

# 人文学・社会科学におけるAI for Scienceについて (案)

令和8年 3月10日

科学技術・学術審議会 学術分科会 人文学・社会科学特別委員会

# 目次

---

1. 背景・意義

---

2. 人文学・社会科学におけるAI for Scienceの効果

---

3. 人文学・社会科学におけるAI for Scienceに関する課題

---

4. 今後の方向性

---

参考資料

## 背景

### ① 研究DX、DHの進展

大規模かつ大量のデータを分析する手法が普及

### ② AI for Scienceの拡大

AIがあらゆる分野の研究活動を根底から変えうるゲームチェンジャーとして深く関与。

我が国もAI for Scienceを国家戦略として位置づけ。

➔ 1. 背景・意義

## 課題

### ① AIへの透明性・信頼性の確保

我が国の文化等について誤った結果が出力される事例が散見。

### ② AIに準拠した研究データインフラの構築

多くの分野では、研究に活用できるだけの質・量のデータは未学習（特に、我が国をはじめとする東アジア特有の書誌体系への対応）

### ③ 研究体制の拡大

より広範囲の領域のデータを分析、解釈する必要性から、チーム型の研究が拡大。

### ④ 研究コストの増加

### ⑤ 権利問題への対処

複雑な権利問題への対応が増加

➔ 3. 人文学・社会科学におけるAI for Scienceに関する課題

## AI利活用による効果

### ① 研究プロセスの高速化・効率化、研究の質の向上

○膨大なデータ解析

これまで人間が処理不能だった規模に対応。

### ② 研究者の創造力の発揮、新たなネットワーク構築への寄与

○新たな知の創出

推論や仮説立てをAIが補助。

○関係性の発見

他分野など、これまでつながらなかった知見や研究者が提示され、より総合的に事象を研究できる。

### ③ 各分野の研究への裨益、研究者の負担軽減

○各分野の研究への裨益

AI活用のためのデータインフラ整備が、我が国の人文学・社会科学の知の蓄積・活用・発信にもなる。例えば、外国からも日本研究をしやすくなるなど、分野の研究にも裨益。

○負担軽減

翻訳、校正等がAIの補助により効率化。

➔ 2. 人文学におけるAI for Scienceの効果

# 1. 背景・意義

## 前提：人文学・社会科学におけるデジタル、AI

- 研究のデジタル化、AIの利活用は人文学・社会科学に革新をもたらし得る。
- 一方、研究の在り方が変遷しても、人文学・社会科学の本質（人間の精神活動の根本・根源の理解や、社会的合意形成等を探求する学問分野であり、社会の在り方や人間の生き方の再考に寄与するとともに、人間や社会に関する「想像力」を広げ知的好奇心を刺激することで、人間の生きる力の根源や社会の根本を支える。）は変わらない。
- あくまで人間が研究の主体・中心であることには変わりがない。

## 〈背景〉

### ①研究DX、DHの進展

- 欧州では、2000年代からデジタル・ヒューマニティーズ(DH)が拡大。これにより、人文学においても大規模かつ大量のデータを分析する手法が広まった。
- 我が国でも、2010年代半ばから、学会の創設、大学設置のセンターによる組織的なDH推進へと拡大。
- 国においても、2024年度から、国内諸機関による共同体制（DHコンソーシアム）の構築、データ規格のモデルガイドラインの策定、データ利活用研究のユースケースの創出、教育プログラムの開発・実施・検証を推進（～2026年度：※人文学・社会科学のDX化に向けた研究開発推進事業）。



## ②AI for Scienceの拡大

- あらゆる研究分野において、AI利活用（AI for Science）が急速に進展。データ活用、仮説づくり、実験のあらゆる段階でAIを活用。
- 人文学・社会科学においても、AIを利活用した研究が行われ始めている。
- 我が国としても、AI for Scienceを国家戦略として位置づけ、計画的かつ集中的に推進していくこととしている。

