

# インフラ維持管理における データサイエンス活用による生産性向上

湧田 雄基・阿部 真育

本稿では、土木インフラ維持管理分野を対象とした業務効率化のためのデータサイエンス活用について述べる。著者らは、データサイエンス試行を支援するシステムを開発しており、中でも特に、データ分析結果の解釈を支援する技術として、分析結果の文章化表示を行うための概念構造「オントロジ」の設計について紹介する。

キーワード：インフラ維持管理, データサイエンス, 機械学習, 結果の解釈性, 業務効率化

## 1. はじめに

近年のICTの急激な進歩により、あらゆる分野において大量のデータが日々蓄積されている。これまでの少ないデータから傾向を予測し意思決定を行うのではなく、大量のデータの中から重要な要素を抽出して意思決定を行う時代になったと言える。従って、世の中の人的ニーズが、少ないデータから傾向を見出す統計専門家から大量のデータの扱い方を把握しているデータサイエンティストにシフトしてきており、データサイエンス人材の育成が急務となっている。図-1はデータサイエンス人材育成の現状を示した図となっており、業界を代表するレベルで5~20人、見習い・独り立ちレベルで3,400人程度が年間輩出されていると言われており、棟梁レベルのデータサイエンス人材が不足していることが各種報告書や国の委員会等にて

課題として挙げられている。しかしながら、次代のデータサイエンス人材のポテンシャルとして大きい部分を占めているのは、“データサイエンティスト以前の方”と定義されているピラミッド構造での裾野人材であると言える。この裾野人材に対するデータサイエンス教育を如何に効果的に行えるかが、来るべき超ビッグデータ社会を乗り切るためには、教育研究機関のみならず産業界にとっても重要なミッションとなり、それ故にデータサイエンス人材育成の確固たるモデルケースを構築することが喫緊の課題となっている。データサイエンティストは、データから得られる結果を解釈し、業務などの応用に対する活用の方策を見定めることで、これらをつなぐスキルを有することが期待される。本研究では、そのシステム開発において最も現場の業務支援に直結する、分析結果の解釈の理解促進に着目し、これを支援する技術の一例を紹介する。

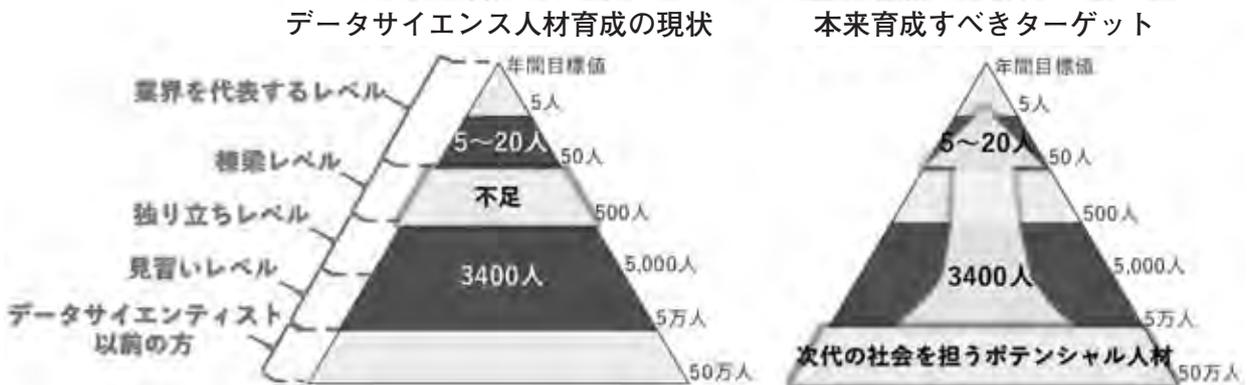


図-1 次代の社会を担うデータサイエンス人材育成のターゲット

(データサイエンティスト協会発行『データサイエンティストスキルチェックリスト Ver.2』, 並びに『科学技術・学術審議会戦略的基礎研究部会 数学イノベーション委員会 (第23回) H27.10.30』より作成)

## 2. インフラ維持管理分野におけるデータサイエンス

アメリカにおいて、1920年代のニューディール政策以降、大規模なインフラ整備が進められ、1980年代にインフラの老朽化問題が、多くの落橋事故により顕在化した。このような時代は『荒廃するアメリカ』と称され、インフラ維持管理の重要性が見直されるきっかけとなった。その後、世界銀行が1994年に公表した World Development Report<sup>1)</sup>において、予防保全のためのインフラ維持管理データの蓄積とデータを用いた維持管理の実施が世界的に推奨された。アメリカのインフラ事情から30年のラグがあると言われる日本においても、近年、インフラの老朽化対策は大きな政策課題となっている。そのため、維持管理の徹底が重要視されており、維持管理の効率化を目的としてデータを扱う様々な研究開発が行われている<sup>2)</sup>。このような社会情勢において、日々蓄積されるインフラ維持管理データの活用を実業務に取り入れ、効率化をいかに果たしていかせるかがインフラ管理者の喫緊の課題と言える。

