

2021年11月27日
@相模女子大学

経営理論II 第8回

未来社会を展望する(1) : ICTの未来

「データ・人工知能・オープン×市民」

合同会社ミドリアイティ

代表 小池 隆

会社紹介

- ▶ 合同会社ミドリアイティ 2014年設立
- ▶ 企業理念は「オープンソース&オープンデータによる、健全で豊かなIT文化の創出」
- ▶ 専門分野は、地理情報システム、データ分析・可視化、AI（機械学習）
- ▶ 主力事業はソフトウェアの受託開発
 - ▶ ただし、同業社からの二次請けはしない
- ▶ ソーシャルビジネス
 - ▶ 地域情報化に関わる支援等 - コロナ禍で縮小中

自己紹介

- ➡ 早稲田大学第一文学部哲学科心理学専修卒
- ➡ 準大手 S I 企業にて受託開発と研究開発に従事
- ➡ 2014年に起業
- ➡ 社会活動
 - ➡ ウィキペディアに郷土史関連の記事執筆、執筆イベント主宰
- ➡ 研究活動
 - ➡ 情報処理学会 人文科学とコンピュータ研究会 会員
 - ➡ 石造物（石仏、石碑）の研究
 - ➡ 石仏情報学会 を立ち上げ
 - ➡ 石造物のデータ化、データに基づく石造物研究

本日のお話

当社の企業理念

「オープンソース&オープンデータによる、健全で豊かな
IT文化の創出」

に絡めて、

- ➡ データ
- ➡ 人工知能
- ➡ オープン×市民

について様々な事例とともにお話します

(当社の関係する事例は見出しに★を付けてあります)

第1部 データ

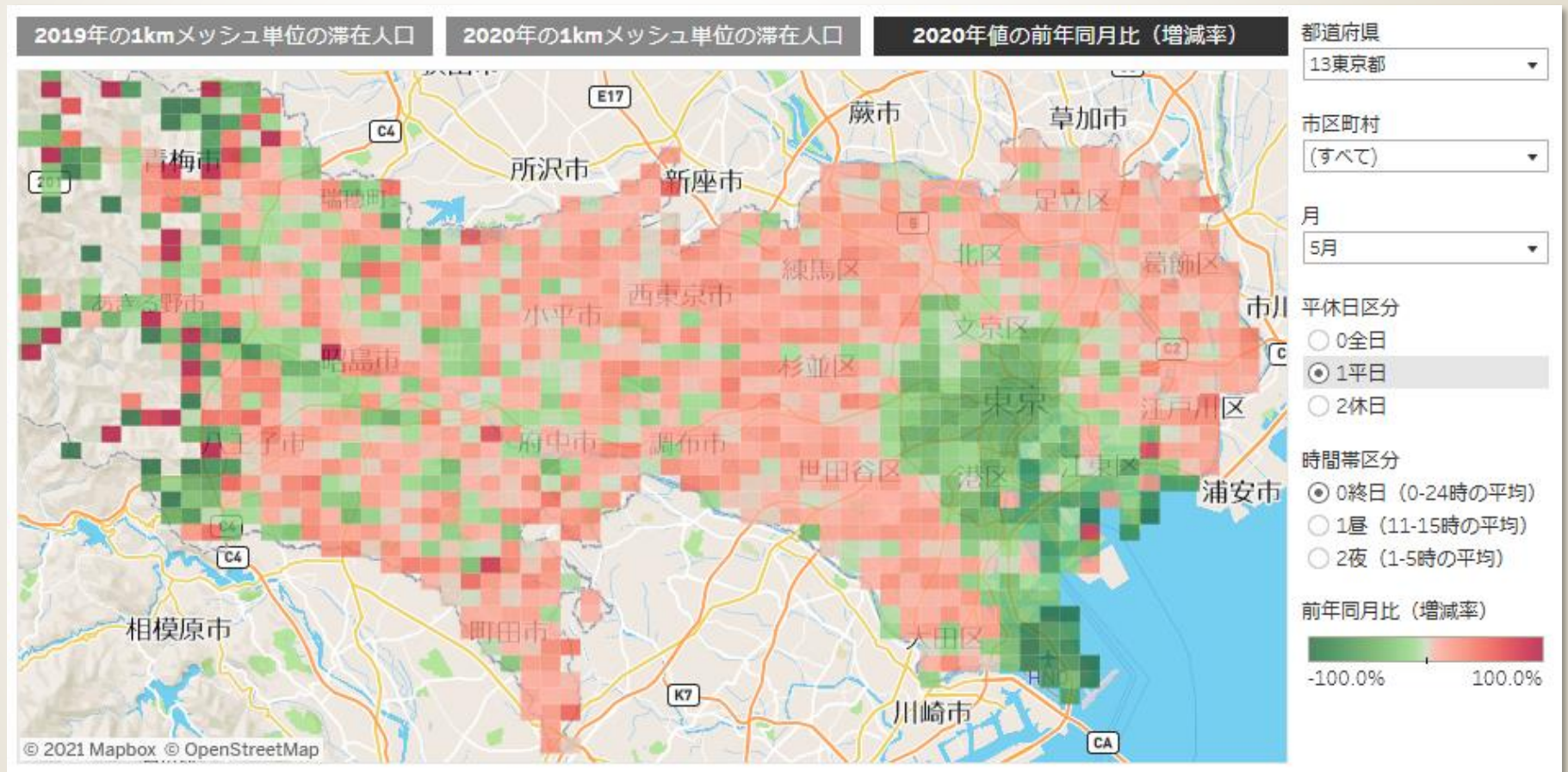
コロナ禍で高まったデータへの関心



<https://corona.go.jp/dashboard/>

「人流データ」

- ➡ 携帯電話端末等の位置情報のビッグデータを集計



東京都 新型コロナウイルス感染症対策サイト

- ➡ コロナ関連のオープンデータを可視化
- ➡ プログラムをオープンソースとして公開し、全国に派生



<https://stopcovid19.metro.tokyo.lg.jp/>

ビッグデータとオープンデータ

■ ビッグデータ

- 大量のデータ
- 非公開 or 有償 or 公開だが自由には使えない
 - 集計されたデータがオープンデータ化されることもある

■ オープンデータ

- 営利目的、非営利目的を問わず無償で二次利用可能
- クリエイティブ・コモンズ (CC) ライセンス
 - CC0 パブリックドメイン (全ての権利を放棄)
 - CC BY 出典を明記すれば改変可能、商用利用可能
 - CC BY-SA (派生物を同一ライセンスで配布)

