



## データサイエンスアワード2017

# データサイエンスと働き方改革 ～安全で効率的なメンテナンスを目指して～

2017年10月23日

学校法人産業能率大学総合研究所

福中公輔

澤田基之

東京地下鉄株式会社

小西真治

川上幸一

三浦孝智

今泉直也

**日々、多くお客様が東京メトロを利用しています**

**お客様の安全を守るためには、毎日の点検を欠かすことはできません**



お客様の安全を守るためには、毎日の点検を欠かすことはできません

**しかし、東京メトロ様が管理する路線は膨大です**



銀座線

丸ノ内線

日比谷線

東西線

千代田線

有楽町線

半蔵門線

南北線

副都心線

**少しの異常が重大な事故につながる可能性がある**

**不断の努力**



**私たち産業能率大学は、  
この東京メトロ様の日々の努力に敬意を払うとともに、**

**「地下鉄を利用するお客様により安全で快適なサービスを提供したい」**

**そんな東京メトロ様の思いに応えるため、データサイエンスの観点から  
ご協力をさせていただきました**

**これは、そんな社会インフラの安心と安全のために  
データサイエンスを活用した  
1つの事例報告です**

# 1. 土木構造物検査データ管理システムの構築

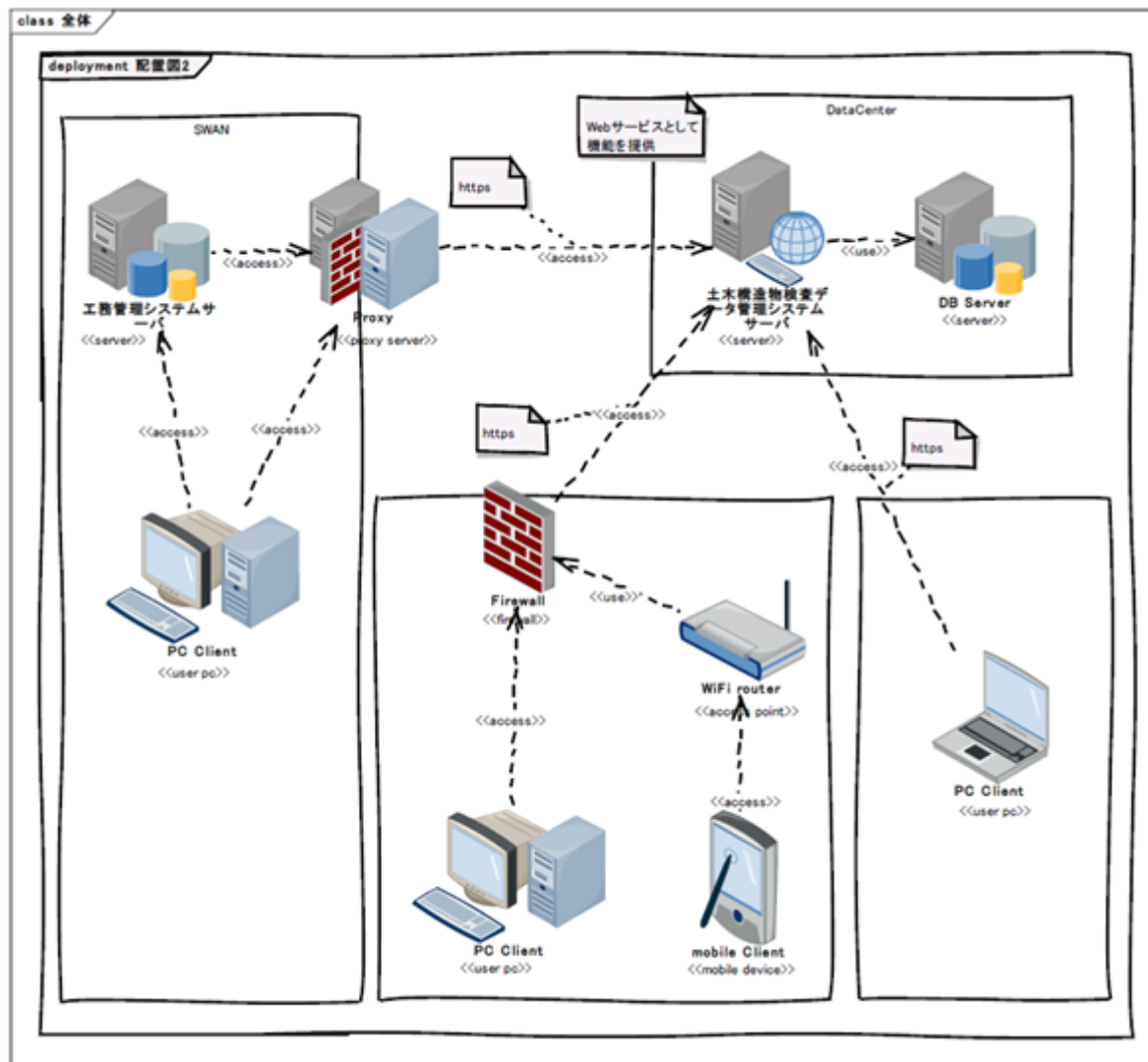
検査結果の共有までに要する時間

最大  
3ヶ月



翌日

# MMMシステムの開発

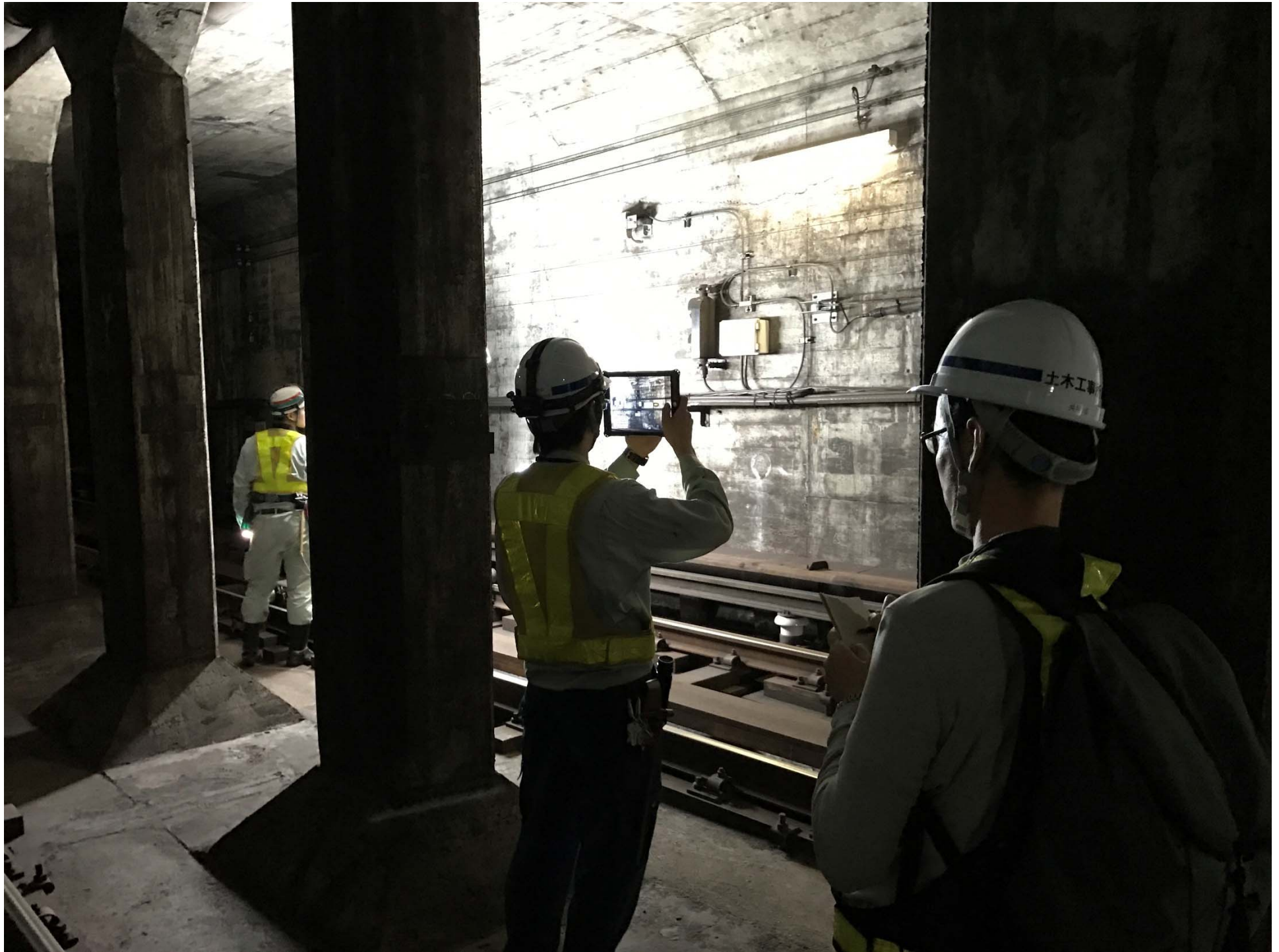


## システムの物理構成

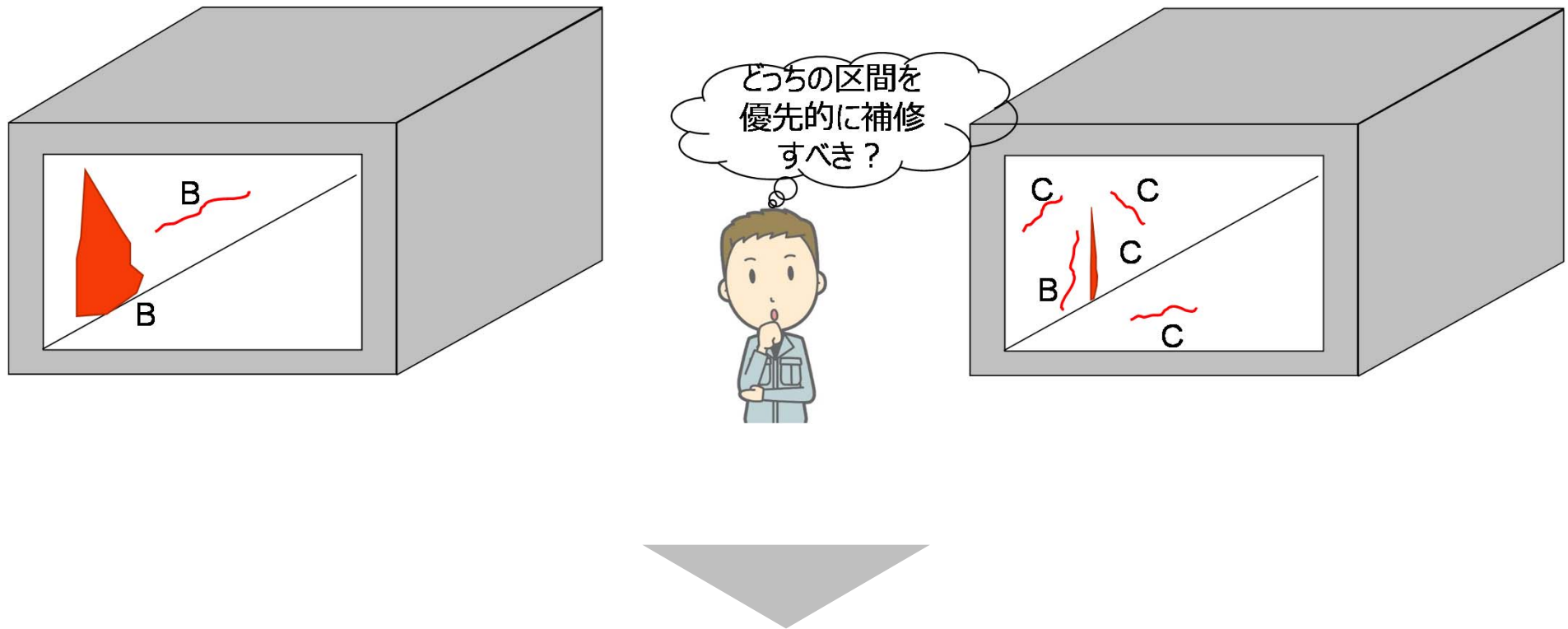
インターネット上のサーバは、各クライアントPCおよびモバイルデバイスに対してWeb Service として機能を提供する。

