

第16回CODHセミナー

歴史ビッグデータと「タイムマシン」構想



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）

<http://codh.rois.ac.jp/>

「まち」や都市のデジタルアーカイブ - 歴史ビッグデータと実世界での利活用



「まち」や都市に関する歴史的な文化資源のデジタルアーカイブ

1. 文化資源をデジタル化するだけでは、**実世界での利活用とのギャップ**を埋めることは簡単ではない。
2. **デジタルアーカイブと実世界**とをうまく結合するための方法論が求められている。
3. **課題と将来像を議論する。**

資料中心型の視点と実世界での利活用

1. **資料中心型の視点**：人文学では、（文書／モノ）資料から読み取れることを、当時の世界を基準に考える。
2. **実世界での利活用**：現在の人々が知りたいことを基準として、過去の世界に関する情報を得たい。
3. **資料中心型の視点でアーカイブを構築**しても、実世界での利活用に適合しない問題がある。
4. **問い**：現在の**実世界を基準**として、**過去の資料を使いやすくする**にするには、どうすればよいか？

歴史ビッグデータの統合解析

<http://codh.rois.ac.jp/historical-big-data/>

過去のビッグデータを統合解析するための基盤技術の研究



歴史的資料
(史料)

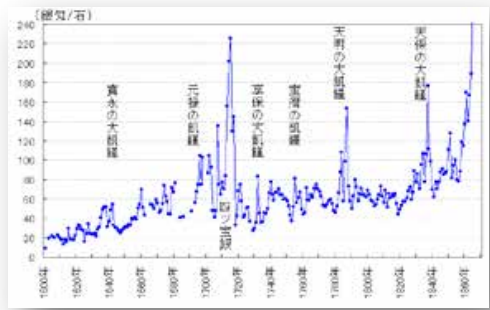
自然科学的
データ

人文
社会的
データ

- 気候
- 地震
- 噴火
- 疫病
- 経済
- 人口
- 政治
- 文化

データ
構造化
ワーク
フロー

歴史ビッグ
データ研究基
盤 (機械可読
データ)



2022/3/28

「まち」や都市のデジタルアーカイブ - 歴史ビッグデータと実世界での利活用 #codh16

歴史ビッグデータとは？

1. 現代のデータを想定して開発されたビッグデータの方法論を、過去の世界の記録に延長すること。
2. データの構造化：人間や機械（AI）による注釈を効率化し、大規模な機械可読データを生成する。
3. 多様（Variety）なデータの統合解析：識別子によってデータを統合し、複数分野にまたがる問いに答える。
4. 人文学研究のデジタルトランスフォーメーション（DX）：様々な資料（文書、モノ等）に対して、デジタル技術を活用した新しい研究方法を提案する。

歴史ビッグデータで収集する記録データ

1. 歴史的状況記録

- 世界の状況を観察した記録（空の状況を視覚的に観察した記録、地震の揺れを聴覚・触覚で感じた記録、地震の被害状況を視覚で観察した記録など）

2. 歴史的行動記録

- 人間の行動に関する記録（人間がある場所から別の場所に移動した記録、店で何かを買ったり食べたりした記録、観光した記録など）

3. 歴史的状態記録

- 人間やモノの状態に関する記録（人間やモノの属性に関する記録、市場取引の価格に関する記録、モノの輸送に関する記録など）

edomi - 江戸をみる／みせるデータポータル

<http://codh.rois.ac.jp/edomi/>



1. 江戸に関する**構造化データを公開するデータポータル**。エンティティ基準でデータを統合し、現代の体系と接続。
2. 現代のYahoo!のように、**カテゴリごとに江戸データを公開**。
3. 現在のカテゴリ：「トラベル」「ショッピング」「マップ」「芸能（試験中）」など。