オープンデータ意義・目的

▶ オープンデータ基本指針

(高度情報通信ネットワーク社会推進戦略本部・官民データ活用推進戦略会議決定)

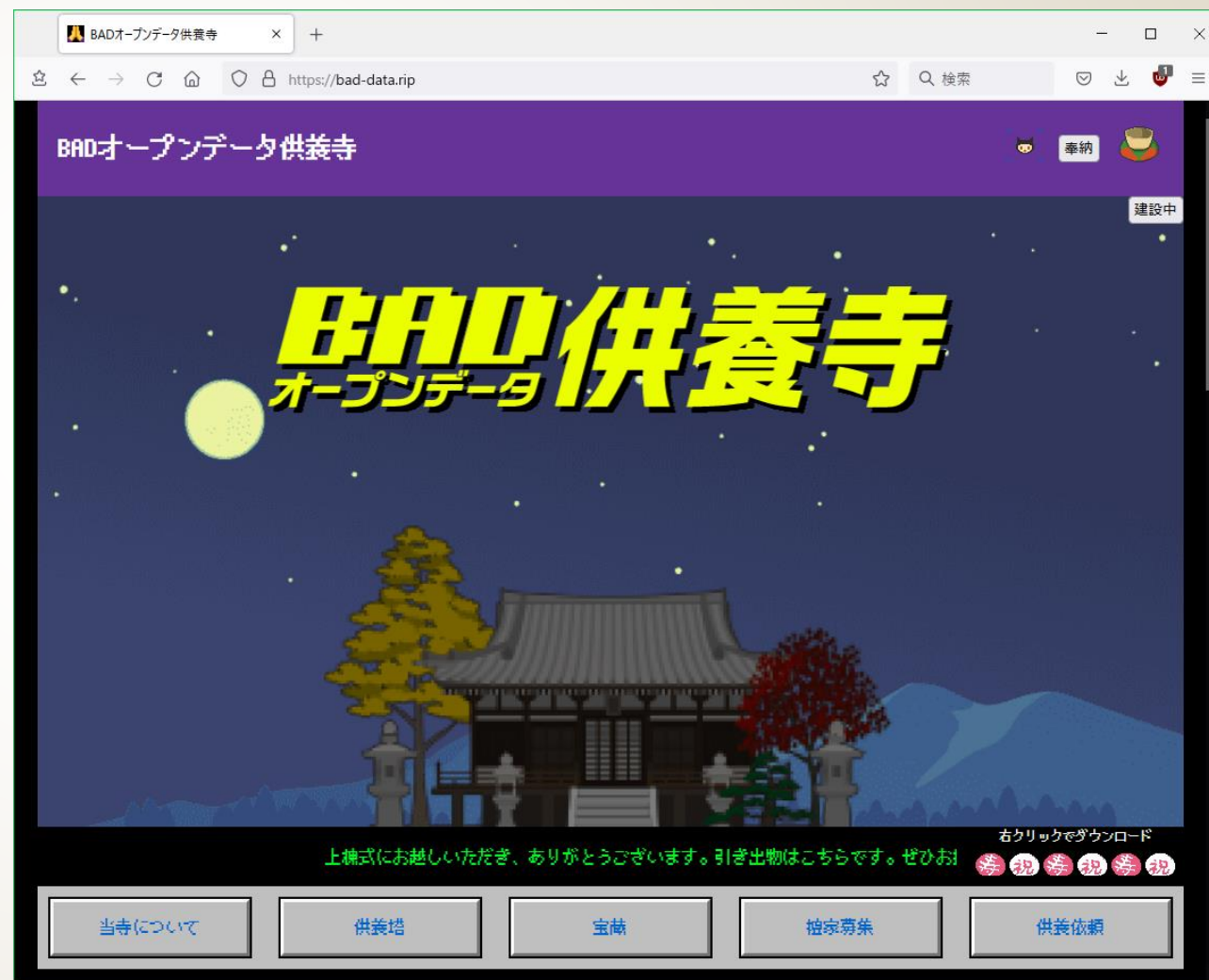
- ▶ 国民参加・官民協働の推進を通じた諸課題の解決、経済活性化
- ▶ 行政の高度化・効率化
- ▶ 透明性・信頼の向上

