

# Project LINKS

国土交通省における分野横断的なデータ整備・活用・オープンデータ化プロジェクト

2025年8月

総合政策局 モビリティサービス推進課 総括課長補佐

Project LINKS テクニカル・ディレクター 内山 裕弥

## 自己紹介



### 内山 裕弥(UCHIYAMA YUYA)

国土交通省 総合政策局 公共交通政策部門 モビリティサービス推進課  
総括課長補佐

Project LINKS テクニカル・ディレクター

PLATEAU Advocates 2024

東京大学 工学系研究科 非常勤講師

東京大学 空間情報科学研究センター 協力研究員

1989年東京都生まれ。首都大学東京、東京大学公共政策大学院で法哲学を学び、2013年に国土交通省へ入省。

国家公務員として、防災、航空、都市など国土交通省の幅広い分野の政策に携わる。

法律職事務官として法案の企画立案や法務に長く従事する一方、大臣秘書官補時代は政務も経験。

2020年からはProject PLATEAUのディレクターとして立ち上げから実装までを一貫してリード。

2023年7月から国土交通省の横断的DXプロジェクト「Project LINKS」を立ち上げ。

2024年4月から現職。

#### 国土交通省略歴

2013年4月 総合政策局 政策課

2015年4月 水管理・国土保全局 水政課 法規係長

2017年7月 航空局 総務課 法規係長

2019年7月 大臣官房 大臣秘書官室 大臣秘書官補

2020年8月 都市局 都市政策課 課長補佐

2023年7月 総合政策局 情報政策課 IT戦略企画調整官

都市局 都市政策課 デジタル情報活用推進室

2024年4月 総合政策局 モビリティサービス推進課 総括課長補佐

# Project LINKS

---

## 1. Project LINKSの概要



POWER of DATA x DATA

## Concept

Project LINKSは、国土交通省の分野横断的なDX推進プロジェクトです。これまで活用されてこなかった様々な行政情報を「データ」として再構築し、これを活用できるようにすることで、データに基づく政策立案の推進(EBPM)や、新たなビジネス創出(オープン・イノベーション)の実現を目指します。

- 急速に進展する人口減少・少子高齢化に対応し、防災、交通、まちづくり等の多様な分野における**官民の生産性を向上**させるため、**政策やビジネスにおけるデータ活用や新サービス創出等を進める**必要がある。
- 幅広い施策・制度・手続を所管する国土交通省には**膨大な行政情報が蓄積**。これを「データ」として整備し、**官民が利用可能な環境を構築**することで、国土交通分野の基礎的な情報インフラを提供。
- オープンデータを利用したビジネス創出**や**政策立案におけるデータ活用**を促進し、社会全体の生産性向上を実現。

### 現状

国土交通分野の膨大な「情報」は保有しているが、利用可能な「データ」にはなっていない「宝の持ち腐れ」状態。

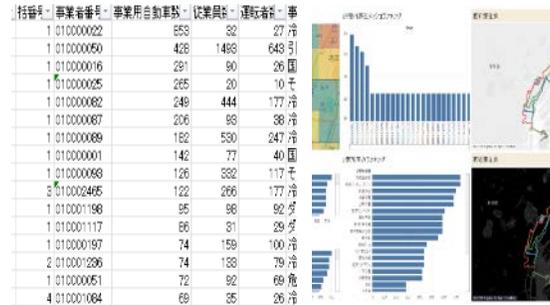


### 目指す姿

#### データ整備スキームの確立

生成AI(LLM)技術を用いてWord等のデータを機械処理・二次利用可能なデータに自動処理する仕組みを開発。

データ整備コストを劇的に低減。



膨大な国土交通分野の行政情報が利用可能に

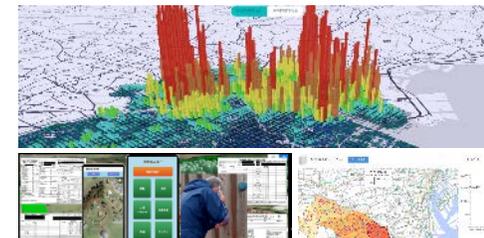


#### オープン・イノベーション創出



イメージ例:  
ドローン事故情報をオープンデータ化することで、ドローン事業者向けの事故情報検索・安全アドバイザリー サービスに活用

#### 政策立案におけるデータ活用



イメージ例:  
観光統計情報(アンケート)をデータ化することで、周遊促進施策の立案やオーバーツーリズム対策等に活用

## 1 | データの拡充



国土交通省保有の行政情報をはじめとする官民保有情報のデータ化を推進する。

## 1. 行政情報のデータ整備スキーム確立(データ作成実証)

- 行政情報をLLMを活用して自動処理でデータ化するための技術実証を実施
- 活用が見込まれる行政情報について各局から紙やword、excel等を収集して実際にデータ作成を検証
- 検証にあたってデータ構築基盤「LINKS Veda」のプロトタイプを構築

## 2 | EBPMの推進



データを用いて政策立案・執行・評価のプロセスを科学化する。

## 2. EBPMベストプラクティス創出(データ活用実証)

- 活用が見込まれる行政情報を調査し、ユースケースを企画立案
- 1.で作成したデータを活用した分析ツールやダッシュボード等のデータ活用環境のプロトタイプを構築
- 10以上の複数テーマ※でデータ活用環境を用いたユースケースの有用性検証を実施

※観光、フェリー、トラック、バス、ドローン、倉庫、空き家等。テーマにより国交省職員や地方自治体職員等が対象としたデータ活用を実証する。

## 3 | オープン・イノベーションの推進



官民のオープンデータを拡充して新たなサービスを創出する。

## 3. オープン・イノベーション創出(オープンデータ活用実証)

- 1及び2を踏まえて公開可能なデータについて、秘匿化処理を適用し、オープンデータ化を推進
- オープンデータ化に必要な秘匿化やデータ作成の手順をまとめたオープンデータ化ガイドラインの策定
- オープンデータを活用したサービスやビジネス創出を促進するため、ハッカソン等のイベントの企画・運営
- LINKSの取組を認知するためのWebサイト企画・運営やメディア活用等のPR活動

# 様々な国土交通分野の課題を解決するデータ活用ソリューション提供を推進

観光



ドローン



旅客船



海運



倉庫



空き家



公共交通



貨物  
運送



道路・  
河川



# MLIT

# 行政文書の現状と難しさ

令和〇年〇月〇日

〇〇運輸局長 殿

〇〇県〇〇市〇丁目〇番〇号  
XXXクルーズ株式会社  
代表取締役社長 XX YY

一般旅客定期航路事業許可申請書

海上運送法第3条第1項及び同法施行規則第2条の規定に基づき、〇〇～〇〇航路における一般旅客定期航路事業を經營したいので、関係書類を添えて申請いたします。

記

1. 住所及び氏名

住所	〇〇県〇〇市〇丁目〇番〇号	
氏名	XXXクルーズ株式会社 代表取締役社長 XX YY	

2. 役員の氏名

取締役会長	XX YY
代表取締役社長	XX YY
代表取締役専務	XX YY
取締役常務	XX YY



イメージ

- ✓ 様式が支局によってバラバラ
- ✓ 紙で提出されたものをPDFスキャンで保管
- ✓ 一部手書きで補完
- ✓ ハンコが丸囲いに重要な要素がある

# 行政文書の現状と難しさ

使用船舶明細書

船名	うさぎ	かめ	(新船)
船舶の種類	汽船	汽船	汽船
船質	鋼	鋼	鋼
進水年月	2012年2月	2013年3月	2014年4月(予定)
船舶所有者	うさぎかめフェリー㈱	うさぎかめフェリー㈱	うさぎかめフェリー㈱
総トン数	12,345ト	3,456ト	約7,890ト
貨物積載容量	12,34 m <sup>3</sup>	3,46 m <sup>3</sup>	なし
自動車航送に係る自動車積載面積	12,345 m <sup>2</sup>	3,456 m <sup>2</sup>	7,890 m <sup>2</sup>
旅客定員	600名	800名	200名
主機の種類	USAGI 12V12C	KAME 33V	USAGI 12V12C
連続最大出力	1,500kw × 2 + 2,000kw	500kw × 3 + 2,000kw	1,500kw × 2 + 2,000kw
航海速度	25.5ノット	28ノット	25.5ノット