エンティティ（実体）基準のアーカイブ

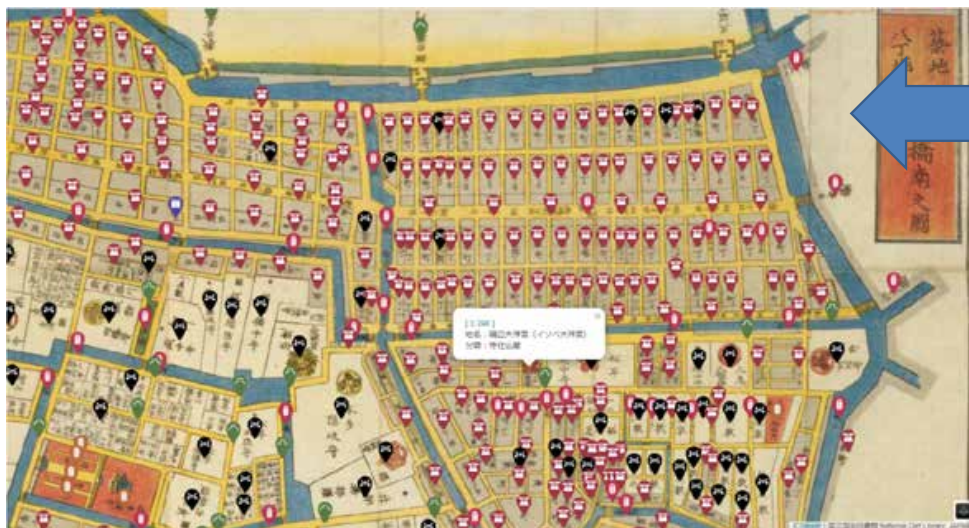
1. エンティティ = 実世界に存在する実在物の単位。空間（地名）、時間、人物、出来事（イベント）など。
2. エンティティには識別子と属性を付与し、これを様々なアプリケーションで共有化することで、多種多様なデータを紐づけられる。
3. どの粒度でエンティティを定義するかはアプリケーションに依存する。

例えば「マイナンバー」は、個人に識別子を付与し、識別子に紐づく各種の属性（氏名）も管理する。

江戸ビッグデータ

<http://codh.rois.ac.jp/edo-maps/>

画像：江戸切絵図（国立国会図書館）
地名：CODHがデータクリエイターと共に、
IIF画像への注釈としてデータ整備



CODH 鈴木 親彦ほか



現代地図との重ね合わせ

地名識別子の整備・共有 (CODH)