▶ EBPM（エビデンスに基づく政策立案）

- ▶ 「その場限りのエピソードに頼るのではなく、政策目的を明確化したうえで合理的根拠（エビデンス）に基づくものとする」
－ 内閣府におけるEBPMへの取組

残念なオープンデータ

- ➡ PDFファイル
- ➡ ネ申エクセル
 - ➡ Excel方眼紙
- ➡ 表記ゆれ
 - ➡ 名寄せできない



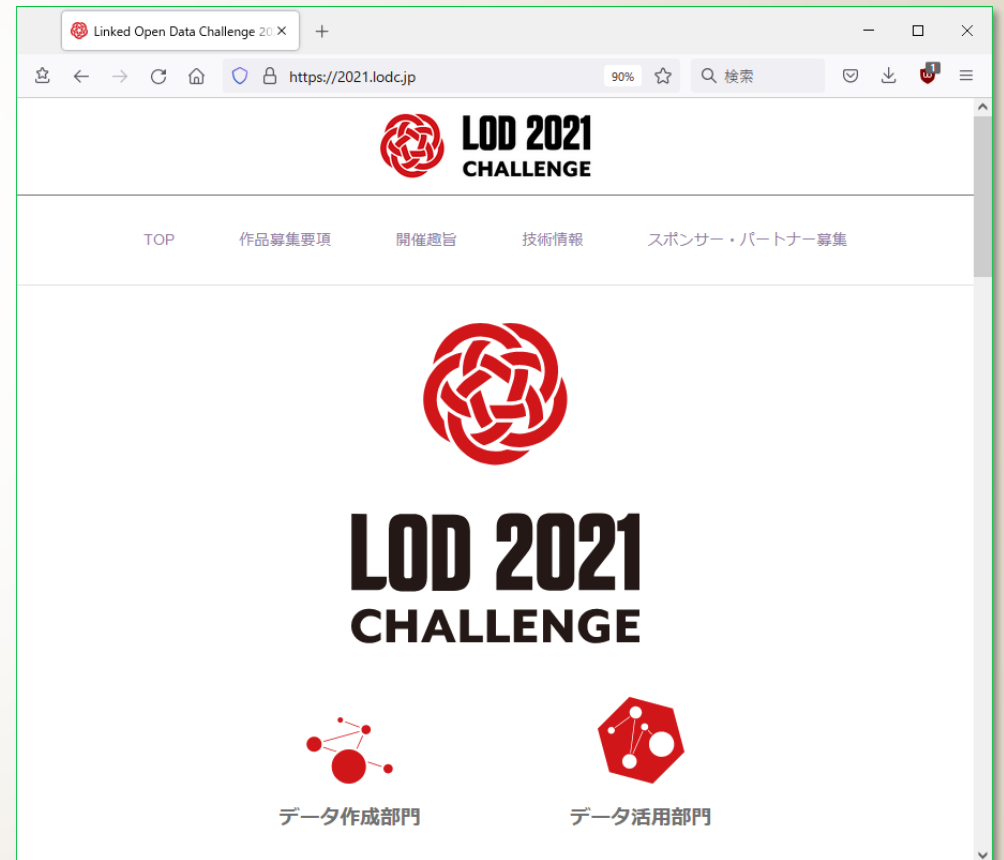
<https://bad-data.rip/>

欲しいデータがない？

- ➡ ないものは作る
- ➡ 公開する
- ➡ コンテストに応募する



<http://linkdata.org>



<https://2021.lodc.jp/>

★江戸後期 武蔵・相模国村名マップ

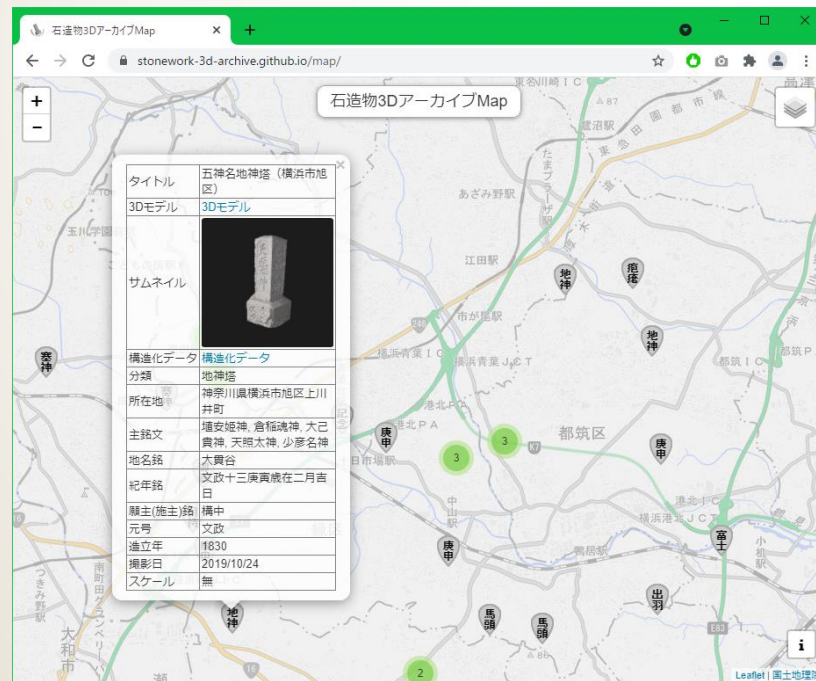
- ▶ 江戸時代の地誌『新編武蔵風土記稿』『新編相模国風土記稿』に記載された村を地図上に配置
- ▶ 国会図書館デジタルコレクションの該当ページにリンク



<https://fudoki.midoriit.com>

★3Dモデルによる石造物デジタルアーカイブ

- ➡ 石造物3Dアーカイブプロジェクト (篠田 浩輔、小池 隆)
- ➡ フォトグラメトリで石造物の3Dモデルを作成して公開
 - ➡ アーバンデータチャレンジ2019 金賞



<https://stonework-3d-archive.github.io/map/>



データは誰のもの？

➡ 情報銀行

- ➡ 個人情報を受け、データから得られた利益を本人に還元

➡ データ税？



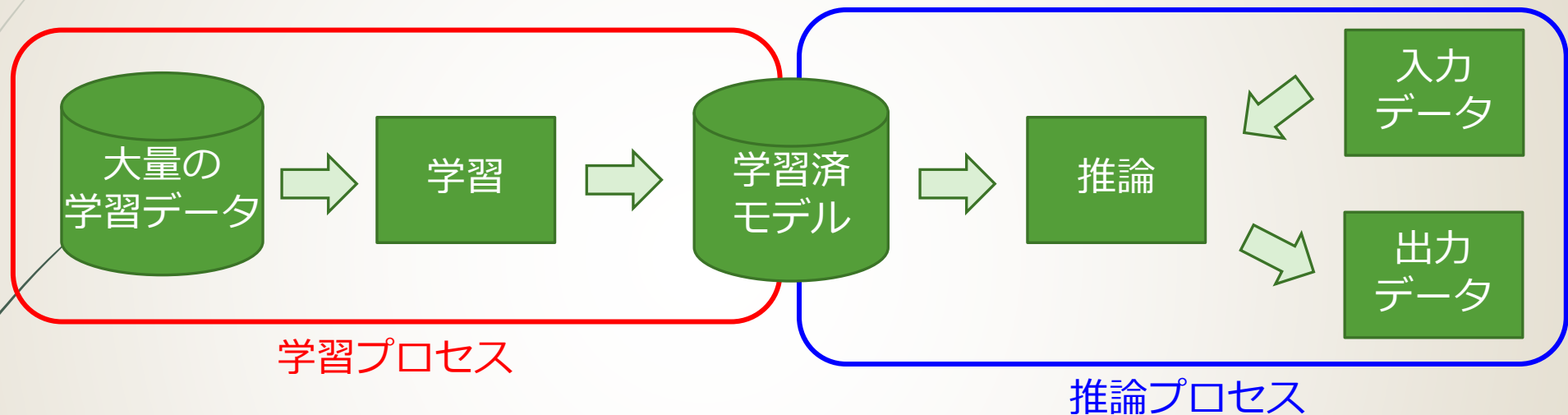
第2部 人工智能

人工知能

- ➡ 人工知能（AI）の2大潮流
 - ➡ 論理知識型AI – 論理的、演繹的
 - ➡ データ駆動型AI – 統計的、帰納的
- ➡ 現在はデータ駆動型AIが主流
 - ➡ 機械学習、特に深層学習（ディープラーニング）
 - ➡ 急速な発展と様々な分野への展開
- ➡ 論理知識型とデータ駆動型の融合（ハイブリッドAI）