インフラ維持管理において蓄積されるデータを活用し、生産性の向上をはかることについては、大きな期待がある。インフラ維持管理分野において、データを活用した生産性向上・効率化については、点検・診断・措置・記録それぞれの業務に応じて様々な研究・実装が進められている。ここでは、代表的な研究事例について紹介する。

まず、インフラの変状検出にデータ活用をする事例である。点検業務により観測され、項目化し記録される「個々の事象」を変状として抽出するが、その際にデータ分析が活用される事例がある。貝戸、小林らの研究<sup>3)</sup>では、目視点検データを対象に、混合ポアソンモデルを用いて、RC床版のかぶりコンクリートの剥離・剥落発生の推定を行っている。変状検出をAIで代替し効率化をはかる研究は多く<sup>4)~6)</sup>、特に、コンクリート表面の変状検出については、多くの研究が存在し、Kochら<sup>7)</sup>の報告の中では、コンピュータビジョン技術の土木技術への適用について述べられ、トンネルのひび割れ検出<sup>8)</sup>やひび割れの背後で生じる透水をモデル化した上で高速にひび割れ検出を行う手法<sup>9)</sup>、SVM (Support Vector Machine) を用いた手法<sup>10)</sup>等が挙げられている。

その他、構造物の状態を総合的に評価し、補修や点検の計画立案に活用する事例も多い。維持管理の対象となる構造物を特定の単位に分割し、状態を総合的に

評価し指標や程度(ランク)で表現するといった事例である。構造物に関する個別変状の程度や発生箇所、進行状況などを総合的に加味し、構造物の健全性、あるいは劣化度、老朽化年齢、リスク、といった総合的な指標を算出する。ここで得られる指標やモデルは、点検や診断の後の措置で行う維持管理の緊急性や優先度の判断にて活用される。

また、判別や診断の根拠の解釈等を目的として、要因分析/データ構造理解としては、変状や劣化等の関心のある事象について、「要因になっている潜在的な情報が何か」といったことを調べる研究が行われている。維持管理においては、劣化要因により、今後の劣化の進行性についての判断を行った上での措置が求められるため、要因の特定は、対策検討のための重要な診断作業となる。従って、変状要因分析やデータに関する構造理解を進める研究については、生産性や業務効率化において現場ニーズも高く、研究事例も多い。例えば、皆川ら<sup>11)</sup>は、RC床版の疲労劣化を対象に、特にひび、剥離、鉄筋露出、遊離石灰といった変状事象の情報を対象に、データの関係性の調査や劣化因子の推定を行っている。有賀ら<sup>12)</sup>は、センサデータモデルと知識モデルとを統合化し、膨大なセンサデータより建物状態に関する知識発見を行うデータマイニング手法を開発している。

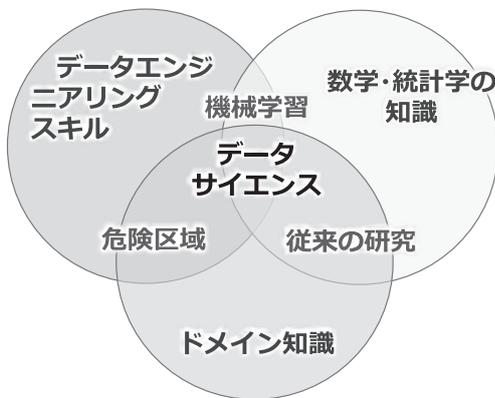
さらに、インフラの将来の状態を予測する研究も多く取り組まれている。構造物状態を見据えた管理計画立案は、維持管理における効率化の重要な課題である。吉田ら<sup>13)</sup>は、構造物健全性の劣化について、経過年による劣化二次曲線に対して劣化機構に関する特性指標を説明変数として加えた回帰式を求めている。この回帰式の推定において、EMアルゴリズム (Expectation-Maximization algorithm) と PSO (Particle Swarm Optimization) を組み合わせた手法を提案している。また、中津ら<sup>14)</sup>は、遺伝的アルゴリズム (GA : Genetic Algorithm) を用いて予防保全を考慮し、補修年度を前倒した橋梁維持管理計画の策定を提案している。