GeoLOD

<https://geolod.ex.nii.ac.jp/>



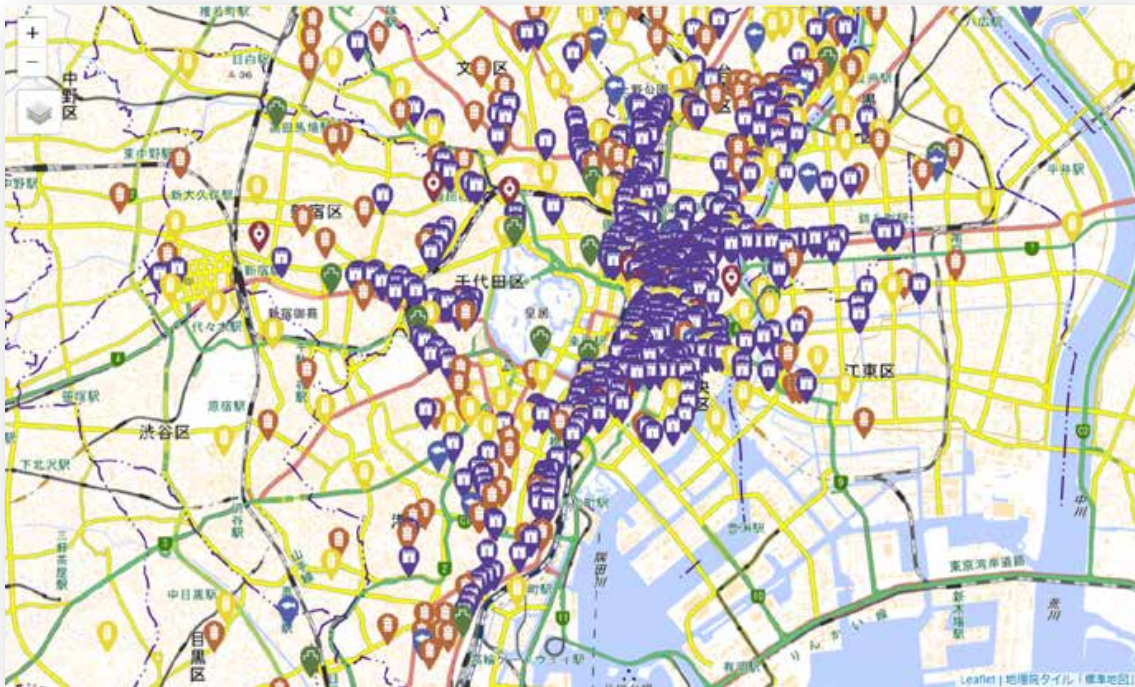
商業ビッグデータ



観光ビッグデータ

江戸の都市空間のデジタル化

<http://codh.rois.ac.jp/edomi/map/>



江戸の商店や観光地の地図。江戸時代の町名を、江戸切絵図を通して、現代の緯度経度とリンク。

清河八郎が安政2年（1855）に江戸を訪問した日記『西遊草』の行動記録。







江戸のクリエイティビティ（歌舞伎）



出典：立命館大学アートリサーチセンター（ARC）、協力：赤間亮教授

浮世絵顔データセット

<http://codh.rois.ac.jp/ukiyo-e/face-dataset/>

Painter	Examples
Hirosada (広貞)	
Kogyo (耕漁)	
Kunichika (国周)	
Kunisada (1st gen) (国貞初代)	
Kunisada (2nd gen) (国貞二代目)	
Kunisada (3rd gen) (国貞三代目)	
Kuniyoshi (国芳)	
Toyokuni (1st gen) (豊国初代)	
Toyokuni (3rd gen) (豊国三代目)	
Yoshitaki (芳滝)	

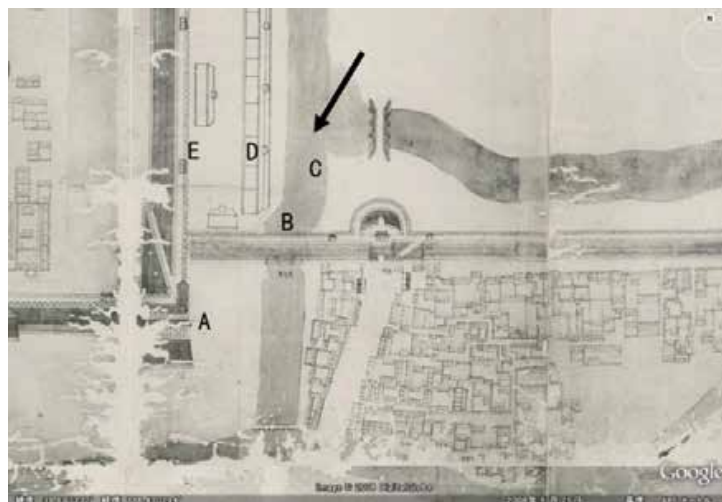
『ARC浮世絵顔データセット』 (Yingtao Tian、ROIS-DS CODH作成、ARCから収集) , <https://doi.org/10.20676/00000394>

1. 立命館ARCがNII-IDR経由で公開する浮世絵データを利用。
2. Google BrainのYingtao Tian氏が、顔検出のAPIを用いて顔データセットを作成。
3. 機械学習 (AI) によるアノテーションで、エンティティの抽出を支援。

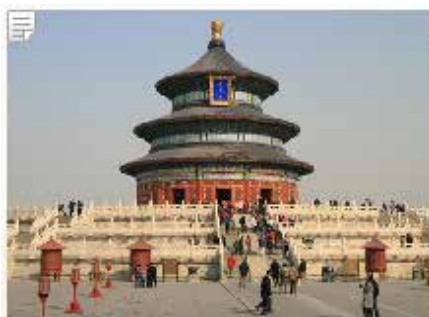
都市や「まち」の景観復元

<http://dsr.nii.ac.jp/beijing-maps/>

ビジュアルな資料を組み合わせた都市景観の復元も重要な課題



北京における、古写真と地図の比較による、3次元的な景観比較



古写真と現代写真・地図上の描画の比較

東洋大学 西村陽子氏との共同研究

メモリーグラフ：記憶 を重ねる新しい写真術

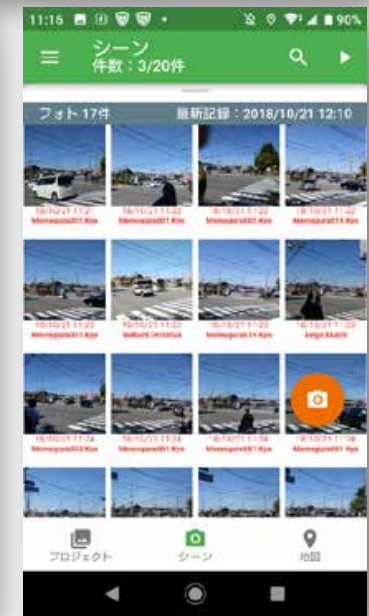
Memory + graph

<https://mp.ex.nii.ac.jp/mg/>

1. ファインダー上で「**撮るべき写真**」のガイドを描画する（構図グリッドの拡張）。
2. **ガイドに風景を合わせて撮影**すれば、同一地点・同一構図の写真が簡単に撮れる。
3. **撮影時のセンサーデータ**を保存すれば、古写真の撮影位置・方向が推定できる。
4. 任意の写真を基準に使用すれば、「**疑似的な定点観測**」も可能となる。