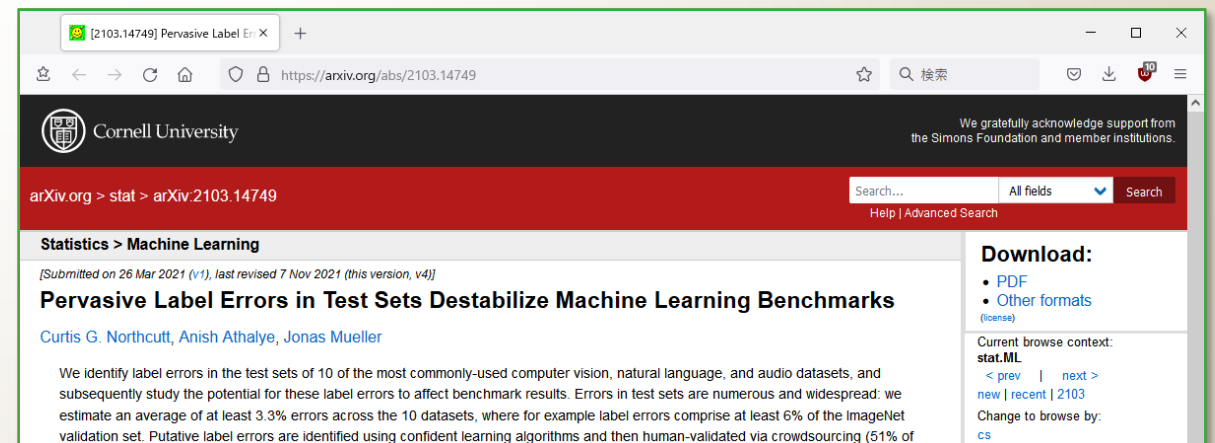
深層学習による学習と推論

- ➡ 学習データから作られる学習済モデルを使用して推論



- ➡ 学習データが重要

- ➡ 公開されたデータには誤りが含まれることも



オープンデータ×AI

➡ 華北交通アーカイブ

- ➡ 日中戦争期の中国北部の写真のデジタルアーカイブ

➡ AIによる

- ➡ 古写真の自動カラー化 (AI彩色写真)
- ➡ 古写真の自動タグづけ
 - ➡ 公開された学習済みモデル Inception-v3を利用
 - ➡ 自動カラー化画像を対象



<http://codh.rois.ac.jp/north-china-railway/>

★車載カメラとAIによる路傍の石造物調査

- ➡ AIで動画から石造物と文字、像を検出
 - ➡ 独自に収集した画像を用いて学習



- ➡ プレスリリース「[車載カメラで撮影した動画からAI（人工知能）で路傍の石仏や石造物を検出する技術を開発](#)」

★調査結果をマップで公開

➡メタデータと動画をオープンデータとして公開

➡メタデータ = データについてのデータ

➡動画の撮影日・場所

➡検出された物体

➡道路種別・名称

車載カメラとAIによる路傍の石造物調査

DATE	2021-02-24
ADDRESS	埼玉県比企郡ときがわ町玉川
OBJECTS	庚申塔、地藏
DETECT	jizo, stone, koshin
ROAD TYPE	一般道府県道 (secondary)
ROAD NAME	ときがわ坂戸線

メタデータ

ダウンロード

[735 records] メタデータ (CSV, GeoJSON - CC0)