使用する船舶

添付資料②-2  
第一号様式

使用船舶明細書

船名	さる	
船舶の種類	汽船	
船質	FRP	
進水年月	令和元年1月	
船舶所有者	㈱さるかにドッグ 船舶借入 一般社団法人鬼ヶ島	
用途	旅客船	
総トン数	100t	
貨物積載容量		
自動車航送に係る自動車積載面積		
定員	旅客	100人
	船員	10人
	その他の乗船者	0人
主機の種類	ディーゼル	
連続最大出力	100PS × 2基 (1000馬力)	
航海速度	20ノット	

✓ 申請項目や審査事項は法令等によって規定されているが、表現方法としての様式や入力規則は管轄区域や事業者によって大きく異なる

→ 単純なOCRやRPAでは規格化されたデータ抽出ができない

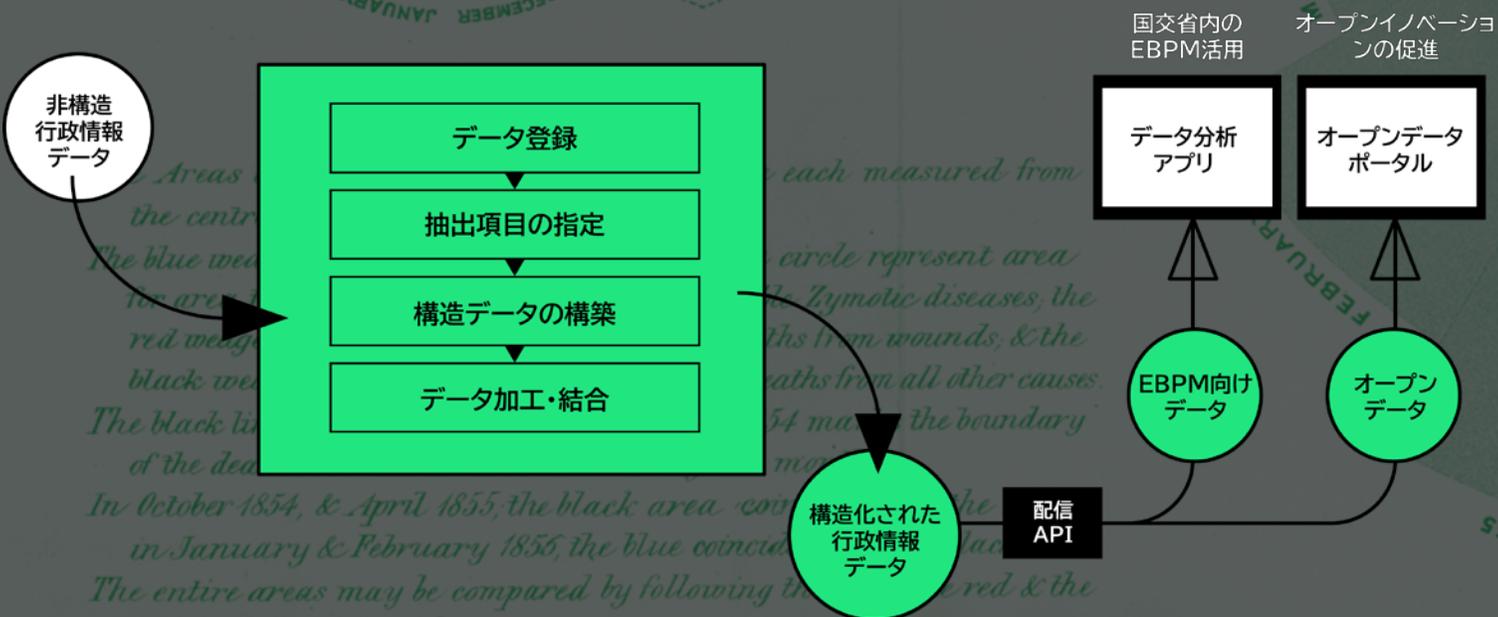
イメージ

# LINKS Veda's Overview

## LINKS Vedaとは

LINKS Veda(ヴェーダ)は、LLM(大規模言語モデル)を用いて自然言語を解析し、非構造データから意味情報を抽出。指定されたカラムに格納することで、テーブルなどに構造化されたデータを自動生成するシステムです。

国土交通省が保有する膨大な行政情報をデータ化し、誰もが探索可能なデータアクセス基盤を実現します。



## アクロニム

Veda:  
**V**erbal **E**xploring system for **D**ata **A**ccess

### Verbal

LLMを駆使したVedaを用いることで、生成モデル(AIがデータを学習し、予測、新たなデータとして生成するモデル)やプロンプトをノーコードで(口語で)誰でも直感的に活用できるようになります。

### Exploring

画像や文章などの非構造データを大規模言語モデルが「探索」し、意味を抽出。機械判読可能なデータとして抽出します。

### Data Access

非構造データを構造化データとして再構築、再生成する仕組みは、これまで「宝の持ち腐れ」となっていた大量の行政情報を「データ」として生まれ変わらせ、アクセス可能とします。