アプリ操作の流れ

旧版がインストール不可となり、新版（Android / iOS）をFlutterで開発中。2022年前半の公開を目指す。



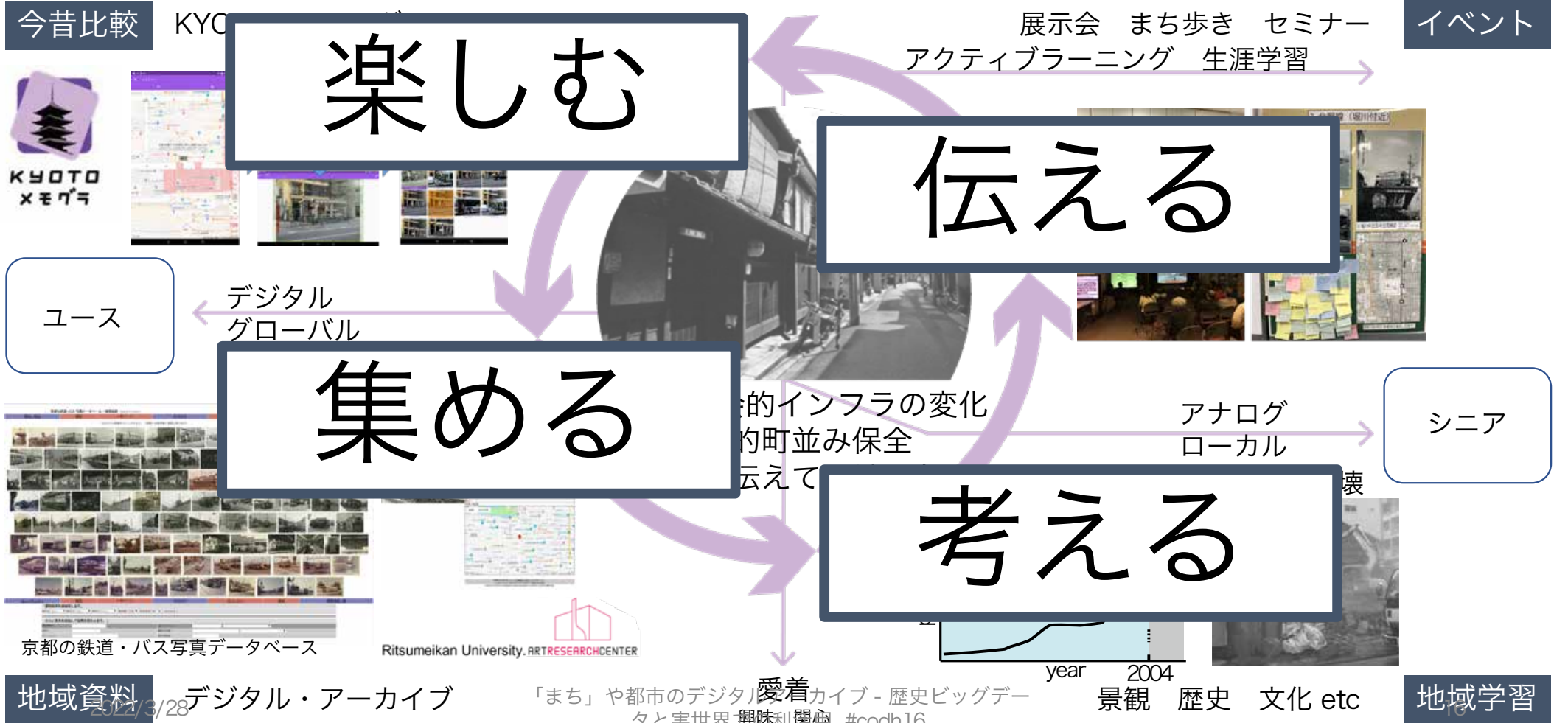
2022/3/28

「まち」や都市のデジタルアーカイブ - 歴史ビッグデータと実世界での利活用 #codh16

15

提供：大阪大学サイバーメディアセンター 高橋 彰氏（DS公募型共同研究の成果）