### 〈意義〉

#### ①AI for Scienceの拡大の人文学・社会科学・自然科学を通じた政策的な目的

##### （1）科学研究の革新と科学的発見の加速・質の変革

研究プロセスの自動化・知能化により知の生産性を向上させ、**研究者がより高次元で創造的な活動に専念できる環境を実現**

##### （2）研究力の抜本的強化と科学の再興

**従来の延長線上では解決できない科学的課題に挑戦**し、世界をリードする信頼性のある科学的成果を継続的に創出

##### （3）国際的優位性・戦略的自律性の確保

主要なAI研究開発拠点として、技術的不可欠性と戦略的自立性を確立し、不可欠な国際研究パートナーとなり、AI for Science 先進国の地位を築く

#### ②人文学・社会科学においてAI for Scienceが拡大することの意義

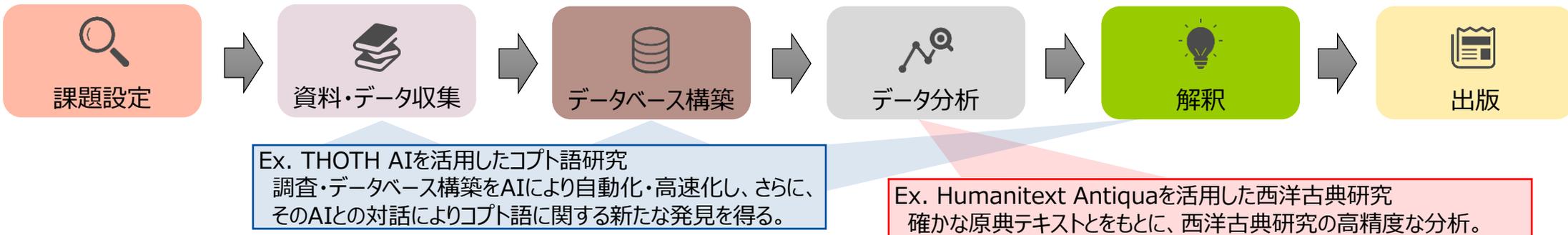
○研究への障壁がなくなり、研究者だけでなく、あらゆる人々が人知へのアクセス・活用が可能になる

**= 人文学・社会科学が「民主化」**

○人文学・社会科学研究を始めるための知へのアクセス性が向上

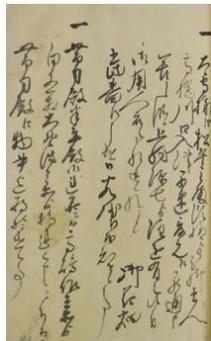
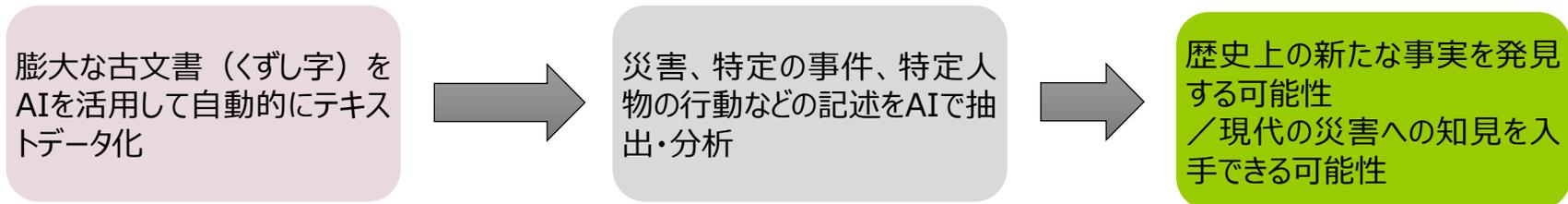
➡ **人文学・社会科学の研究に携わる者の増加**

# 人文学・社会科学の流れとAIの関係

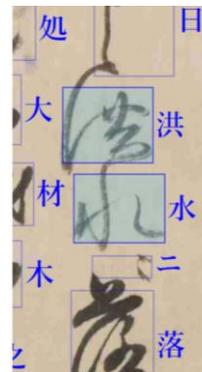


※研究の流れ図は、Citizen Humanities (The Science of Citizen Science (P97-118)) を参考に文部科学省作成

## 人文学・社会科学のAI利活用の例



細川家文書  
約5万枚・約950万文字



【参考文献】  
・熊本大学とTOPPAN、くずし字AI-OCRを活用した古文書の大規模調査のための独自手法を開発  
([https://www.holdings.toppan.com/ja/news/2024/07/newsrelease240726\\_2.html](https://www.holdings.toppan.com/ja/news/2024/07/newsrelease240726_2.html))  
・歴史ビッグデータ構造化による安政江戸地震被害記録の分析 北本朝展(ROIS-DS人文学オープンデータ共同  
利用センター/国立情報学研究所) 橋本雄太(国立歴史民俗博物館) 大邑潤三・加納靖之(東京大学地震研究所)「  
人文学とコンピュータシンポジウム」2024年12月

## 2. 人文学・社会科学におけるAI for Scienceの効果

### ① 研究プロセスの高速化・効率化、研究の質の向上

- AIによる分析はこれまで人間が処理できた範囲を超え、かつその速度も人間のそれを超えたものとなっているため、人文学・社会科学のあらゆるプロセスを高速化・効率化。
- 人文学・社会科学の研究者にとって重要な解釈のプロセスに自身の時間を割くことができるようになり、研究がより精微化することから、研究の質の向上にも寄与。