各クライアントは専用のプログラムによって、httpsを用いてサーバへアクセスする。

工務管理システムとの連携は、工務管理システム側からWeb Serviceへアクセスすることによって行う



## 2. 土木構造物の健全性を評価し、比較するための指標を開発



トンネルの健全度を定量評価したい

## 測定モデル： 維持管理指標 $\theta$ がガンマ分布に従う検査項目反応モデル

$$P_{ij} = \frac{1}{1 + \exp(-a_j(\theta_i - b_j))} \quad (1式)$$

$$\theta \sim \text{Gamma}(\text{shape}, \text{scale})$$

ここで、 $a$ は識別力を、 $b$ は困難度を表しています。

$P$ は変状箇所の観測確率を表しています。

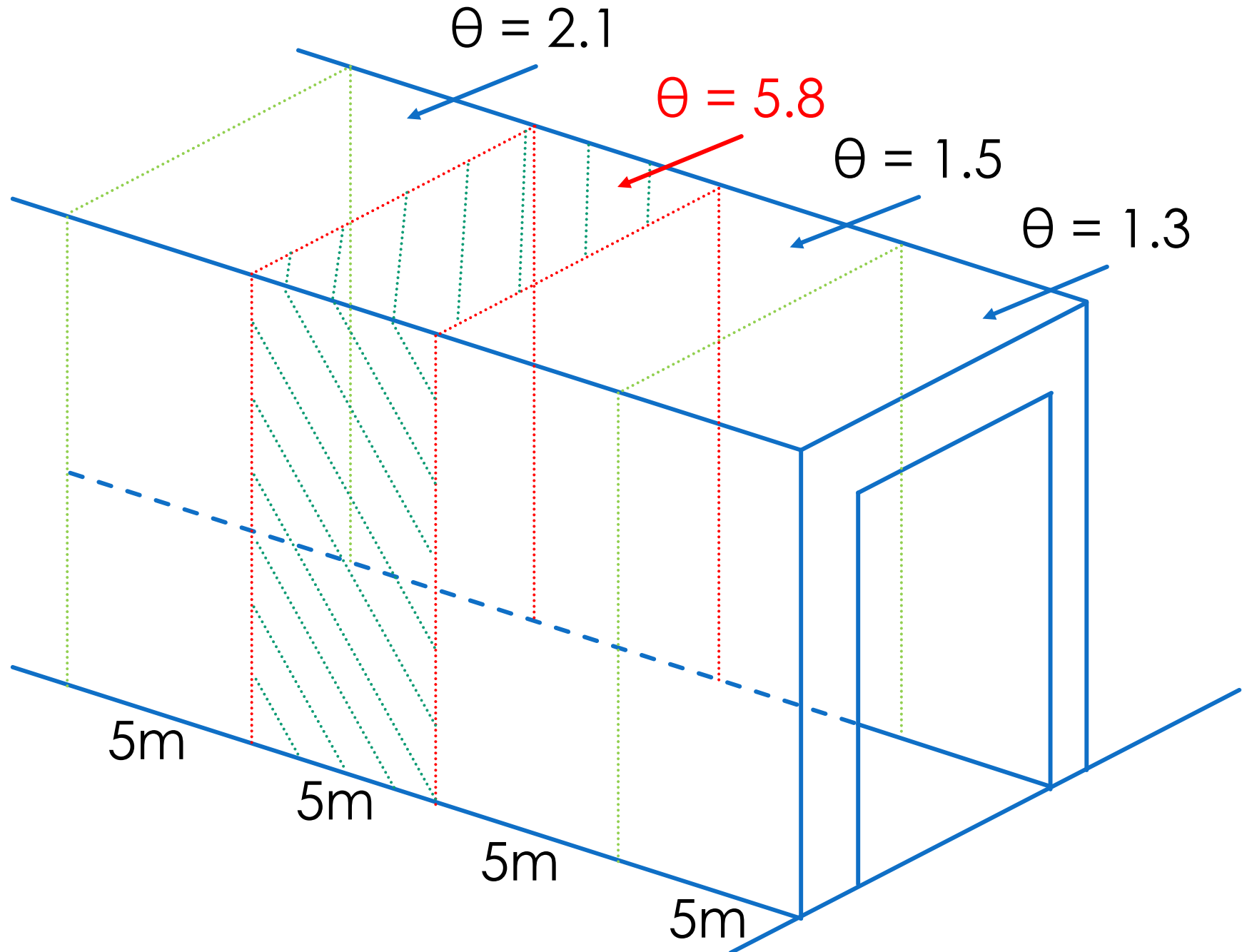
添え字  $j$  は検査項目の変数の種類を表しています。

添え字  $i$  は路線区間を表しています。



分析用データに対して上記の数理モデル(検査項目反応モデル)を当てはめ、マルコフ連鎖モンテカルロ法を用いたベイズ推定により各種パラメータ( $\theta$ ,  $a$ ,  $b$ )を推定します。