メモグラを活用した地域学習の枠組み



クライシスによる変化と都市のアーカイブ

<http://dsr.nii.ac.jp/bam/>

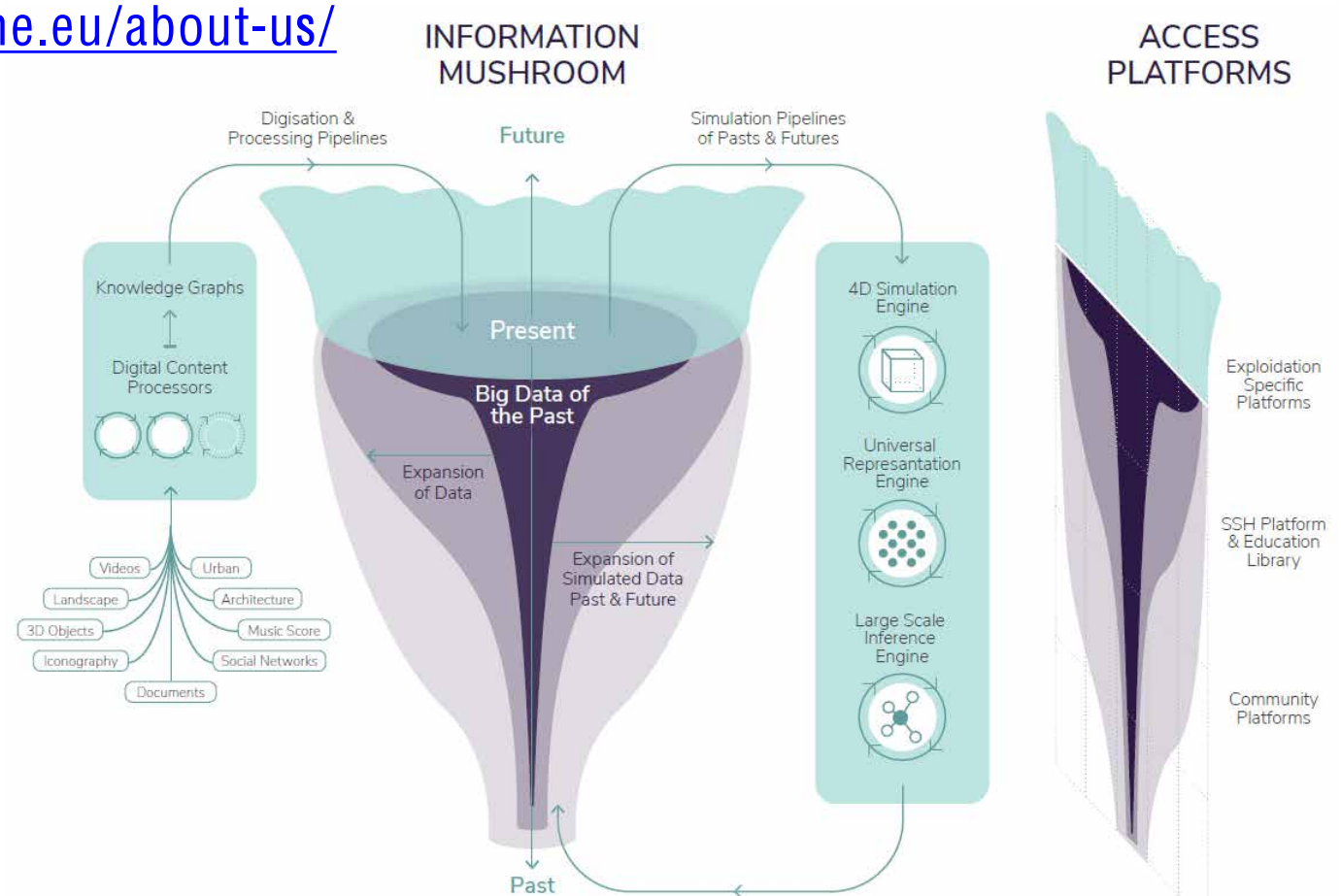


2003年の地震で破壊された**イラン・バム遺跡**を3次元復元。テヘラン大学のエルハム・アンダルーディ氏を中心としたチームが、**10年ほどの時間**をかけて、都市の主要部分を復元。