### ② 研究者の創造力の発揮、新たなネットワーク構築への寄与

- AIは分析や翻訳、論文校正等の補助的な役割のみならず、推論や仮説を立てる、研究計画を提案するというような研究の内容面にまで関わる共同研究者のような振る舞いをするまでに至っている。研究者はAIの提案も基にすることで、一層の創造力を発揮し、新たな知を創出することができるようになる。
- AIによる膨大かつ広範なデータ分析により、これまで人間が気づけなかった事柄どうしの関係性を発見することが、これまで関わりのなかった異分野の研究や研究者どうしをつなげる可能性がある。

### ③ 各分野の研究への裨益、研究者の負担軽減

- AIのために整備されたデジタルアーカイブ・データインフラは我が国の人文学・社会科学の知の蓄積や、国内外の研究者による利活用、我が国の知の発信に寄与。それにより、日本研究のような国外で日本について研究する研究者に適切な情報を提供する手段となるなど、個々の分野の研究の発展にも寄与。
- 翻訳、論文校正等をAIが補助することにより、研究者の負担軽減に寄与。

### 3. 人文学・社会科学におけるAI for Scienceに関する課題

#### ① AIへの透明性・信頼性の確保

- 現在主流となっている対話型生成AIは、膨大な学習データを機械学習し、使用者の入力に対する回答としてふさわしいものを推測するという仕組み。
- AIの主たる提供者の海外企業は、学習データの内容をほとんど明らかにしておらず、出力結果の適切性を判断することが困難。そのため、透明性・信頼性に課題。
- 学習データが欧米中心の資料に基づいている可能性が指摘されており、我が国の文化等について誤った結果が出力される事例が散見。
- AIの仕組み上、データの偏りやバイアス、事実に基づかない誤った情報をもっともらしく生成すること（ハルシネーション）は大なり小なり避けられないが、人文学・社会科学への活用にあたっては、それに耐えうる精度まで向上させることが必要。
- こうしたAIの特性を理解し、AIによる生成物を鵜呑みにせず、一次資料による裏付けや、人間（人文学・社会科学研究者）がAIの処理プロセスに介入し、監修、検証、解釈を行うことが必要。

## ②AIに準拠した研究データインフラの構築

- AIの進展によって、これまでの研究DXよりも一層高いレイヤーで高精度、大量なデータへの需要が増加。
  - ※構造化データ（文字列や絵図等が持つ情報（意味、音韻、位置等）を付与し、機械が内容を可読できるようにしたデータ）、非構造化データのいずれも需要が増加する。
- 構造化データについても、AIへの利活用を前提とした場合、我が国をはじめとする東アジア特有の書誌体系（縦書き、漢字等）への対応が必要。
- AIへのデータ利活用に当たっては、人文学・社会科学の研究者もデータやAIの仕組み等を理解していることが重要であるため、人材育成の枠組みが必要。
  - ※例えば、東京大学大学院では、人文社会系研究科の大学院生を対象に、AIによる研究手法を教育する講座を令和7年度から開始しているなど、大規模大学では枠組みが構築され始めている。
- 高精度かつ大量なデータを構築することが正当な業績として評価されることが必要。

## ③研究体制の拡大

- AIの進展により、人文学・社会科学の研究者と情報学の研究者の連携ニーズが増加。また、人文学・社会科学の中でもより広範囲の領域にわたるデータを分析、解釈することとなり、研究者どうしの連携が増加。
  - ➡ チーム型、共同研究型の研究手法が拡大
- そのような時代の変化に即した研究環境整備、支援の枠組みが必要だが、本来役割を果たすべき人文学・社会科学系の共同利用・共同研究拠点は減少傾向

#### ④研究コストの増加

- 従来の人文学・社会科学に比べ、AIを利活用する研究は各種システムの構築・運用に係る経費やライセンス料といった費用を必要とし、従来の人文学・社会科学よりも研究遂行に必要な資金が増加する。

※令和7年度補正予算において、「AI for Scienceによる科学研究革新プログラム」が創設され、あらゆる分野の研究者がAIを活用して科学研究の高度化・加速化を図るため、計算資源の確保等の研究環境を整備し、アカデミア全体にAI for Scienceの波及・振興を推進し、意欲ある研究者による次の種や芽となる新たなアイデアへの挑戦への支援が行われることとなった。

#### ⑤権利問題への対処

- 人文学・社会科学が研究で使用する原資料やデータには、所有者や原資料に投影・記述されている者、データ作成者等との間に複雑な権利関係（所有権、肖像権、著作権等）が存在。
- AIの進展により、権利問題の複雑性は増大（権利者や研究者の意図しない形での二次利用のリスク等）。
- 複雑な権利問題について、研究者から相談を受け、適切な対応を助言・補助できる人材が必要だが、現状では、そのような専門人材を育成・配置する環境に乏しい。