θの値が高いところはリスクが高い



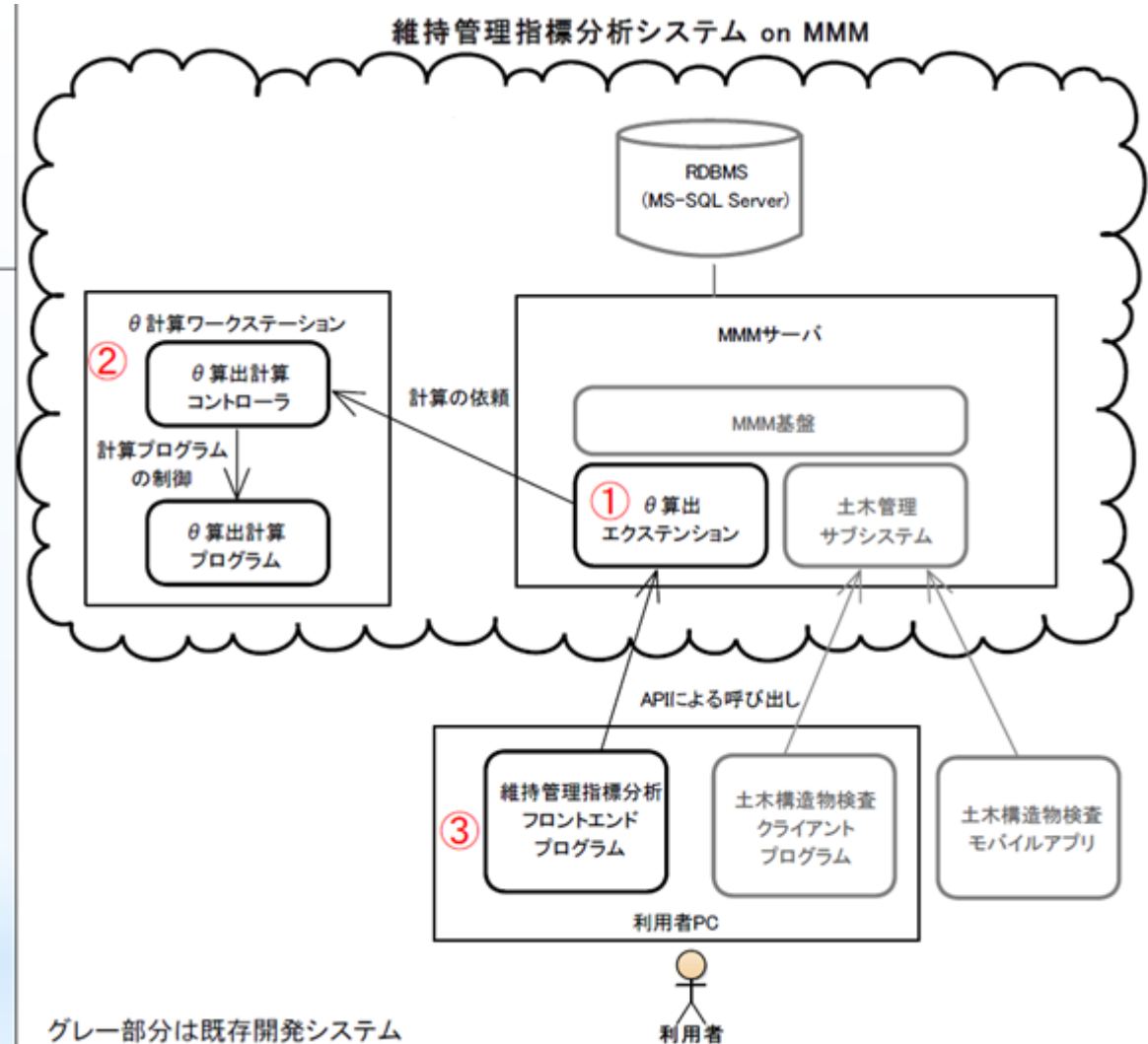


平成28年3月25日  
特許第5904651号を取得  
維持管理指標算出装置及び維持管理指標算出方法

# 開発

- ① 既存のMMMサーバ上に  
Θ算出拡張モジュールを追加
- ② Θ計算は非常に重い処理のため  
計算用のワークステーションをメトロ  
クラウド内に新規導入
- ③ 利用者PC上にインストールされた  
フロントエンドプログラムから  
計算実行および分析を行う