# 様々な国土交通分野の課題を解決するソリューション提供を推進中

観光



ドローン



旅客船



海運



倉庫



空き家



公共交通



貨物  
運送



道路・  
河川



# MLIT

# 多様な行政情報のオープンデータ化を推進中

#一般旅客定期航路事業データ



#無人航空機飛行計画データ



#貨物自動車運送事業データ



#内航海運業事業データ



#貨物船、貨物鉄道データ



#国土交通省組織管内図データ



#倉庫業データ



#自動車運送事業事故データ



#GTFS





# ABOUT

「交通空白」解消など地域交通の  
「リ・デザイン」の全面展開を進めるため、

サービス

データ

マネジメント

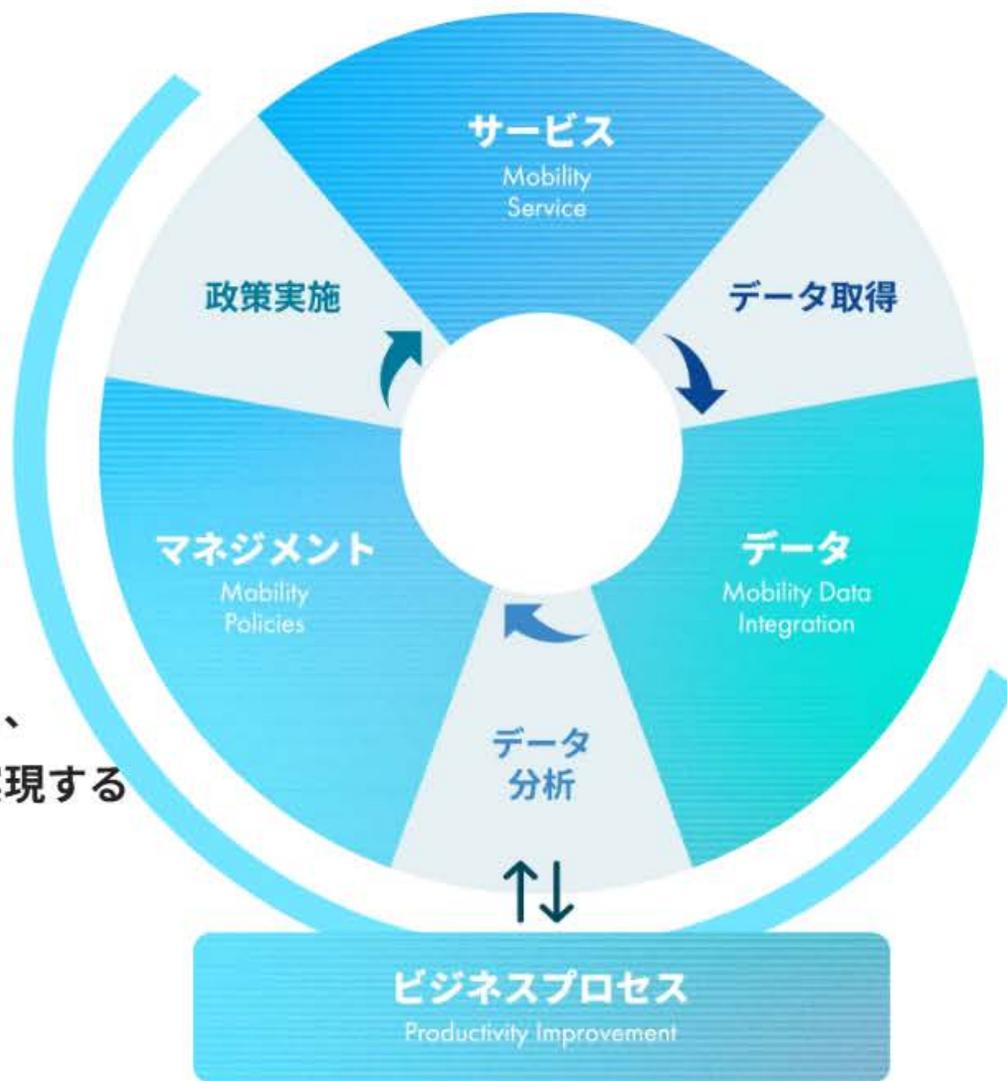
ビジネスプロセス

の4つの観点からデジタル活用を一体的に推進し、  
地域交通の持続可能性、利便性、生産性向上を実現する

地域交通 DX 推進プロジェクト

**COMmmONS** (コモンズ)

を新たにスタートします。



## Googleマップ経路検索へのデータ掲載を支援する施策(モビ課)

- 地域交通におけるデジタル技術活用を推進するモビリティサービス推進課では、GTFSデータの整備及びGoogleマップをはじめとする各種乗換案内サービスへの交通情報の掲載を支援するため、以下の施策等を実施しています。

### ① データ整備の財政支援

- 「日本版MaaS推進・支援事業」において、**MaaSの基盤的データとしてGTFS整備の支援を実施中。**

※ FY2025予算規模:20億円程度を予定



継続実施中

### ② GTFS国内標準アップデート/先端技術実証

FY2025新規

- GTFSの「作りやすさ」と「品質」を向上させるため、**国交省が策定する国内標準「GTFS-JP」のバージョンアップ**を実施予定。
- デマンド交通(タクシーやデマンドバス)に対応した**新規格「GTFS-Flex」等の技術実証**を実施予定。(LINKS連携)



### ③ 官民連携によるデータ整備支援

- 「公共交通オープンデータチャレンジ」等の**官民連携イベントを通じて全国でGTFS整備を促進。**
- GTFSコミュニティと連携した**GTFS整備研修・情報発信イベント等を実施。**(LINKS連携)



継続実施中

### ④ データ整備ツールの提供

FY2025新規

- 中小事業者等によるGTFS整備を支援するため、**職員によるGTFS内製ツールを開発・オープンソース化**予定。(LINKS連携)
- ※ フリーソフトやマニュアル等は既に存在するが、使い勝手や更新性が悪く、「得意な人」がいなくなるとデータ更新が止まる現象が頻発。





## Point

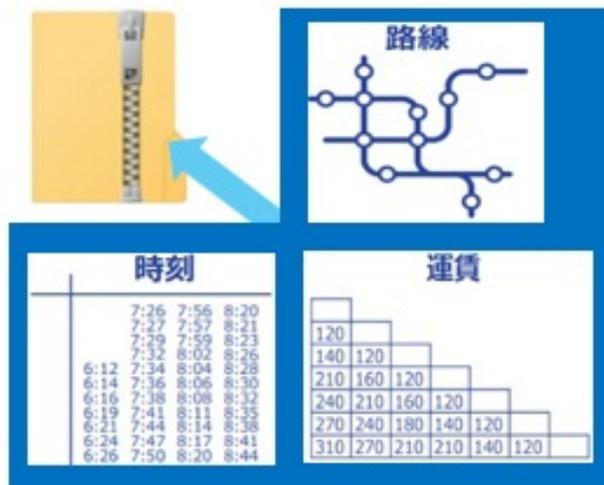
- GTFSの国際標準の動向と国内の整備・活用状況、学術における活用実態等を踏まえ、GTFS国内標準 (GTFS-JP)のアップデートを行う。
- GTFS品質の向上や整備促進により、乗換案内や様々なアプリによる活用を推進。

## スコープ

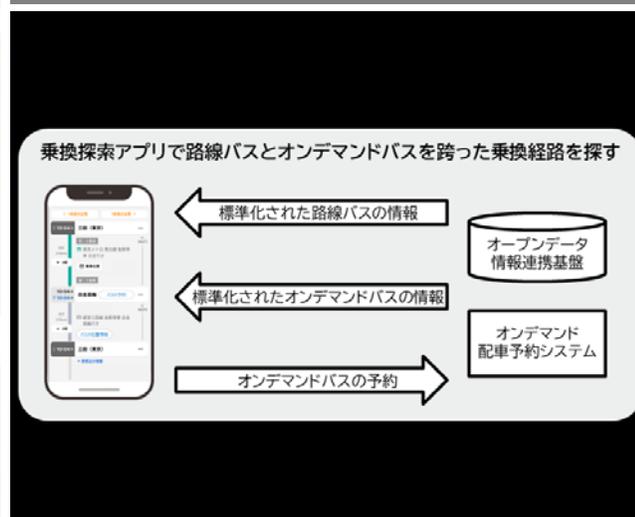
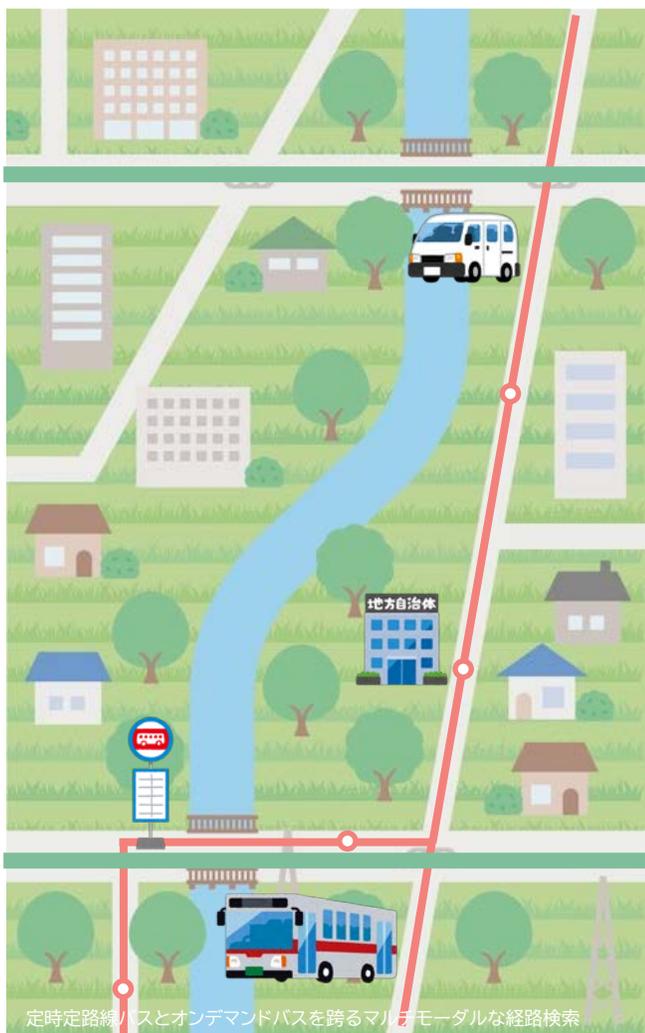
- GTFS-ScheduleやGTFS-RT、GTFS-Flex等の最新の国際標準の国内標準への取込みを図るとともに、利用実態を踏まえたデータ項目の再編や入力規則の策定等によりデータ品質向上とデータ整備の容易化を推進。
- GTFSデータの整備者・利用者の両者に分かりやすく明確な技術解説資料を作成するとともに、普及活動を実施。

## 提供価値

- 国内標準の最新化や技術解説資料の公開により、交通事業者や自治体等によるデータ作成の促進や、データ作成システムの開発の活性化を実現。
- 品質の高いGTFSデータの整備・公開を促進することで、GoogleMapsをはじめとする多様な情報提供サービスへの地域交通情報の掲載を促進。