※原資料やデータの取扱いについては、法的な権利関係のみならず、その所有者と研究者の信頼関係や、研究者の研究者の倫理観、振る舞いが重要であることに留意が必要。

#### ⑥その他

- 翻刻や注釈等のこれまで学術的業績とみなされていた行為がAIによって可能となることで、研究評価の観点が増える可能性。

※研究にAIの利活用が浸透していくなかで、どのような仕事が増えるべきか議論が必要。

## 4. 今後の方向性

### 〈前提の再掲〉

- 研究のデジタル化、AIの利活用は人文学・社会科学に革新をもたらし得る。
- 一方、研究の在り方が変遷しても、人文学・社会科学の本質（人間の精神活動の根本・根源の理解や、社会的合意形成等を探求する学問分野であり、社会の在り方や人間の生き方の再考に寄与するとともに、人間や社会に関する「想像力」を広げ知的好奇心を刺激することで、人間の生きる力の根源や社会の根本を支える。）は変わらない。
- あくまで人間が研究の主体・中心であることに留意し、今後の方向性を検討する必要がある。

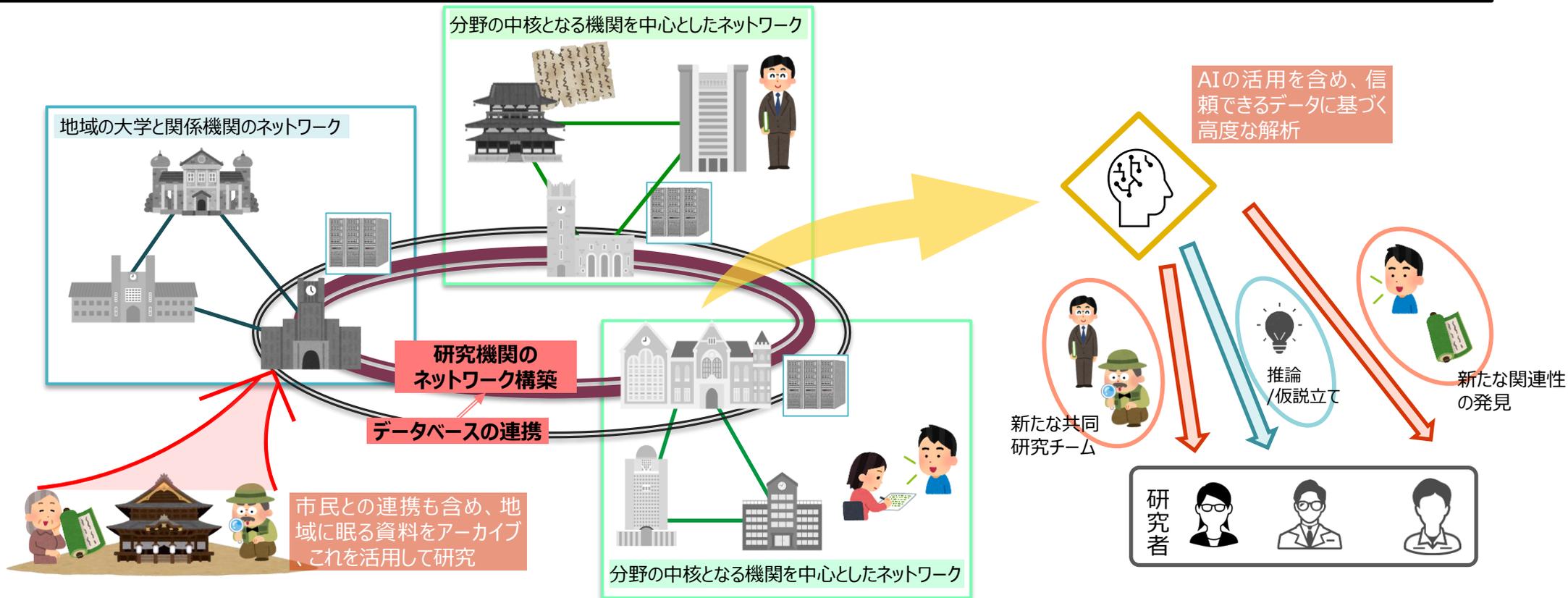
○人文学・社会科学においてAI for Scienceの考え方を取り入れるためには、以下について対応が必要。

- (1) AI利活用研究に必要な正しい知識基盤、信頼できるデータを構築するための  
〔 設備（サーバ等） / システム（研究用アプリケーション、ローカルAI等）  
データインフラ（データ共有、閲覧プラットフォーム）  
教材（AI Readyデータ作成、権利問題への対応マニュアル、講義動画等）  
の整備、研究者への共同利用 〕
- (2) AI利活用研究を支援する、  
〔 データキュレーター / 権利問題専門人材 〕  
の配置、及び当該人材による学内外の研究者への支援、**原資料の所有者との橋渡し**
- (3) AI利活用研究の開始・推進に向けた、  
〔 人文学・社会科学－情報学 / 人文学・社会科学内の複数領域 〕  
**それぞれの研究者の橋渡し（あっせん、紹介、共創の場の提供）体制の構築**

