：2024年2月12日 13:00 ~14日 15:00
- 場所：御殿場高原 時之栖
- 参加者：40名 (ハッカソン参加者：29名)
- 目的：計算物質科学分野におけるAI技術の活用



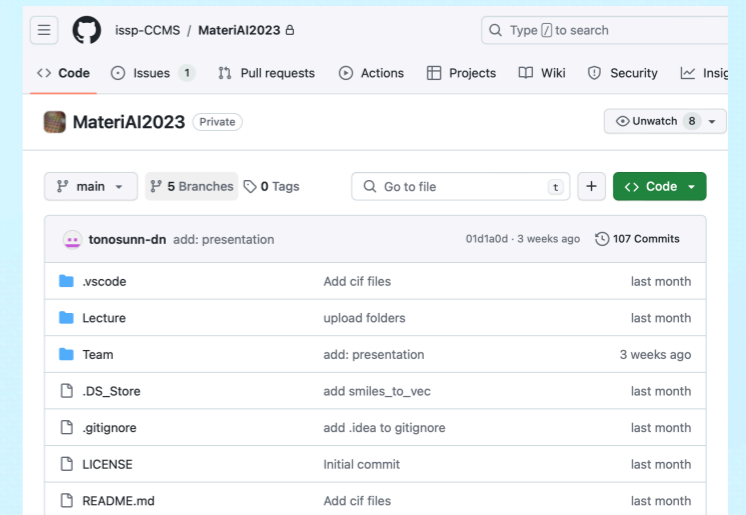
- AIを活用したコーディング技術の効率化、物質科学情報の自動収集、AIコンシエルジュの開発などを実施。AIの基礎から応用までの知識を深め、物質科学分野の課題解決におけるAIの応用方法を学び、それらの経験を共有する。
- チームに分かれたプロジェクト作業を行うことで、実践的なスキルに加え、チームワークの経験を積むこともできる。



# MateriAI2023 (3)

## ハッカソンの流れ

- ・ チーム (5 or 6名で構成) ← 事前アンケートで先にチーム編成
  - ・ コーディング班 2班、データ解析班 2班、情報収集班 1班
- ・ タイムスケジュール
  - ・ 1日目：13:00～18:00
    - ・ AIを活用した事例の紹介+データ解析環境の構築  
(ChatGPT, Slack, GitHub, Google Colab)
    - ・ チームの自己紹介+目標の設定
  - ・ 2日目：09:00～18:00
    - ・ チーム作業メイン (関連講演 2件)
  - ・ 3日目：09:00～15:00
    - ・ チーム作業+成果発表



# MateriAI2023 (4)

## 各チームで取り組んだ内容(の一部)



- **コーディングチーム1** : Materials Projectから結晶構造のデータを抽出。それらを元に、原子の種類と単位胞内の各原子の相対位置から単位胞の体積をAIで予測する。
- **コーディングチーム2** : 結晶構造ファイルから特徴量を導出。特徴量から超伝導転移温度をAIで予測する。
- **データ解析チーム1** : 異なる実験 (XRDなど) 毎に特徴量を求め、それらのmapを作成することで、特徴量の関係性がどのように現れるかを考察する。
- **データ解析チーム2** : Large Language Model(LLM)を使って、論文を学習データに使い、一般的な物性 (バンドギャップなど) を回帰する。
- **情報収集チーム** : Webページから情報を収集するスクリプトを作成し、有機材料を対象にした物性値のデータベースを生成する。



# MateriAI2023 (5)

## 参加者の声 (アンケートより：17名回答)



- 講義内容・ボリューム：ちょうどいいという回答がほぼ得られた。
  - ただし、外国人の参加者がいたので、Zoom翻訳機能を使用したけど、設定が適切ではなく出鱈目な翻訳になる場合が多く、機能しなかったのが反省点。
- 研究会の日程：「ちょうど良い」が7割。「もう少し長い方がいい」が3割。
- 次回の参加希望：7割が参加したいと回答。それ以外はテーマ・進路によって、参加を検討したいという回答。ただし、1名は参加しないという回答(海外参加者)。
  - 要望：目標をあらかじめ設定して、その達成を競うハッカソンをやってみたい。

次年度以降も開催予定。興味がありましたら、ぜひご参加ください！

# Summary

- データの創出、品質、および環境提供

- ✓ MateriAppsプロジェクト
- ✓ ソフトウェア開発・高度化プロジェクト
- ✓ 物性研究所データリポジトリ



- 知の共有、人材育成、および多様性の促進

- ✓ ソフトウェア講習会 (CCMSハンズオン講習会)の開催
- ✓ ハッカソン(MateriAI)の開催



- 持続可能性および相互運用性 ← 今後の課題

- ✓ リソースの長期的な持続可能性と拡張性
- ✓ データとツールの相互運用性





# 謝辞

- ・ 今回の発表にあたり以下のプロジェクト・チーム・開発メンバーの方に感謝の意を表します。
- ・ MateriApps開発メンバー
- ・ ソフトウェア高度化・開発プロジェクト関連メンバー
- ・ データリポジトリ事業関連メンバー
- ・ トヨタ自動車株式会社 先端材料技術部