# GTFS-Flex及びGTFS-Ondemandの技術実証プロジェクト | TIS(株)



## Point

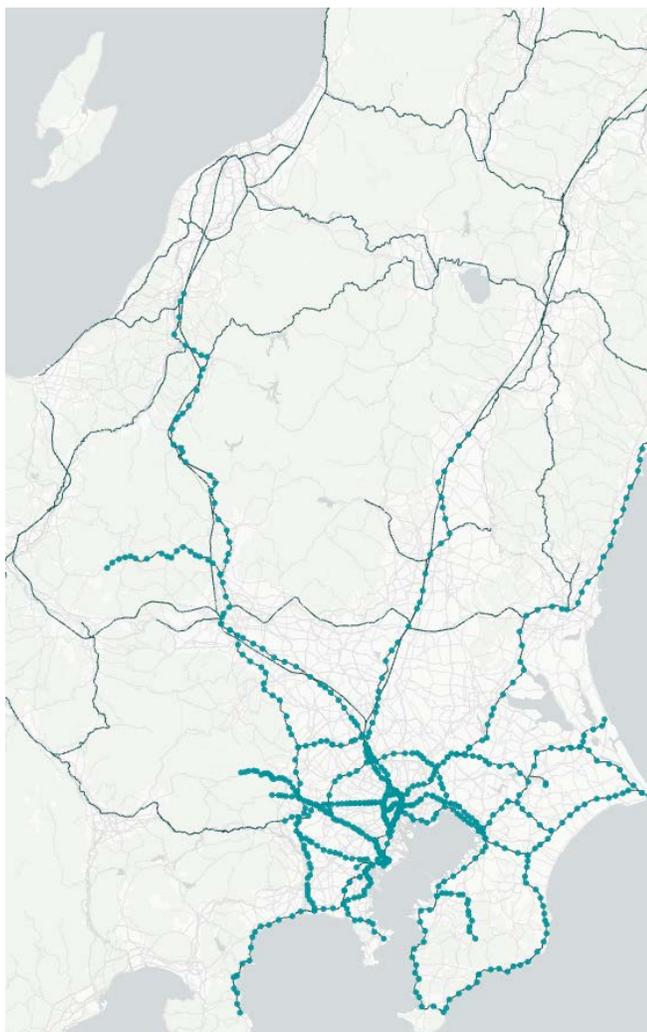
- オンデマンドバス等のデマンド型交通の経路検索を可能とする新たなデータ形式GTFS-Flex及びGTFS-Ondemandの技術検証を実施。
- 検証成果をオープンナレッジ化することで、デマンド型交通の経路検索サービスへの掲載を促進。

## スコープ

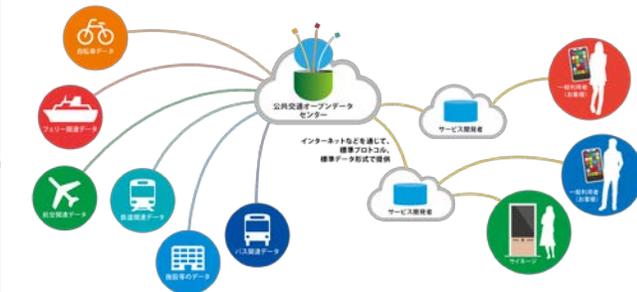
- 国際標準規格GTFS-Flex及び技術提案規格GTFS-Ondemandのデータ有用性やアプリ実装の実用性等を検証するためのデータ整備及びアプリ開発を実施。
- サービスの有用性検証のため、札幌市において定時定路線バスとオンデマンドバスを対象とした実装レベルに近いサービス提供を行う。

## 提供価値

- 経路検索サービスからオンデマンド型交通へのアクセスを改善することで、認知に課題のあるデマンドバス等の収益性を改善するとともに、定時定路線型交通と組み合わせたマルチモーダルかつシームレスな移動を実現。
- 区域運行型交通の経路検索を可能とする最先端規格GTFS-Ondemandの実証成果を公開することで、国際的な技術革新へも貢献。



公共交通オープンデータチャレンジ2024 -powered by Project LINKS-



## Point

- GTFSデータ等の整備・活用を促進し、地域交通におけるイノベーションを創出するため、オープンデータを利用したアプリコンテストを開催。
- 鉄道、バス、航空、フェリー、シェアサイクル等多様な事業者の協力を得て大規模なデータ整備・公開を実施。

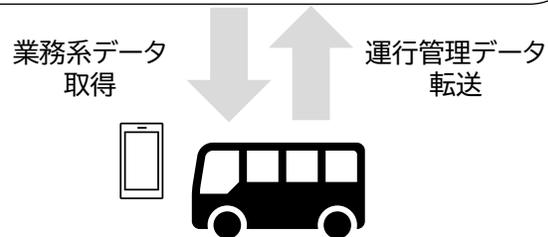
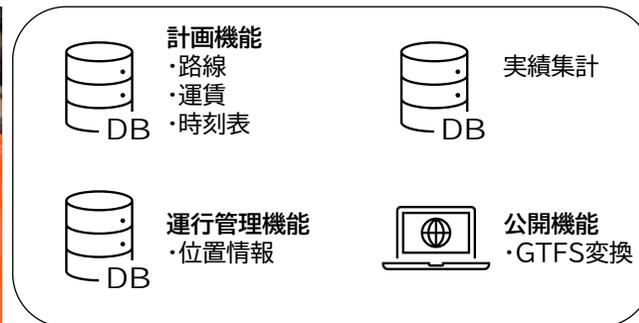
## スコープ

- 公共交通オープンデータ協議会と連携し、公共交通分野のオープンデータを活用したアプリ開発コンテスト「公共交通オープンデータチャレンジ」を開催。
- コンテストを契機とし、事業者・自治体向けウェビナーの開催、関連イベントの実施、事業者との連携等を行い、オープンデータの整備と活用を促進する。

## 提供価値

- オープンデータの活用を活性化させることで、地方公共団体や交通事業者によるGTFSデータ等の整備とオープンデータ化の意義の認知拡大を図る。
- 交通領域以外の多様な技術分野の参画を促すことで、地域交通における新たなサービスや業務改革などオープン・イノベーションの創出を促進する。