○さらに、**地域の諸機関と大学**の間に、**データと人材をつなぐネットワークを構築することで、新たな異分野融合研究、共同研究を促進し、人文学・社会科学を総合的に発展させる必要がある。**

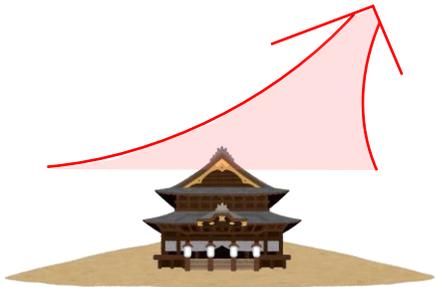
# 人文学・社会科学のAI利活用を支える研究ネットワークの構築

人文・社会科学は、**資料・データの解釈・批評、理論化等を通じて人間・社会を読み解くもの**。信頼でき、かつ、偏りのない十分な量の資料・データを整備し、さらに、AIを活用する、AIを通じて可視化された研究者と共同研究を進める、**より多く幅広い資料・データを、より多角的に**研究することで、人間・社会についての深く新たな発見・理解が可能になる。このため、各データベース同士の連携、アーカイブ機関と強く連携する地域/分野の中核となる大学等研究機関のネットワークの構築により、データの利活用を進めることが考えられるのではないか。



## 期待される効果

- ✓ ネットワークに属する機関/地域の**研究者の可視化**、及び、それによる**共同研究・融合研究の促進**
  - ➔ **人と地域をつなぐ（＝信頼関係）の強化、地域に眠る人々の知の吸い上げ促進による、不可視だった人々の価値の表出**
- ✓ **データの探索の高速化**、中核となる大学等が引き続き管理することによる**データ管理の適正の確保**
  - ➔ **AIの利活用が進むことで人文学研究が深化・高度化**
- ✓ 研究資源のデータが国内外から可視化されることにより、日本研究など**分野の研究の活性化**
  - ➔ **日本関連の人文学研究の国際プレゼンスの向上に寄与**



**春日大社** 768年 創立(社伝上)  
1998年 世界遺産認定

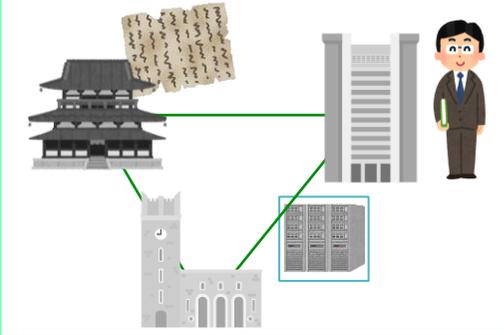


(<https://www.kasugataisha.or.jp/>)

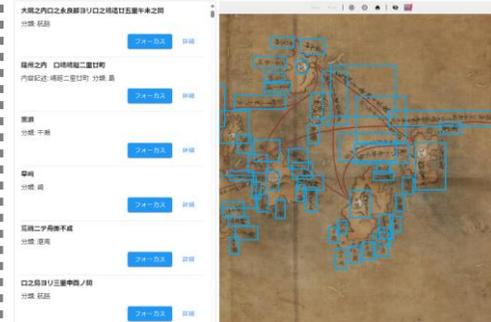
- ・奉納和歌、歌会資料(非公開)を大量に保有
- ・平成24年:国文学研究資料館が調査を開始
- ・令和7年:国文研国書データベースへの公開を開始

➡ 研究者と寺社との間の信頼関係の醸成が資料の公開・研究への利活用に結実

分野の中核となる機関を中心としたネットワーク



**東京大学史料編纂所**



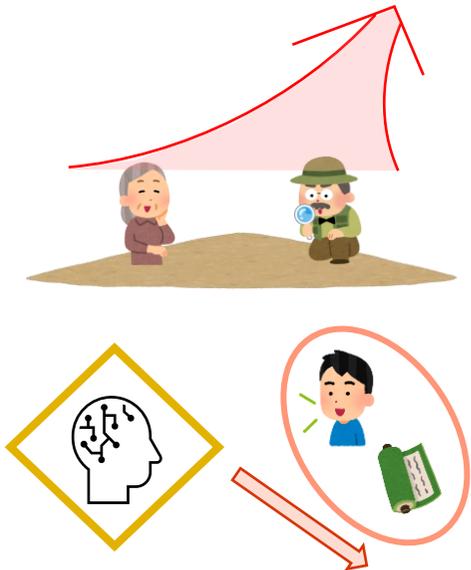
- ・県立金沢文庫、都城島津邸等の地域史料の修理、デジタルアーカイブを支援
- ・史料編纂所のデータベースでもデータ化→公開