計算完了後、利用者にメールで通知



### 3. ベイジアンネットワークによるトンネル内変状観測確率の検討

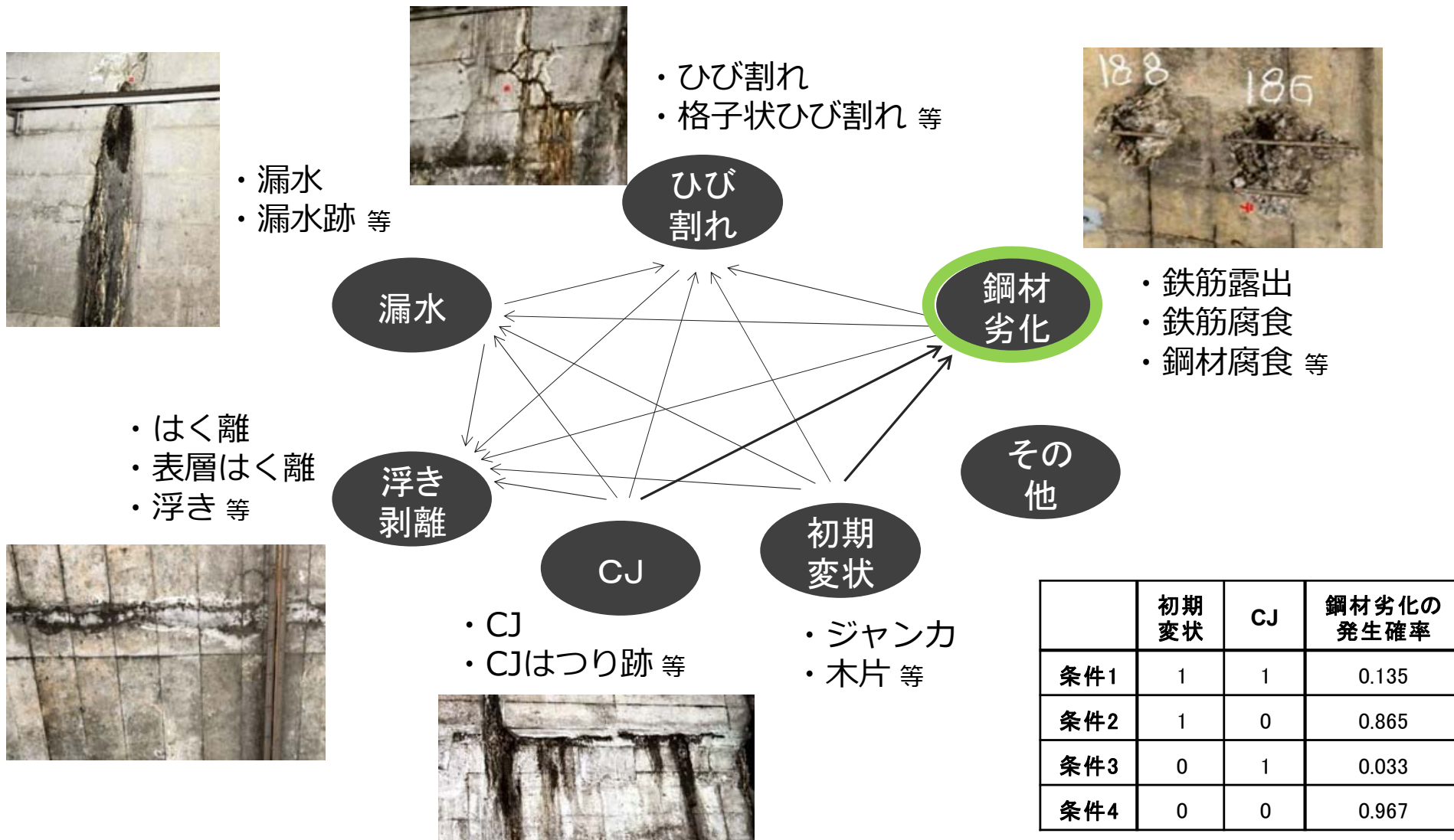


目視主体  
変状発生箇所に見逃しはないか？



ベイジアンネットワークによる変状観測確率の計算

- 平成28年3月25日 特許第5904650号を取得
- 第21回地下空間シンポジウム(土木学会主催)にて講演奨励賞を受賞



1・・・あり、0・・・なし

## 4. 社内データサイエンティスト育成のための教育研修を実施

東京メトロ様の様々な部署（工務部・電気部・車両部・ICT戦略部）の中からメンバーを選抜して統計学やデータ解析に関する教育を行い、**自社の業界に特化した