## コミュニティバスキット開発プロジェクト | (株)Will Smart


 Point

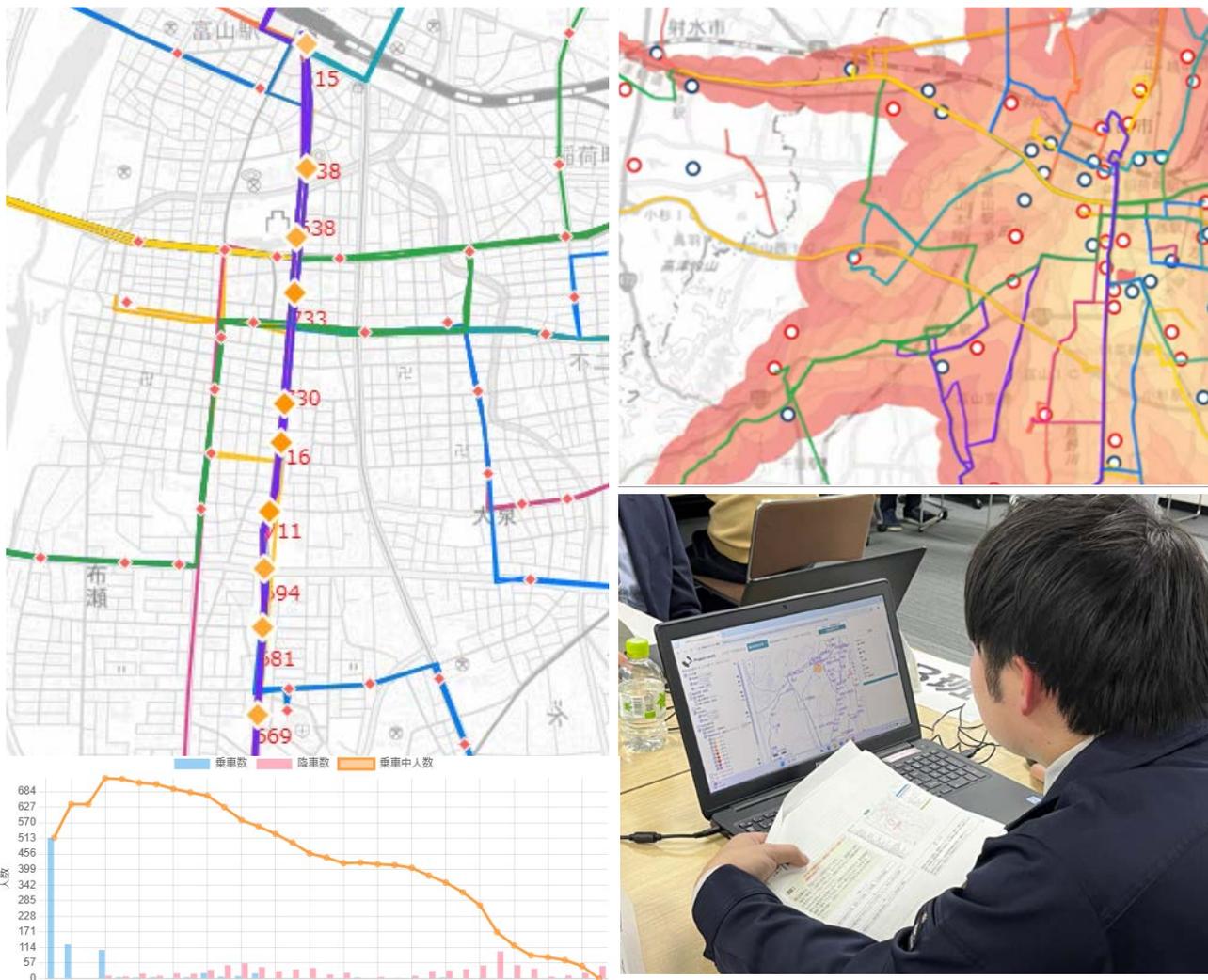
- 地方の小規模バス/コミュニティバスの運行効率化と経路検索アプリへの情報掲載促進を図るため、ダイヤ編成やGTFS出力を可能とする支援ツールを開発。
- 開発成果をオープンソースとして公開し、全国へ横展開を推進。

## スコープ

- 1台~数台程度の小規模事業を想定し、事業運営に必要なダイヤ編成機能や運行実績管理機能、売上実績管理機能等を持つツールを開発し、オープンソースとして公開。
- 特に地方部の観光地等において経路検索アプリへ情報が掲載されない問題に対応するため、ダイヤ編成機能とGTFS出力機能を連携させ、簡易にデータ整備可能な環境を構築。

## 提供価値

- 安価・簡易な利用環境を整備することで、DX投資に余力のない小規模事業者における業務のデジタル化・負荷軽減を促進。
- 業務支援システムとGTFS出力機能をノーコードツールとしてセットで提供することで、GTFSデータ作成の属人化を回避し、持続可能なデータ整備を実現。



### Point

- GTFSデータや乗降実績データ等を活用し、ブラウザ上で容易に地域交通の現状可視化や分析を可能とするウェブツールを開発。
- 開発成果はオープンソースとして公開。自治体職員等がデータを活用した地域公共交通計画を立案するための環境を整備。

### スコープ

- 到達圏分析やOD輸送量など、地域交通計画の検討に必要な諸情報を地図、グラフ、表で簡単に可視化できるツールを開発し、オープンソースで提供。
- バスのルートや停留所の配置、運行本数変更等の検討に必要なシミュレーション機能を簡易なUI/UXで提供し、非専門家でもデータ分析可能な環境を構築。

### 提供価値

- GTFSデータや標準的な乗降実績データなど、標準化されたデータを入力可能な汎用インターフェースをノーコードで提供することで、地域交通政策の検討におけるデータ活用のハードルを低減。
- 施策効果のシミュレーション等の意思決定支援機能を提供することで、検討資料作成や説明にかかる時間短縮や品質向上(EBPM)を実現。

# Project LINKS

---

## 3. LINKS Vedaの技術

### 3. LINKS Vedaの技術

LINKS Vedaは、以下の特徴を具備することで、ユーザーに快適なデータ生成体験を提供します

#### 特徴

1

ノーコードでのデータ生成・データ加工

- データ生成にかかる処理を概念化することで、ユーザーが直感的に操作できる
- 処理をテンプレートとして保存することで、誰でも簡単に再帰的にデータ構造化ができる

2

AI技術を活用したデータ生成処理

- LLM、OCR、NLP、ベクトルDBの利用により、非構造データをデータ化し、システムで利用可能にする
- GISの知識、ツールがなくともUI上の簡易な操作のみで地図で可視化可能なデータ生成ができる

3

セキュアな基盤の提供

- 厳格なアクセス制御や監査ログ、リアルタイムログ収集によるシステムの安全性を確保
- データ量やアクセス数に応じたスケーラブルなシステム構成と運用体制の維持



[開発中]