<https://stone.midoriit.com/>

AI（機械学習）にとって学習データがすべて

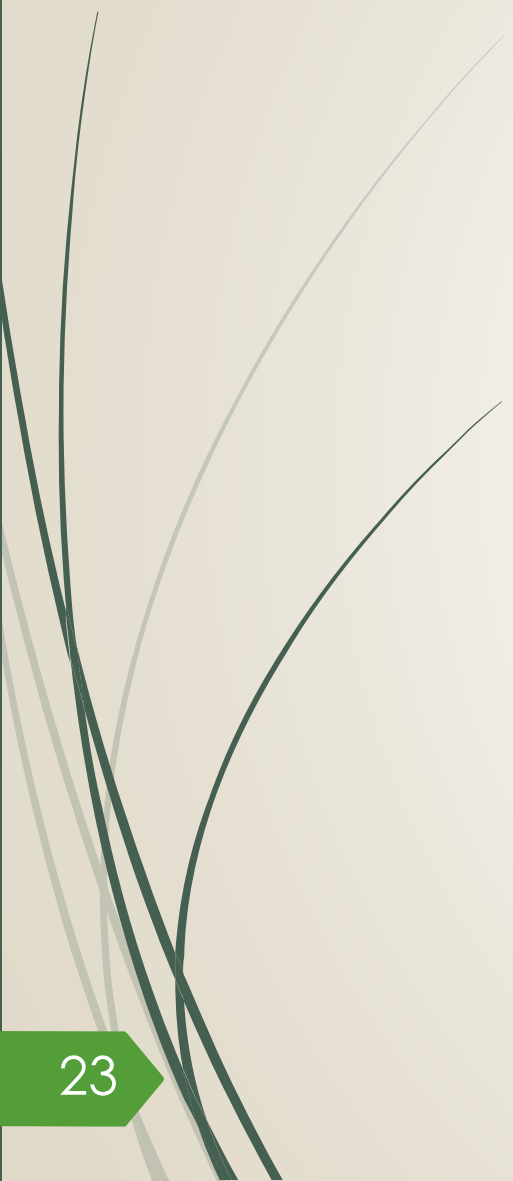

- ➡ データの偏りの危険性
 - ➡ 人種バイアス
 - ➡ ジェンダーバイアス
- ➡ 現状の反映→差別を助長
- ➡ AI=客観的 とは限らない



<https://jp.reuters.com/article/amazon-jobs-ai-analysis-idJPKCN1ML0DN>



<https://twitter.com/WSJJapan/status/616418552829079552>



第3部 オープン×市民

オープン〇〇の広がり

➡ 政府・自治体

➡ オープンガバメント

- ➡ 透明性の確保、市民の政治・行政参加

➡ OpenGLAM

- ➡ 文化施設（Gallery, Library, Archive, Museum）のオープン化

➡ 学術界

➡ オープンサイエンス

- ➡ 学術研究の成果・研究活動のオープン化

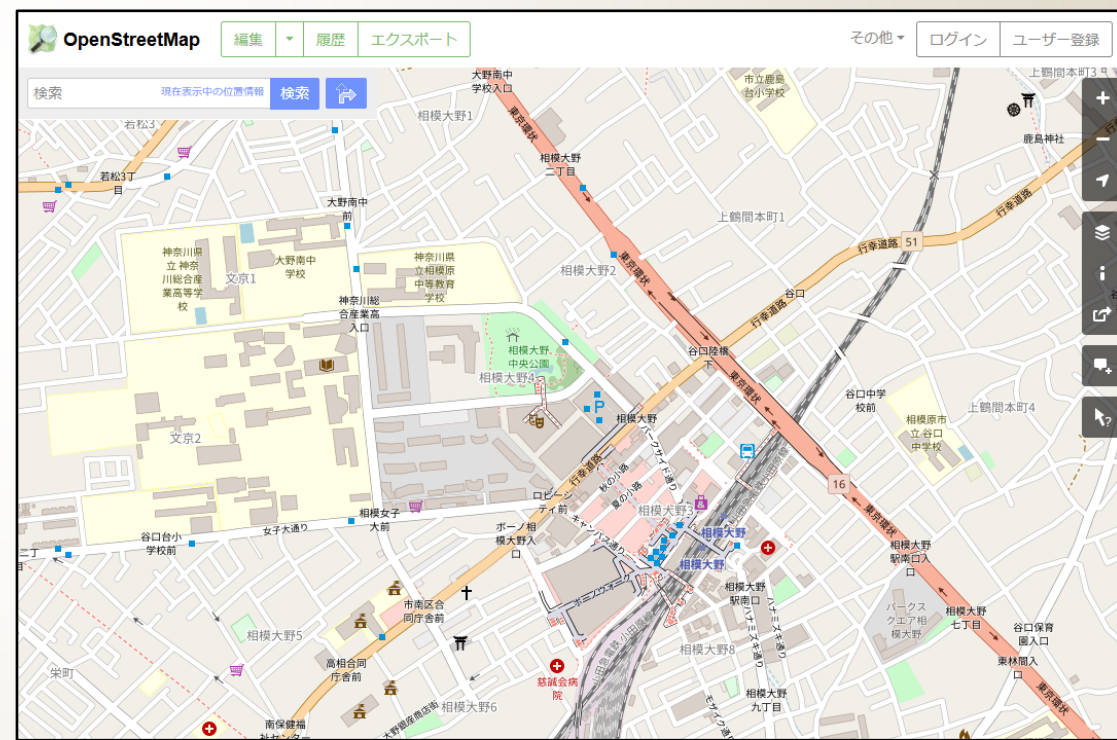
➡ オープンアクセス

- ➡ インターネットで誰でも無料で学術論文を読める