社内データサイエンティスト**として活躍できるだけの知識を身につけ、業務における**「データ解析プロジェクト」**を自分たちで遂行できるようになることを目指します。



- 期間は3年間
- 【基礎】統計学・データ解析の基礎的な知識を習得します
- 【応用】応用的なデータ解析のノウハウを実践形式で習得します
- 【実践】課で必要なデータ解析プロジェクトの遂行法を学びます

## 総合研修訓練センター

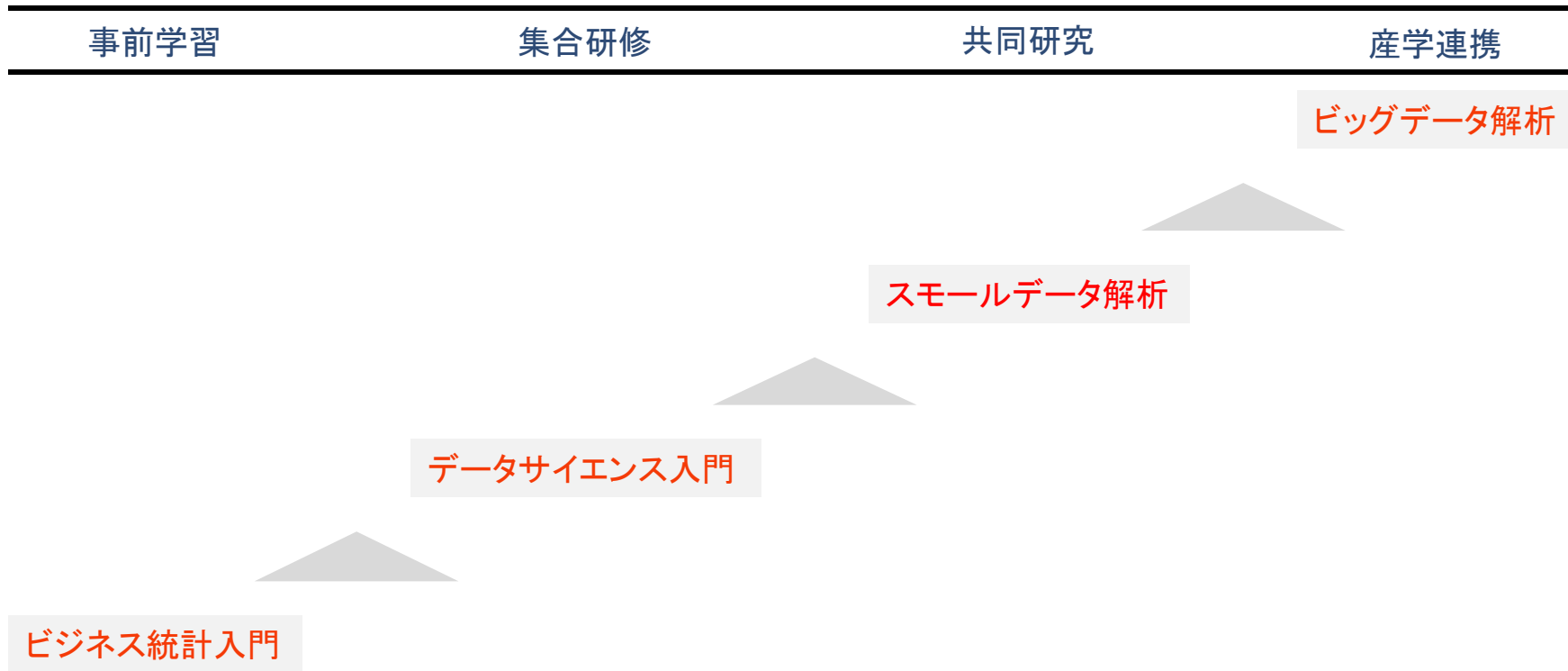
当センターは、これまで各所に点在していた各部門の研修施設を統合し、研修棟と営業線に準じた訓練線を有しており、時間的・空間的制約を受けずに複合的かつ総合的に知識・技能を習得できる施設です。