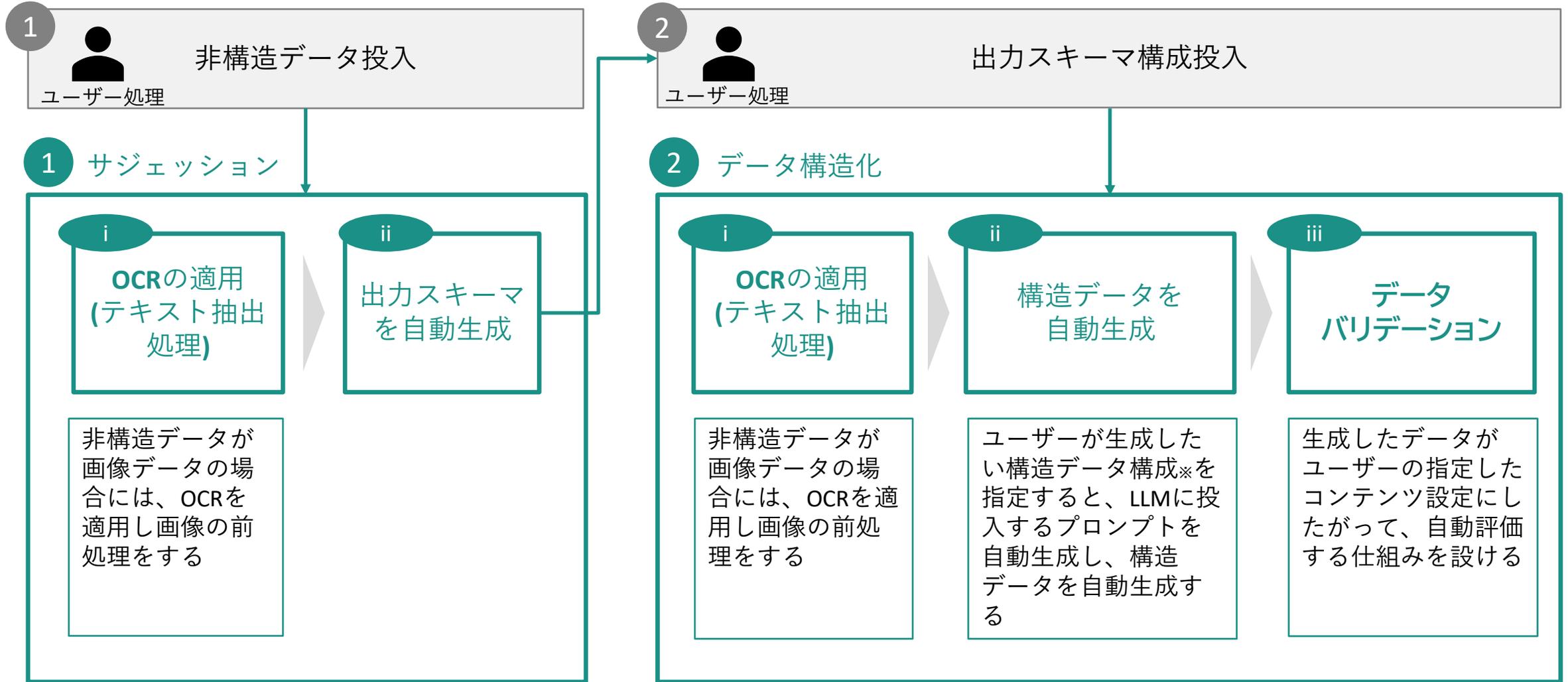
## AI技術を活用したデータ生成処理

LINKS Veda(プロトタイプ版)では、主に以下のAI技術を活用しています

API名称	モデル名称	説明	選定理由
Amazon Bedrock API  Claude 3	<b>Claude3.5 Sonnet</b>	データ構造化処理において、ユーザーの指示に基づき、指定した非構造データから該当する情報を抽出し、構造データ化に利用。また、チャット機能において、ユーザーからの問い合わせ文に対して、関連する文書からの回答生成処理に利用。	日本語処理性能や長文に対する処理能力が高く、LLMリーダーボード上も言語理解能力・応用能力・アライメントが上位にあり、東京リージョンにおいて利用可能であることから、本プロジェクトに適していると判断。
	<b>Amazon Titan Text Embedding v1</b>	チャット機能において、RAG(外部知識を活用した情報抽出・回答)のために文章を数値ベクトル表現変換(Embedding処理)に利用。	大規模な事前学習データセットで学習されており、Embedding処理において、テキストの意味情報を捉えやすく、日本語文章にも対応しているため、本プロジェクトに適していると判断。
Azure AI Vision API 	<b>Azure AI Document Intelligence (version 3.1)</b>	データ構造化処理において、手書きが記載されたPDF資料などを用いる場合に利用。(OCR処理)	日本語の文字認識精度が高く、領収書、請求書、名刺など一般的な書式に対する事前学習済みモデルを利用可能。文書中の段組み・改行・テーブルやチェックボックスなどレイアウト構造の取得が可能であり、本プロジェクトに適していると判断。

## AI技術を活用したデータ生成処理

LINKS Veda内では、以下のフローをシステム内で自動処理しデータ構造化を実現します



※カラム構成、データ型等の詳細条件

## AI技術を活用したデータ生成処理

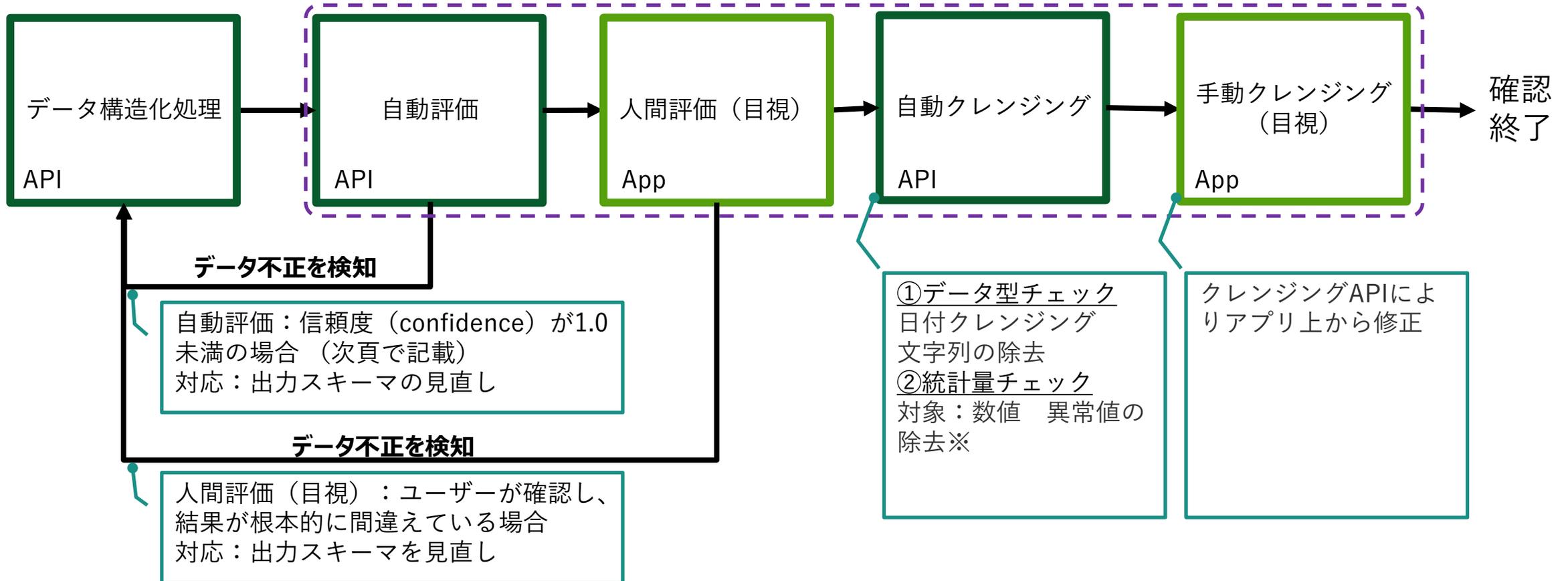
### 2 データ構造化プロセス

iii

### データバリデーション プロセス(全体フロー)

データ構造化プロセスでは、自動処理と手動処理を組み合わせたバリデーションプロセスを経る。不正を検知した場合は再生成、データ修正を実行する。

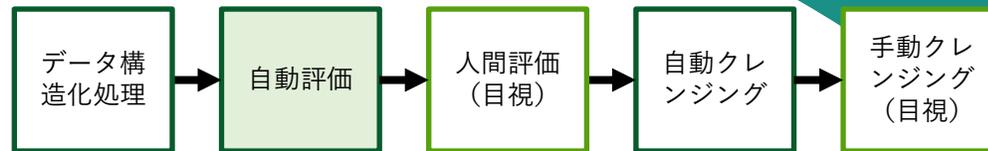
#### バリデーションプロセス



AI技術を活用したデータ生成処理

2 データ構造化プロセス

iii データバリデーションプロセス(自動評価)



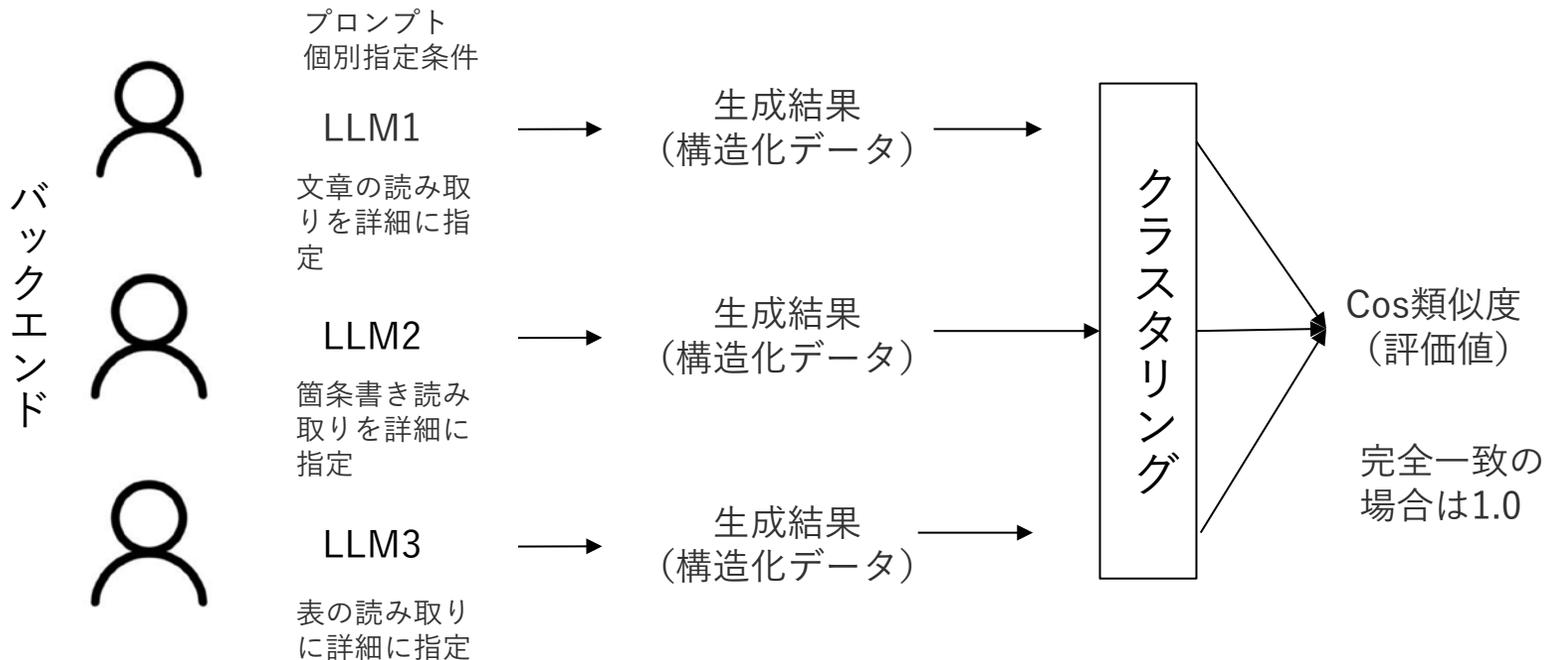
- 生成したデータがユーザーの指定したコンテンツ設定にしたがって、自動評価する仕組みを設ける。ハルシネーションを防ぐため、対象データソースごとに適切に生成されたかを定量的に示す信頼度confidence（定義域：0-1.0）を算出する。