市民参加により創られるオープンデータ

➡ ウィキペディア、OpenStreetMapなど

➡ 誰でも編集できるフリーな百科事典／地図

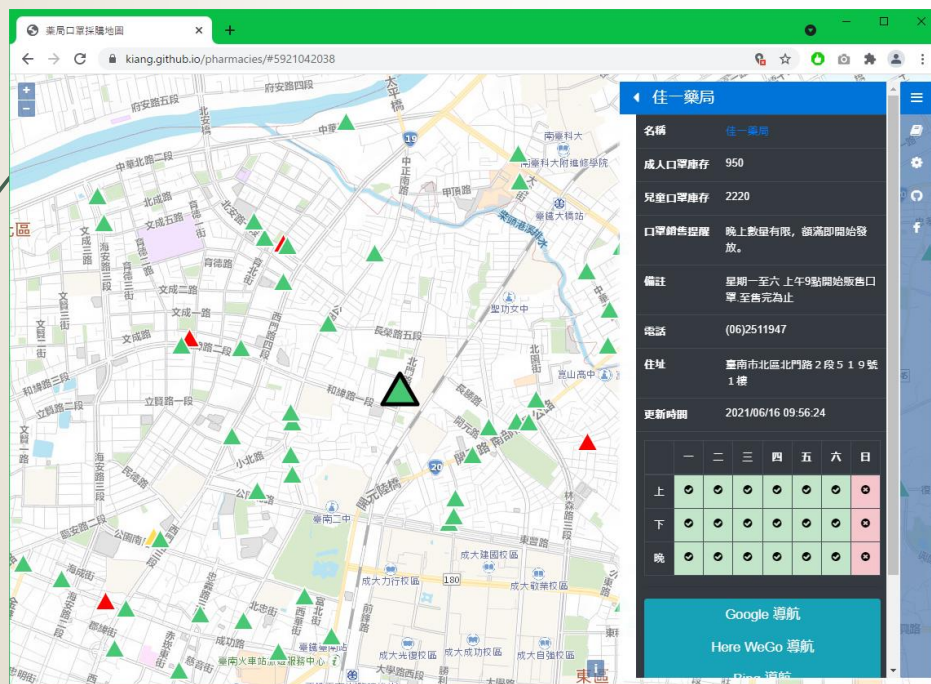


テキストはCC BY-SA
写真等はCC0、CC BY、CC BY-SA等

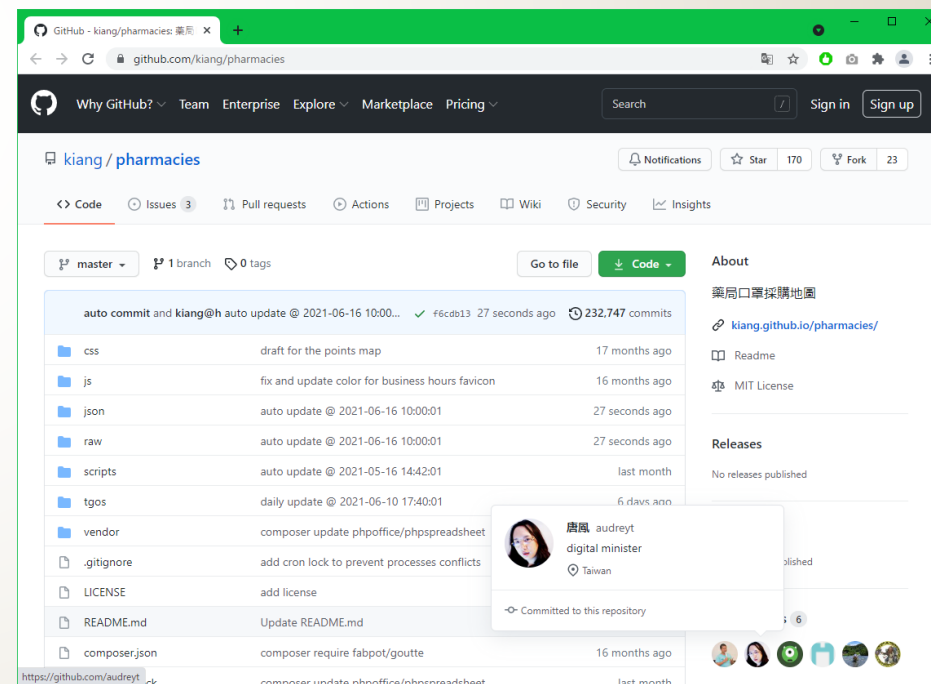
地図データはODbLライセンス
地図画像はCC BY-SA

コロナ禍で脚光を浴びた、シビックテック

- 市民が中心となり、ITを用いて社会課題を解決
 - オープンデータ、オープンソースの活用
- 台湾、マスクマップ、オードリー・タン氏



<https://kiang.github.io/pharmacies/>

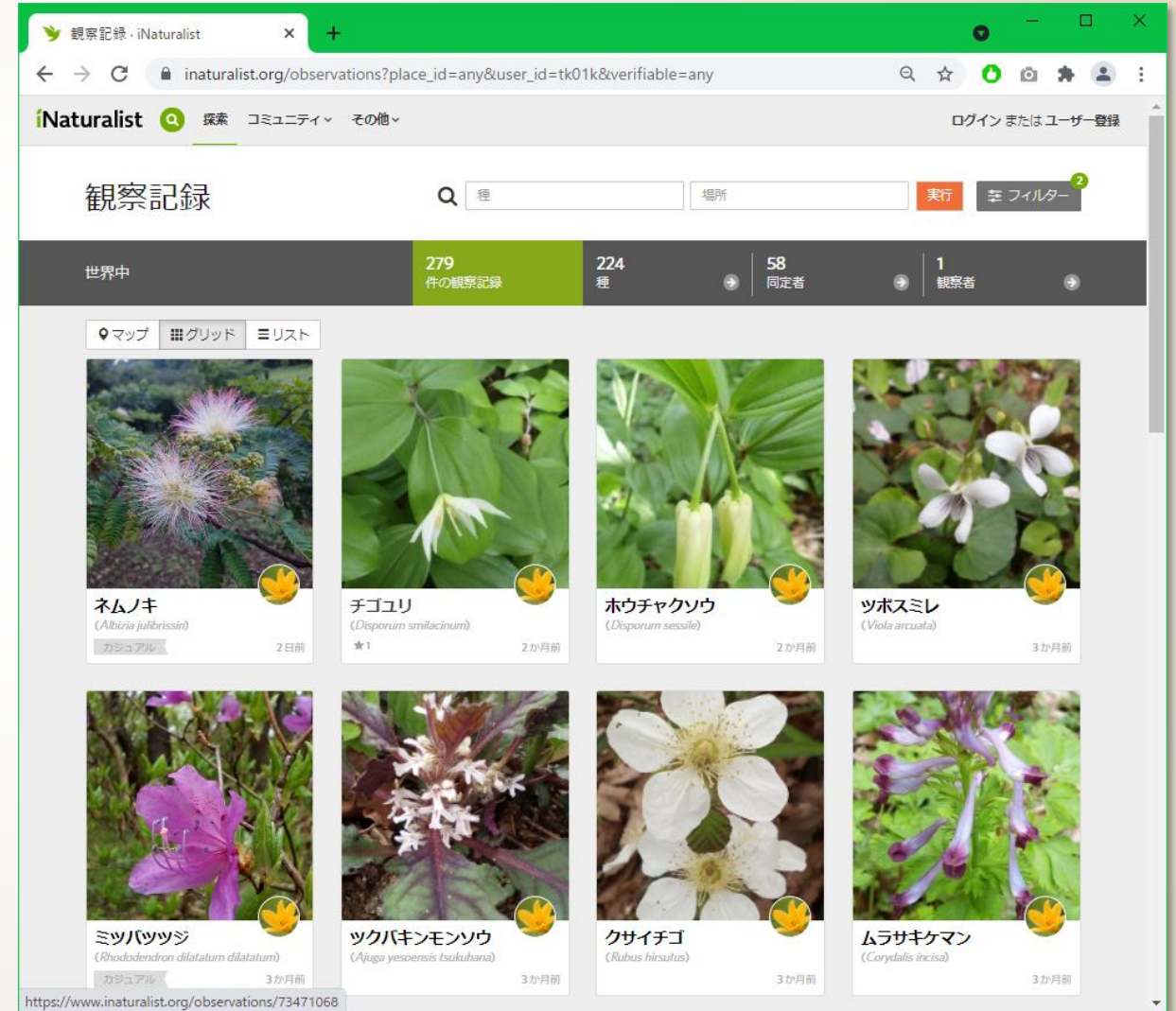


<https://github.com/kiang/pharmacies>

シチズンサイエンス（自然科学）

➡ iNaturalist

- ➡ 生物の観察記録を公開・共有
 - ➡ 種の同定にはプロの科学者も協力
- ➡ データのライセンスは参加者が選択
 - ➡ オープンデータとは限らない