東京地下鉄様ホームページより抜粋



# データ解析プロジェクトを効率的に進めるために

データ解析に関連する知識は多岐にわたるため、その全体像を同時に把握することは困難です。そこで本学では、無理なくムダなくプロジェクトを進められるように、以下のような段階的なアプローチで教育を行いました。



自社でデータ解析プロジェクトを実施するためには、統計学の基礎知識だけでなく、データ解析の一連の流れを知っておく必要があります。これは外注する場合でも同様で、本格的なプロジェクトを遂行する前に、データサイエンティストなどの専門家と対話するための「共通言語」を身につけておくと、プロジェクトを円滑に進めていくことが可能になります。

最初から高コストなビッグデータ解析に手をつけるのではなく、まずはデータ量の少ないスモールデータ解析から始めることを強くおすすめします。スモールデータ解析で事業の方向性や分析方針に当たりをつけてから、ビッグデータ解析プロジェクトを始めることで、最小のコストで最大の効果を引き出すことが可能となります。

## 【基礎】スケジュール

初級編のスケジュールは次のとおりです。1会合あたり、4時間程度で実施しました。  
統計学の基礎とデータ解析の一連の流れをマスターすることが目的となっています。  
同時に、中級編で使用するデータセットの検討および作成を行いました。

日程	内容
8月	(本プロジェクト計画決定、対象者の決定)
9月	事前学習(通信教育4ヶ月間) 受講開始
9月	オリエンテーション:統計解析ソフトRの入門、ビジネス課題の抽出
9月	データ解析基礎教育①:データサイエンス入門、データマイニング標準プロセス
10月	データ解析基礎教育②:統計学の基礎、ビジュアライゼーション
10月	データ解析基礎教育③:確率分布、統計的仮説検定
11月	データ選定と分析①:回帰分析、Rの振返り
11月	データ選定と分析②:母数の推定
12月	データ選定と分析③:ダミー変数の導入
1月	データ選定と分析④:ロジスティック回帰分析
2月	プレゼンテーション

## 【応用】スケジュール

中級編のスケジュールは次のとおりです。1 会合あたり、4 時間程度となっています。  
人工知能に実装されている機械学習をマスターすることが目的です。  
同時に、初級編で作成したデータセットの分析を行います。

日程	内容
8月	オリエンテーション： データマイニング入門
9月	機械学習①： アソシエーションルール
9月	機械学習②： ベイジアンネットワーク
10月	機械学習③： 決定木(ランダムフォレスト)
10月	機械学習④： サポートベクトルマシン
11月	機械学習⑤： クラスタ分析
11月	機械学習⑥： ニューラルネットワーク(深層学習:ディープラーニング)
12月	研究法①： 応用データ解析の実践
1月	研究法②： 応用データ解析の実践
2月	プレゼンテーション(研究成果の報告)
3月	論文指導