プロンプト | 評価結果 | ユーザーへの通知

レイアウトの違いにより抽出ミスを検知できるようにLLMごとに個別条件を設定

複数LLMを用いて、同様の結果が返されるかで結果の妥当性を評価

ハルシネーションが起きる場合には、LLMによって生成結果が異なるため、複数LLMを用いることで検知



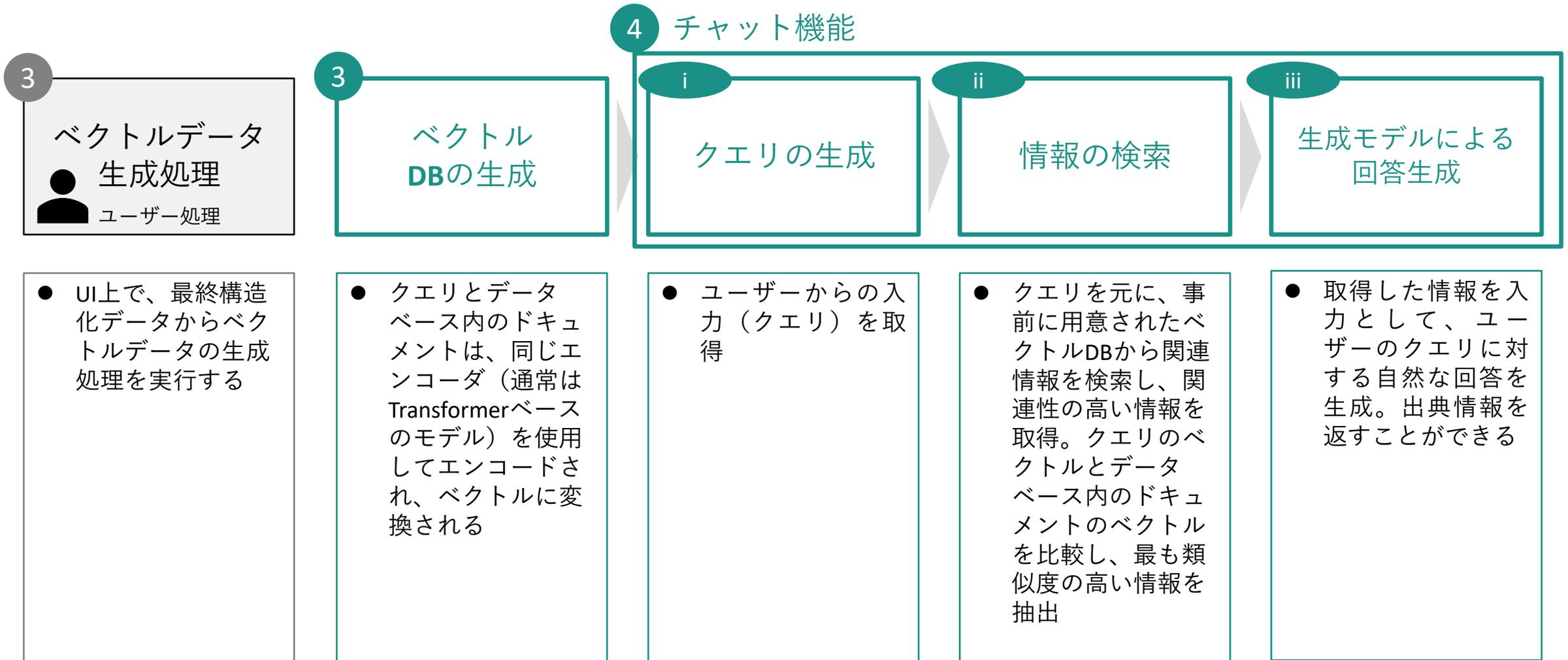
ID	FileName	2020 (R2) 年 目標の旅行消費額(全体)	2020 (R2) 年 目標の旅行消費額(インバウンド)	2020 (R2) 実績の旅行消費額(全体)
0001	【下呂温泉観...	32000	162000	20965
0002	【飛騨・高山...	161000	32000	135000
0003	【飛騨・高山...	32000		32493
0004	【飛騨・高山...	90000	10200	46577
0005	【中央日本総...	90000	20965	46577

参考：<https://docs.databricks.com/ja/generative-ai/agent-evaluation/llm-judge-metrics.html#llm-judge-trust>

[開発中]

## AI技術を活用したデータ生成処理

Veda内では、以下のフローをシステム内で自動処理しチャット機能を実現します



## LINKS Veda開発の振り返りと課題

## 本格実装に向けた課題

FY2024の業務では、システムの最小限の機能要件をおおむね満たしていたものの、一般ユーザーの利用や多様かつ大量な行政データの処理をスムーズに行うためには、以下の課題解消が必要。

課題	概要
1. データ品質	<ul style="list-style-type: none"> <li>データによってシステム自動処理による抽出ができないものや精度が低いものもあり、ユーザーによる手動修正が大量に入ってしまう場合がある</li> </ul>
2. ユーザービリティ・UI/UX	<ul style="list-style-type: none"> <li>概念のモデリングが複雑であり、データと処理の関連性の理解が難しくユーザービリティを下げている</li> <li>専門用語が多かったり処理手順が煩雑なため、直感的に処理できるUIになっておらず、十分なUXを提供できていない</li> <li>データ登録時のエラーチェックがないため、ルールに則らないデータも登録されてしまう。結果、本来は不要な修正作業が発生してしまう</li> <li>ノーコードで操作できる汎用的なBIツールの作りこみに至っていない</li> </ul>
3. パフォーマンス	<ul style="list-style-type: none"> <li>処理速度、スループット(処理能力)、スケーラビリティ(最大同時実行数など)、省内全職員での利用を想定した場合に耐えるシステム構成になっていない</li> <li>大容量データを扱う際に画面が重くなる、フリーズするなど、ユーザー操作を阻害する場合がある</li> </ul>
4. システム運用体制・ガバナンス	<ul style="list-style-type: none"> <li>オープンデータの運用体制(フロー、チェック体制など)が明確に定まっていない</li> <li>システム管理者、一般ユーザーといったミニマムの権限管理のみで、原課レベル、データレベルでの細かい権限管理ができていない</li> </ul>