タイムマシンの概念図

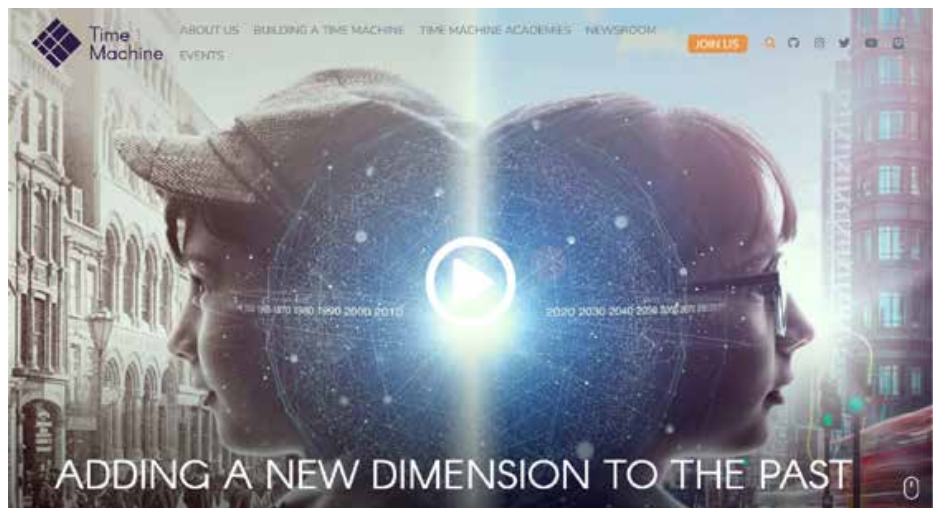
<https://www.timemachine.eu/about-us/>

1. **デジタル化**で過去のデータを増やす。
2. **AI**で構造化データを増やす。
3. **シミュレーション**で欠けているデータを補う。
4. **過去を未来に延長する (4D)**。



「まち」や都市のデジタルアーカイブ - 歴史ビッグデータと実世界での利活用 #codh16

European Time Machine



- **欧州**を中心とした4000年の文化遺産デジタル化と活用
- **予算**：当初予算は10年で1000億円（現在は大幅減額？）

Hanyang Time Machine



- **韓国・ソウル**の3次元復元（メタバース）と知識ネットワークの構築
- **予算**：7.6百万ドル／年

日本における「タイムマシン」の展開



1. 江戸（東京）に加え、文化の中心である京都も必須！
2. 江戸の幕藩体制（200藩以上）は、地域に独自の文化を生み出した。
3. 地域（まち）の文化をデジタル化して実世界で活用する仕組みを、各地域が独立しつつ連携して展開することは可能か？

武鑑全集「居城マップ」

<http://codh.rois.ac.jp/bukan/book/200018823/map/>

2022/3/28

「まち」や都市のデジタルアーカイブ - 歴史ビッグデータ
と実世界での利活用 #codh16

21

参考文献

1. **江戸**：北本 朝展, 鈴木 親彦, 寺尾 承子, 堀井 美里, 堀井 洋, "地理的史料を対象とした歴史地名の構造化と統合に基づく江戸ビッグデータの構築", 人文科学とコンピュータシンポジウム じんもんこん2020論文集, pp. 171-178, 2020年12月
2. **バム**：Asanobu KITAMOTO, Elham ANDAROODI, Mohammad Reza MATINI, Kinji ONO, "Post-Disaster Reconstruction of Cultural Heritage: Citadel of Bam, Iran", 人文科学とコンピュータシンポジウム じんもんこん2011論文集, pp. 11-18, 2011年12月
3. **北京**：西村 陽子, 北本 朝展, "Google Earthと『乾隆京城全図』を用いた北京歴史空間の情報基盤", 人文科学とコンピュータシンポジウム じんもんこん2008論文集, pp. 81-88, 2008年12月
4. **メモリーグラフ**：北本 朝展, "災害の非可逆性とアーカイブの精神ーデジタル台風・東日本大震災デジタルアーカイブ・メモリーグラフの教訓", デジタルアーカイブ・ベシックス2 災害記録を未来に活かす, 今村文彦 監修/鈴木親彦 責任編集 (編), pp. 169-197, 勉誠出版, ISBN 978-4-585-20282-0, 2019年8月
5. **欧州タイムマシン**：北本 朝展, "「過去のビッグデータ」を探る欧州タイムマシン研究計画", カレントアウェアネス, No. E2248, 2020年4月