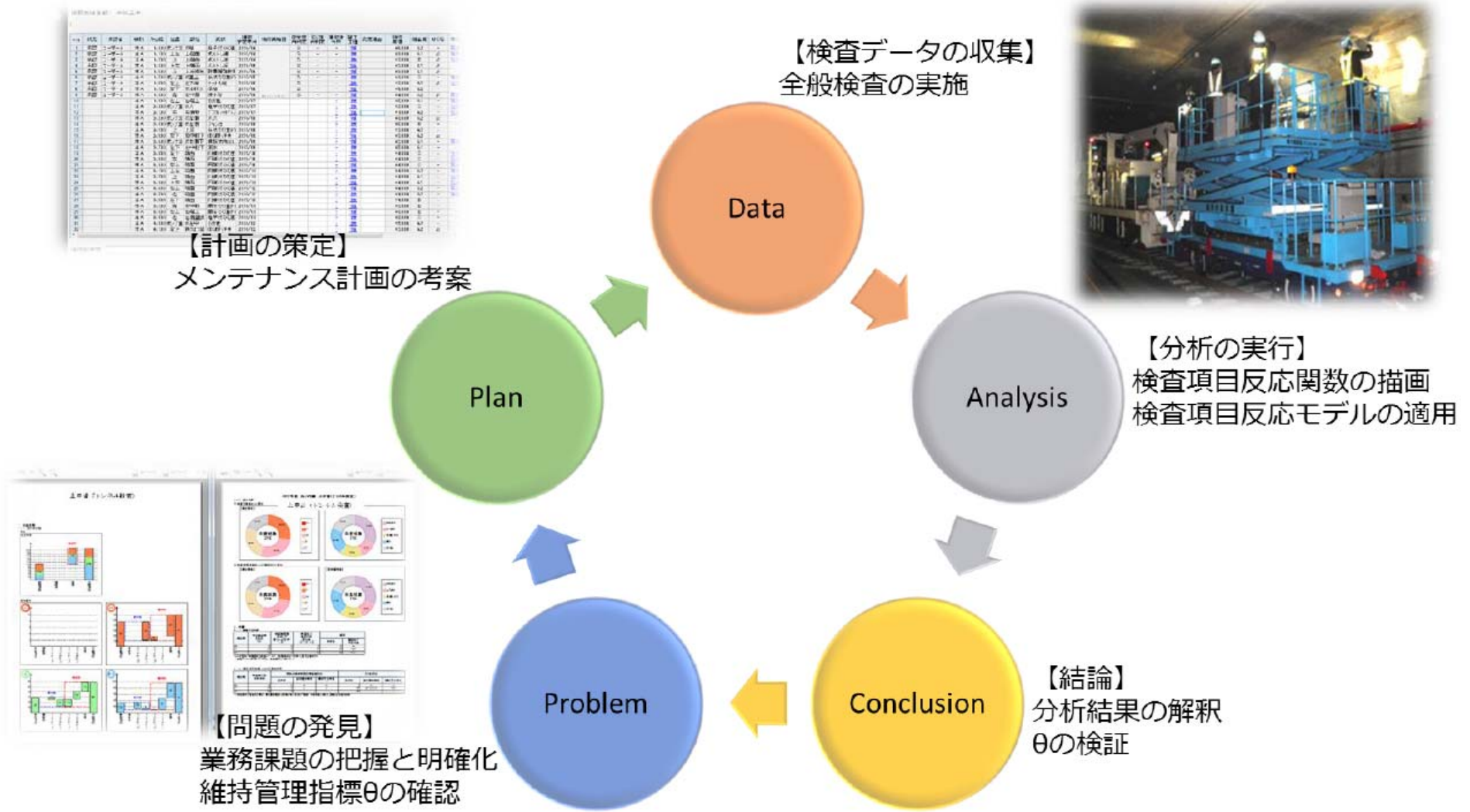


## 【実践】スケジュール

実践編のスケジュールは次のとおりです。1 会合あたり、3時間程度となっています。  
新規に開発した**検査項目反応モデル**や**ベイジアンネットワーク**について、理論と応用の仕方を学ぶとともに、実際のプロジェクトとして業務を遂行する方法について理解することが目的となっています。

日程	内容
5月	オリエンテーション, 統計解析ソフトRの復習
6月	解の推定(最尤推定・ベイズ推定)
7月	マルコフ連鎖モンテカルロ法・検査項目反応モデルの数理的構造の理解
8月	検査項目反応モデルの運用の仕方を理解する
9月	自部署のデータに対するベイジアンネットワークの応用
10月	社内データサイエンティストに求められる役割りを知る

# 通常業務での運用



東京外口様

産業能率大学



まとめ

**2013年度より、継続的に社内のデータを分析・報告してきました**

**この取り組みは、  
地下鉄を利用するお客様の安全を守るための、  
データサイエンスの先進的な活用事例となっています**

## **実績**

**平成28年1月22日 早稲田大学国際会議場  
第21回地下空間シンポジウム(土木学会主催)  
ベイジアンネットワークによる地下鉄トンネルの変状観測確率の検討  
講演奨励賞を受賞**

## **実績**

**平成28年3月25日  
特許第5904650号を取得  
管理指標算出装置及び管理指標算出方法**

**平成28年3月25日  
特許第5904651号を取得  
維持管理指標算出装置及び維持管理指標算出方法**



**高度なデータサイエンスの研究と実務レベルでの応用が  
東京メトロ様の中で実を結びました**

**しかし、どんなに高度な知見も理解されなければ役に立ちません**

**理解できなければ実務で応用し続けることは困難です**

**それゆえに、社内でのデータサイエンティストの育成が不可欠なのです**

**東京メトロ様は  
これまでも、これからも、  
お客様がより安全に  
そして、より快適に地下鉄を利用できるよう、挑戦し続けます**

**私たち産業能率大学も  
この東京メトロ様の真摯な思いに応えられるように  
可能な限り、支援し続けたいと思います**

**今日と同じ明日を**

**ご静聴、ありがとうございました**