シチズンサイエンス（人文科学）

▶ みんなで翻刻

- ▶ くずし字で書かれた古文書を活字化
 - ▶ AIくずし字認識がサポート
- ▶ 翻刻文データのライセンスはCC BY-SA

約8,000人の人々により
約2,000万文字もの
史料を翻刻

みんなで翻刻 MINNA DE HONKOKU

歴史資料を読み解く 歴史資料を解読するために お問い合わせ EN 参加する

みんなで翻刻 MINNA DE HONKOKU

仲間と、AIと協力して、
古文書解読に参加しよう！

JOIN 参加する

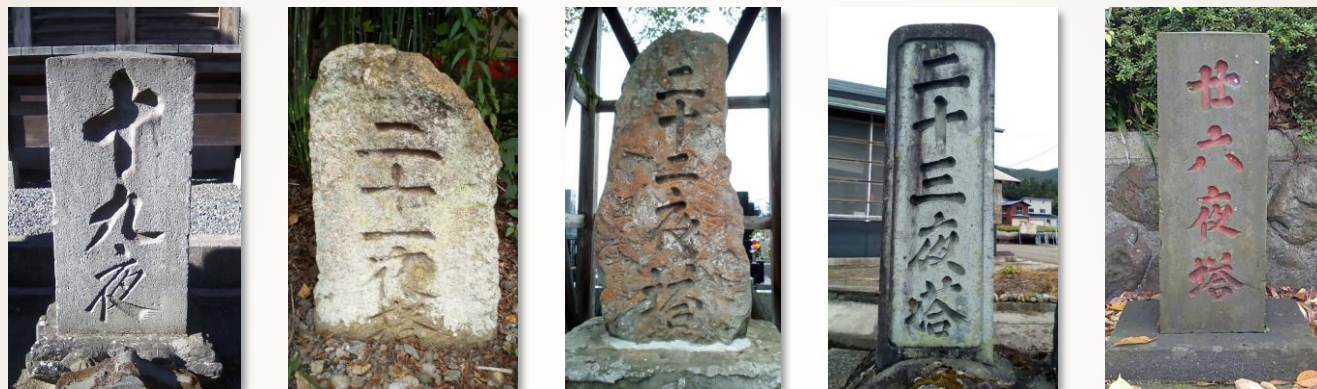
※現在新バージョンの試験運用中です。
安定稼働するまで旧バージョンも併せてご利用ください。

いいね! ツイート B! ブックマーク LINEで送る

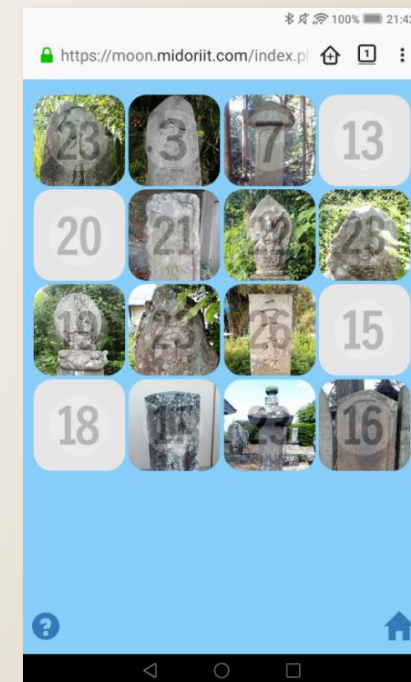
<https://honkoku.org/>

★月待ビンゴプロジェクト

- ➡ 月待塔 - 特定の夜に信仰行事を行う講中が造立した石塔

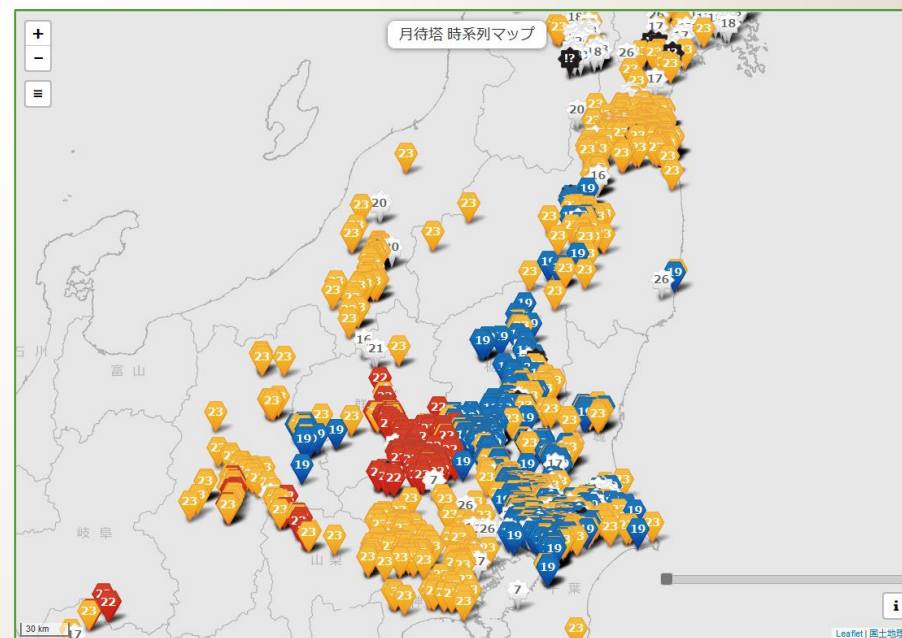
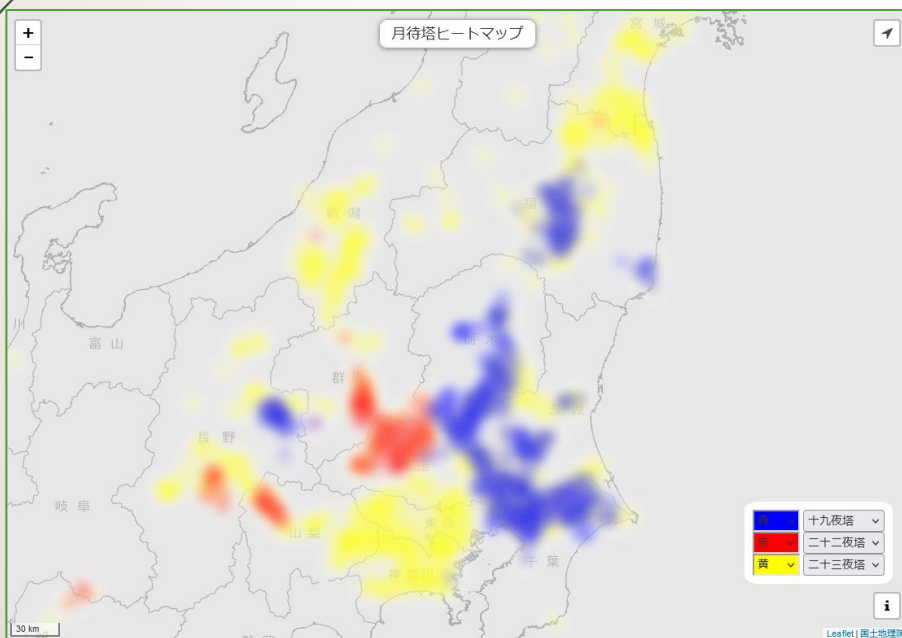