- ・海外の機関が所蔵している日本の歴史資料にも着手

「都城島津邸所蔵「琉球并諸島図」アーカイブ」

**江戸期の地震被害の分析**

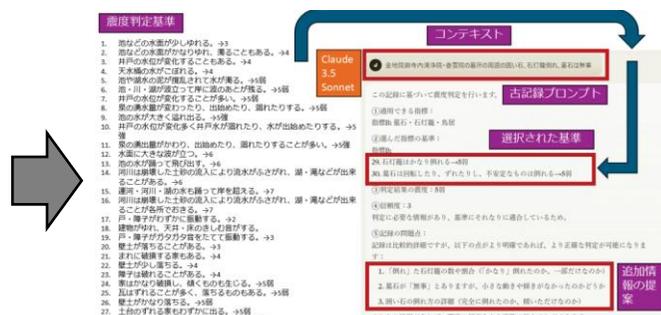
- ・古文書に書かれている地震の揺れに関する情報(例:池の水が踊って飛び出す)を市民の力を借りて抽出 → データ入力(“みんなで注釈”)
- ・入力したデータをLLMで分析 → 震度情報に変換(例:池の水が踊って飛び出す→震度6相当)
- ・防災科研の地盤データと比較し、当時の被害状況を現代風地図に可視化



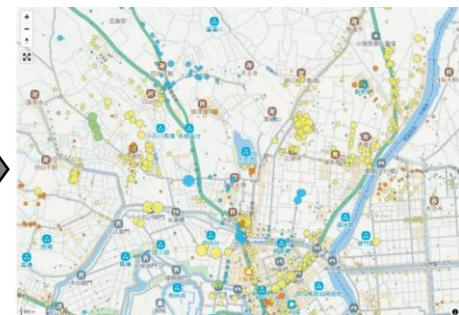
「みんなで注釈」の入力画面



震度判定基準の一部とLLMを用いた震度判定の出力例



地震被害記録の可視化(建物)



※歴史ビッグデータ構造化による安政江戸地震被害記録の分析 北本朝展(ROIS-DS人文学オープンデータ共同利用センター/国立情報学研究所) 橋本雄太(国立歴史民俗博物館)大邑潤三・加納靖之(東京大学地震研究所)「人文科学とコンピュータシンポジウム」2024年12月

# 参 考 資 料

## A データ規格のモデルガイドラインの策定、データ利活用研究のユースケースの創出

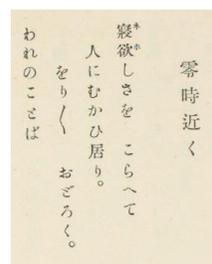
## 領域①文字資料 領域②地図・地誌類

### ■ データ規格モデルガイドラインの策定

ガイドラインに盛り込む規格の例 (イメージ)

#### 【文字資料】

改行、字下げ、字あけ (日本語固有の表記手法) をコードで記す方法を統一



```
<head>零時近く</head>
<| mxl:id="ori-haru-0073">
寝欲しさをこらへて<|b style="text-indent: 1em">
人にむかひ居り<|pc type="句点">。</pc>
<|b style="text-indent: 2em">
をり| |<|pc type="空白"></pc>お
どろく<|pc type="句点">。</pc>
<|b>われのことば
</|>
```

<https://doi.org/10.20730/300004817>

#### 【地図・地誌類資料】

・方位の設定、記述 (登録) 方法の統一  
・基準地点の設定、登録方法、対象地の統一

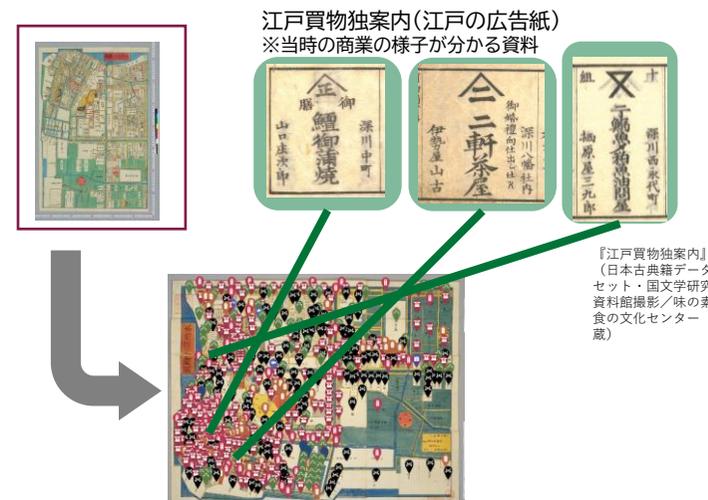


重ね合わせた上での研究が容易に

### ■ データ利活用研究のユースケースの創出

創出するユースケースの例 (イメージ)