# LINKS Veda開発の振り返りと課題

## FY2025の取組 | データ品質の向上

行政文書特有のレイアウトや図形要素、イレギュラーな記載パターンを正常に処理できるよう、OCRモデルとLLMのファインチューニングによる精度向上を目指す。

No	課題詳細	対応策
1	OCRの前処理として該当箇所の画像を切り出して個別処理するなどの精度向上を行ったが、誤認識やご抽出のケースが依然として存在する	<ul style="list-style-type: none"> <li>• <b>OCRモデルのファインチューニング:</b> 行政文書特有のレイアウトや図形要素に対応するため、OCRモデルのファインチューニングを導入し、追加学習を行う必要がある。これにより、丸囲みや矢印などの特殊な視覚要素の認識精度向上が期待できる</li> <li>• <b>信頼度スコアの導入:</b> OCRの結果ごとに認識結果の妥当性を評価する信頼度スコアを設け、ユーザーに判断材料を提供する。これにより、自動抽出が不確かな箇所をユーザーが容易に特定し、必要に応じて手動で修正できるようになる</li> </ul>
2	訂正線などのイレギュラーなパターンでは誤検出が目立ったため、積極的にNULL値を採用した。その結果、人間による手動修正に頼ることになり、操作負荷が上がってしまった	<ul style="list-style-type: none"> <li>• <b>画像認識アプローチの導入:</b> プロンプトによる制御には限界があるため、訂正線や修正内容を検出するための専用の画像認識モデルを導入する。これにより、テキスト認識の前段階で訂正パターンを特定し、適切な処理を行うことが可能となる</li> <li>• <b>ファインチューニングの実施:</b> 行政文書特有の訂正パターンに対応するため、OCRモデルやLLMのファインチューニングを行い、イレギュラーな記載に対する認識精度を向上させる</li> </ul>
3	LLMはプロンプト次第で汎用的にデータ抽出が可能である一方、抽出内容やボリュームによっては、一度の実行だけでは抽出ミスが起こりやすい。	<ul style="list-style-type: none"> <li>• <b>LLM実行の分割:</b> この場合に複数回に分けて実行することが望ましい。試験的な抽出を行った結果、自動的に実行単位を分割、プロンプトを設定できる機能を設ける(UC14一般旅客船やUC20自動車事故)</li> </ul>
4	データバリデーションにおいて、複数のLLMを用いて抽出結果の投票型のチェックを行っているが、抽出精度の向上が必要である。	<ul style="list-style-type: none"> <li>• <b>データバリデーション精度向上:</b> データ抽出精度向上のためのLLMのファインチューニングを行う</li> </ul>

**FY2025ではLLMハッカソンを実施予定!**

**プロンプトエンジニアリングやアプリケーションサイドの技術・アイデア、OCR技術、データ抽出技術などなど・・・幅広いプレイヤーとの協業を募集中です!**

# R7年度 オープンイノベーション創出に向けたイベント(予定)

LINKSのオープンコミュニティを醸成し、オープンデータの利用促進やイノベーション創出の活性化を図ることを目的として、多数のコミュニティイベントを開催。

## ハッカソン

### LINKS:POWER of DATA × DATA 2025

PRイベント	8月29日(金)
アイデアソン	10/25(土)
ハッカソン本番	11/29(土)~11/30(日)
主催	国土交通省

#### 開催の狙い

LINKSのオープンデータを活用した一連の開発イベント。開発者やプランナー、デザイナー等の民間人材にLINKSオープンデータに触れるきっかけを提供することで、オープンデータ活用の潜在的価値を引き出すとともに、具体的なイノベーションにつなげることを目指す。

半年間ほどの期間で、プロモーションイベント、アイデアソン、ハッカソンを連続的に開催することで、データへの習熟やチームビルディングを促し、最終的なゴールとなるハッカソンにおける品質の高い作品発表を促す。

## ライトニングトーク

### LINKS LT 2025(仮)

開発イベント	12月
主催	国土交通省

#### 開催の狙い

LTとは、自ら開発したサービスやアプリを短時間でプレゼンし合うピッチイベント。LINKSデータ活用をテーマとしたライトな発表の機会を提供することで、開発者等のモチベーションアップやプロジェクトへの認知拡大を図る。

## 官民連携技術イベント

### 国交省 x Progate プログラマー育成WS(仮)

WSイベント	未
主催	Progate

#### 開催の狙い

LLMモデルやプロンプトエンジニアリングの最新技術を発掘するためのAIに特化した開発者向けイベントを実施。

MSやProgateといった既存の技術者コミュニティを持つプレイヤーと連携することで、技術者や学生などにLINKSのデータに触れてもらい、新サービス創出等の契機とする。

+ α



2024年度のLINKS DATA x Hackathonの様子

## アプリコンテスト

### 公共交通オープンデータチャレンジ 2025 ~ powered by Project LINKS ~

最終審査会	2/21(土)
一次審査会	1/24(土)~1/25(日)
応募期間	10/1(水)~1/12(月)
プレイベント	7/1(火)から多数
主催	国土交通省

#### 開催の狙い

公共交通オープンデータ協議会(会長:坂村健)との共催により継続的に開催。GTFSをはじめとする公共交通オープンデータを活用したイノベーション創出の活性化を狙った国内最大級のアプリ開発イベント。

Project LINKS以外にも、PLATEAU、COMmmONS、国土交通データプラットフォーム等と連携。数多くの交通事業者と連携し、公共交通分野におけるオープンデータ化を推進する契機とする。



2024年度の公共交通オープンデータチャレンジの様子

## 国土交通省内向け データ活用研修

### 業務で役立つ! QGIS研修

第1回	8月上旬
第2回	9月下旬
主催	情報政策課

#### 開催の狙い

省内実務担当者における日常業務におけるデータ活用の普及を図るため、著名なオープンソースGISであるQGIS活用をテーマとして、LINKSのビッグデータを活用してパワポ資料作成等に役立つグラフや地図ビジュアルの作成方法を学ぶ研修を提供。

## ITケイパビリティ向上勉強会2025

全6回	10月~12月
主催	情報政策課

#### 開催の狙い

昨年度好評のため継続開催。省内政策立案担当者(係長~総括・企画官)向けにシステム開発プロジェクトのノウハウを提供する勉強会。

# LINKS Vedaプロトタイプ版のOSSを公開！

<https://github.com/Project-LINKS-mlitoss/LINKS-Veda>



## Project LINKS OSS organization

Project LINKSとは、幅広い施策や手続を所管する国土交通省において、これまで活用されてこなかった 様々な行政情報を機械判読や二次利用が可能な「データ」として再構築し、これを活用したデータに基づく政策立案（EBPM)の推進や、新たなビジネス創出（オープン・イノベーション）の実現を目指すDXプロジェクトです。

1 follower

<https://www.mlit.go.jp/links/>

@LINKS\_MLIT

README MIT license

- ブラウザ : Microsoft Edge、Google Chrome

## 5. インストールとセットアップ

システムは、Dockerを用いて開発環境を構築します。以下の手順に従って、開発環境をセットアップしてください。

### 1. 環境変数の設定

```
cp web/.env.example web/.env
```

### 4. アプリケーションの起動

- アプリケーションは <http://localhost:3000/> で起動します。
- Storybookは <http://localhost:6006/> で起動します。
- Prisma Studioは <http://localhost:5555/> で起動します。
- Firebase Emulatorは <http://localhost:4000/> で起動します。

5. Linting Biomeを使用してLintingを行います。詳細なインストール手順は[こちら](#)をご覧ください。

### 6. ビルド

- 開発環境ではHono Vite dev serverを使用しています。
- 本番環境では、 `server/build.ts` が `build/server/index.js` にバンドルされます。
- Dockerデプロイ時には、 `build` フォルダと `node_modules` をコピーします。

公式ウェブサイトを開発中！

<https://www.mlit.go.jp/links/>

DATA には不思議なチカラがある。

それは様々な真実を見せてくれる。

データとデータがかけあわさるとき、今まで見えなかったものが明かされる。

国土交通省では、直面する様々な社会課題を解決するため、

DATA x DATA のチカラを最大化するProject LINKS（リンクス）を始動。

LINKS を通じてデータの可能性を引き出し、新しい価値を創り出します。

POWER of DATA x DATA

Project LINKS



活用

国土交通省では無人航空機の安全飛行を目的に事故情報や機体情報など様々な情報を保有しています。

備・活用

全国的な空き家の増加が課題になっていますが、その実態把握は容易ではありません。国土交通省、地方公共団体及び民間事業者が保

LINKS:DATA x Hackat

国土交通分野のオープンデータ活用チャレ

Project LINKS のオープンデータを使って新しいサービスやコンテンツ

ハッカソンイベントを開催します。

大学、行政など、様々な立場の方のご参加をお待ちしております。

LINKS の生み出すデータと、みなさまのアイデアや技術を組み合わせた自由な発

2024  
09/06 Fri.  
キックオフ  
Kick off

オンライン開催



2024  
10/05 Sat.  
アイデアソン  
Ideathon

現地開催



# LINKSの最新情報はXで随時発信中！



# @LINKS\_MLIT

フォローよろしくお願いします！





**MLIT**  
Information Policy Division