- ➡ Twitterを用いて「月待塔」の情報を収集・公開
 - ➡ 価値ある情報を膨大なツイートに埋もれさせない
 - ➡ 必要なのは #月待ビンゴ と位置情報と写真
 - ➡ ゲームフィケーション
 - ➡ Twitterの拡散力、コミュニティ形成力



★月待塔オープンデータを[GitHub](#)で公開

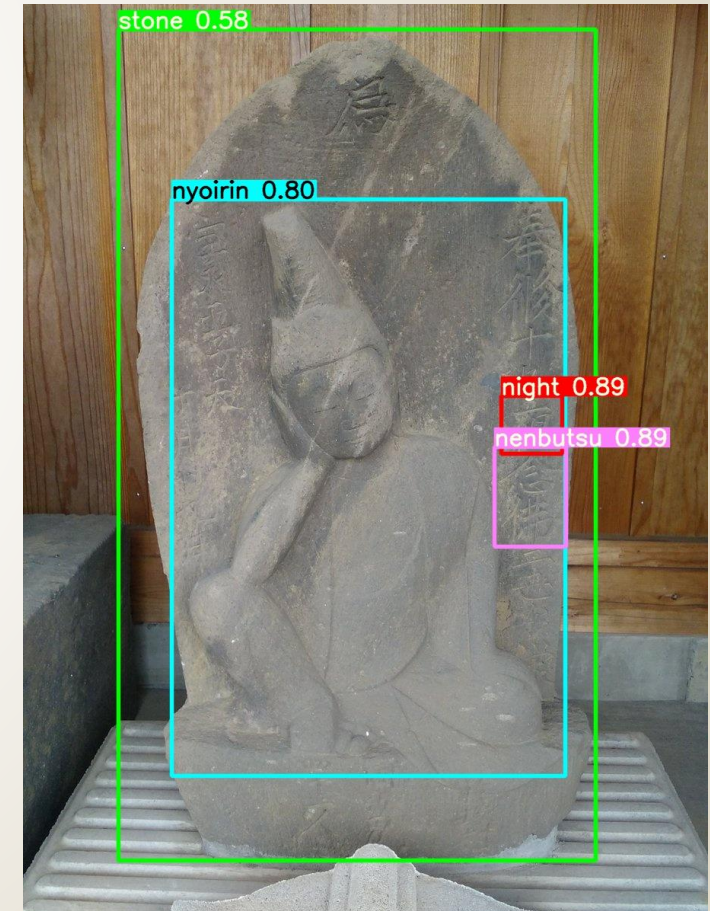
- ▶ 毎月10名前後の参加者による、累計4,800件超のデータ
 - ▶ Web地図上で[攻略マップ](#)として公開
- ▶ [ヒートマップ](#)、[時系列マップ](#)として可視化
 - ▶ [LODチャレンジ2019](#) データ分析・可視化部門 優秀賞



★AIによるデータエンリッチメント

- ➡ AIによる物体検出の結果を月待塔オープンデータに付加
- ➡ 「念佛」と刻まれた月待塔について分析
 - ➡ 月待塔のうち、十九夜塔に「念佛」が多い（既知）
 - ➡ ただし、「念佛」が多いのは関東の十九夜塔で、福島・長野の十九夜塔では少ない（新発見?）

第126回 人文科学とコンピュータ研究発表会
「深層学習を用いた石造物の検出と分類」
([スライド](#)、[論文](#))



まとめ

データとAIは今後さらに重要な役割を果たす

- しかし、課題も多い
 - データの信頼性
 - データ処理プロセスのブラックボックス化
- 大切なのは、オープンであること
- 市民の関与は不可欠
 - 情報リテラシー
- オープン化の流れを妨げてはならない

朝日新聞デジタル > 記事

自治体サイト受注めぐり他社を妨害か 2社に立ち入り 公取委

会員記事
田中恭太 2021年11月2日 17時40分

[list](#) 37 [メール](#) [印刷](#)



スマートバリューのウェブサイト

地方自治体が発注するウェブサイトの受注をめぐり、競合他社を妨害した疑いがあるとして、公正取引委員会は2日、システム開発会社スマートバリュー（本社・大阪市、東証1部上場）など2社に対し、独占禁止法違反（不公正な取引方法）の疑いで立ち入り検査に入った。関係者への取材でわかった。

関係者によると、2社は複数の自治体に対し、サイトの管理システムを発注する時の仕様書に「（ソフトの設計図が一般公開されている）オープンソースソフトウェア（OSS）で構築しているシステムは不可」との要件を盛り込むよう働きかけ、他社の受注を不当に妨害している疑いがある。OSSは、システムを開発する際に活用すればコストが抑えられるなどのメリットがあるとされる。

ご清聴ありがとうございました

midoriit.com