- ・デジタル化した古典籍から昔の災害発生地の特定制  
➔防災分野への活用
- ・文献資料との組み合わせによる、江戸時代の地域商業の解明



江戸買物独案内 (江戸の広告紙)  
※当時の商業の様子が分かる資料

『江戸買物独案内』  
(日本古典籍データセット・国文学研究資料館撮影/味の素食の文化センター蔵)

『江戸マップデータセット』  
(ROIS-DS人文学オープンデータ共同利用センター作成)

#### データ規格の統一によるメリット (例)

古典籍テキストデータの解析、比較、検索が格段に効率化

時代の変遷等の比較研究が効率化

異分野の資料や異なる資料 (文献と地図等) の組み合わせなど分析の幅が広がる

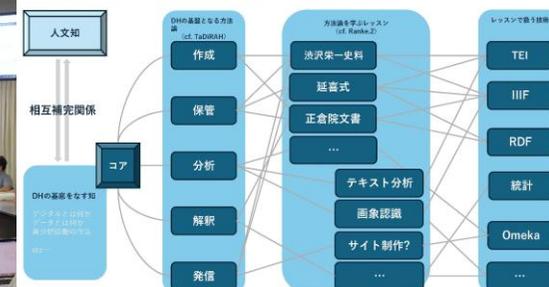
## B データ構築やデータ・AI利活用研究を可能とする人材の育成

- 人文系データ人材の育成に係る調査
- 教育プログラム (オンデマンド、オンライン、実習) の開発、実施・検証

### 教育プログラム



2025年10月作成版 DiHuCO人材育成プログラム全体像の素案



### カリキュラムマップの策定

## C 「デジタル・ヒューマニティーズ・コンソーシアム」の設置・運営

- コンソーシアムの進捗管理・取りまとめ・調整
- データ規格の国際対応
- コンソーシアム活動の成果の普及・啓発

# (参考) AI for Scienceに関する議論の状況

## (3) 今後の方向性 (目的)

### AI for Science の方向性

AI for Science は、AI政策、科学技術・イノベーション政策、安全保障政策を横断する複合的課題

AIを活用して科学研究を根本から革新し、AIによる研究パラダイムの転換と、科学技術立国としての戦略的自律を図る。

- ✓ 研究環境をAIで刷新する
- ✓ 研究システム・研究プロセスをAIで変革する
- ✓ 研究のあり方をAIで革新する
- ✓ AI for Science 時代の組織変革（分野・組織を越えた連携等）

2030年代・・・

- ✓ AIが研究の自然な一部として活用される環境
- ✓ AI for Science が日本の研究力の中核として国際的に認知
- ✓ 分野横断型人材が学術・産業双方で活躍
- ✓ 自律性と信頼性を備えた研究国家としての確立

### ■ 政策的な目的

#### 1. 科学研究の革新と科学的発見の加速・質の変革

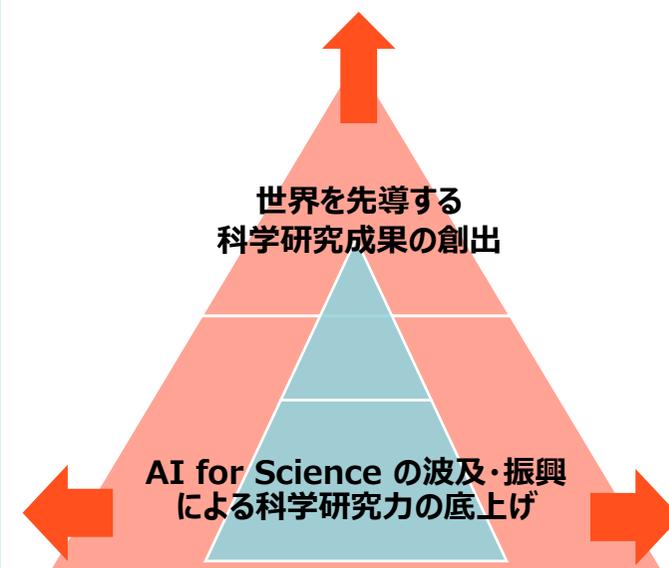
研究プロセスの自動化・知能化により知の生産性を向上させ、研究者がより高次元で創造的な活動に専念できる環境を実現する

#### 2. 研究力の抜本的強化と科学の再興

従来延長線上では解決できない科学的課題に挑戦し、世界をリードする信頼性のある科学的成果を継続的に創出する

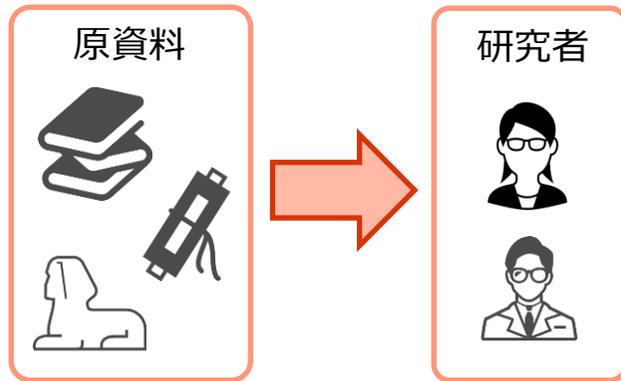
#### 3. 国際的優位性・戦略的自律性の確保

主要なAI研究開発拠点として、技術的不可欠性と戦略的自律性を確立し、不可欠な国際研究パートナーとなり、AI for Science 先進国の地位を築く

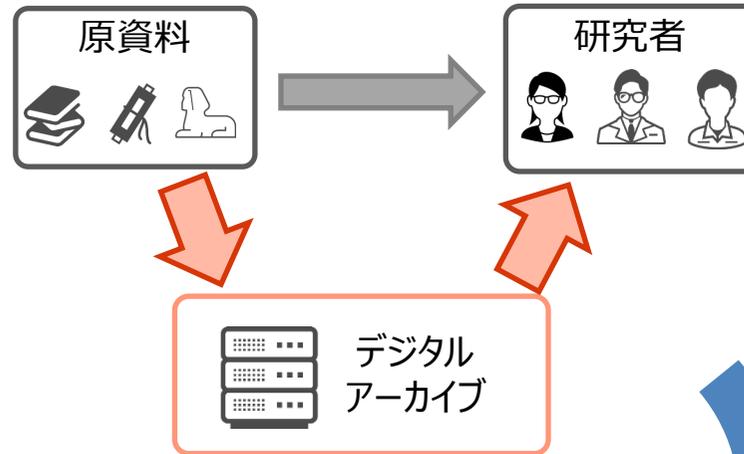


# デジタル化・AI利活用の人文学・社会科学研究への影響

研究のデジタル化が進む前



資料のデジタル化が進展した場合の研究への影響

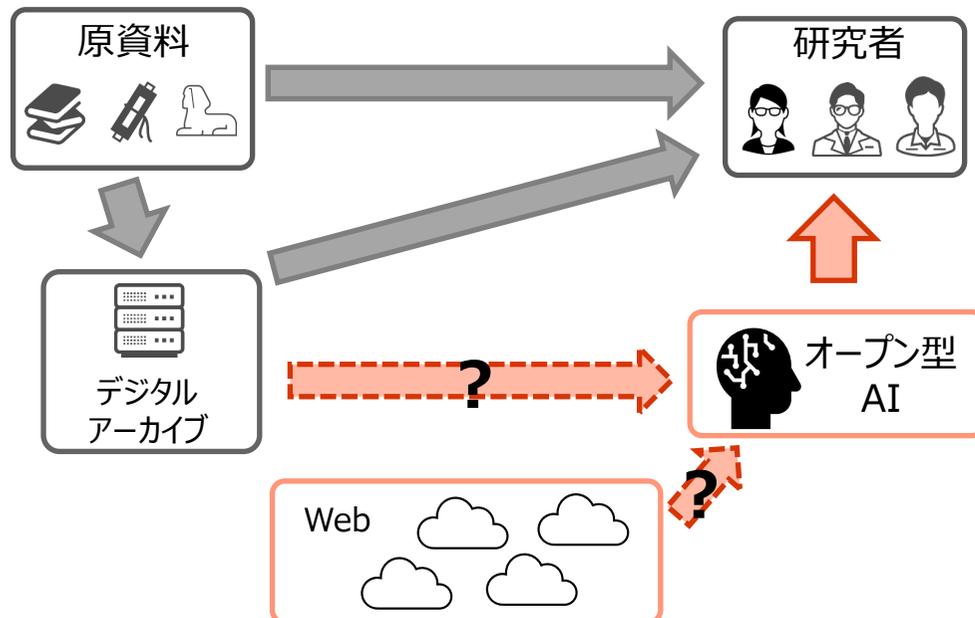


物理的にアクセスできない原資料も  
一定程度参照できるようになる。  
(但し、原資料の保全・参照の価値は  
低減しない)



著作権・肖像権等の権利問題など、  
資料の所有者と利用者間で調整  
を要する事項が顕在化。

AIの活用が進展した場合の研究への影響



人間が処理できなかった膨大なデータがAIによって解  
析可能となり、研究が加速化・効率化。  
AIが研究のあらゆるプロセスをサポートすることで、研  
究の深み、幅が拡大。



オープン型AIがどの情報をもとに回答を生成しているか  
は必ずしも明らかでなく、研究に活用するには透明性・  
信頼性の観点から課題。

## 第28回人文学・社会科学特別委員会のヒアリング事例

第28回人文学・社会科学特別委員会（令和7年12月5日）の委員会では、透明性・信頼性を高めるために信頼できるデータのみを参照するAIを研究に活用する事例と、AI利活用の前提となるデジタルアーカイブの振興にも必要な権利問題への対応事例についてヒアリングを実施。