このように、インフラ維持管理分野におけるデータ活用・データサイエンスに関する研究は、それぞれの業務にあわせて積極的に研究が進められている。ここ数年については、深層学習の活用により特に報告数が多くなっているが、データ分析結果の解釈を扱う研究は少ない。また、インフラ分野におけるデータサイエンス試行を支援する仕組みについての報告は見当たらない。

### 3. インフラ維持管理分野におけるデータサイエンスの実践

データサイエンスの実践には、Conwayによる定義<sup>15)</sup>によれば、「データエンジニアリング力」、「数学・統計学の知識」、「実質的な専門知識（ドメイン知識）」の3つの要素が必要とされている（図一2）。これらは、一般社団法人データサイエンティスト協会が示すデータエンジニアリング力、データサイエンス力、ビジネス力にそれぞれ対応すると考えられる。また、この図において、3つの要素をバランス良く満たす取組がデータサイエンスであると述べられている。また、例えば、「数学・統計学の知識」が不足する取組は、誤った処理や解釈を行ってしまう可能性があるとして危険区域とされている。

このような3つの要素を併せ持つ人材を育成するためには、例えば、データエンジニアリング力や数学・統計学の知識を有する「分析の専門家がドメイン知識を習得する場合」や、あるいは逆に、「ドメイン知識を有する者が、データエンジニアリング力や数学・統計学の知識を学ぶ」というアプローチが考えられる。



図一2 データサイエンスの構成要素<sup>15)</sup>を元に一部改変)

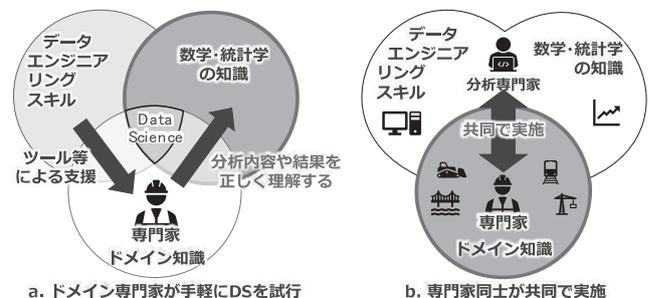
しかし、インフラ維持管理分野における専門知識の習得には、現場の経験や工学的知見などが必要であり、これらを習得し、なおかつ数学・統計学の知識までを深く学ぶことは非常に時間がかかると考えられる。

従って、本研究では、ドメイン知識を有する者が、手軽にデータサイエンスを試すことができるよう、この試行を支援する仕組みの構築を目指す。データエンジニアリングスキルのない者をツール等で支援し、さらに、データ分析結果の解釈を促す（図一3 a）。

また、インフラ維持管理分野においては、データ分析の専門家とドメイン知識の専門家が共同でデータサイエンスを行う場合も多いと考えられる（図一3 b）。この場合にも、分析結果についての解釈を共有することが必要となる。

### 4. インフラ維持管理データサイエンスの支援

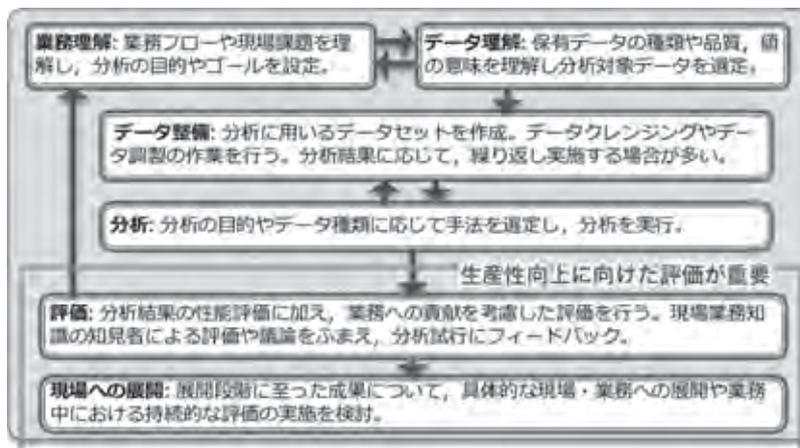
インフラ維持管理業務を対象としたデータサイエンスの流れについて、著者らは図一4のように6つのプロセスにより定義している。ここで示す業務プロセスは、CRISP-DM および KDD を基に、インフラ維持管理に適合するようカスタマイズを施したものである<sup>16), 17)</sup>。まず最初に、「業務理解」と「データ理解」を行う。ここで、データ活用により改善を行いたい対象業務に



a. ドメイン専門家が手軽にDSを試行

b. 専門家同士が共同で実施

図一3 データサイエンス実践のパターン



図一4 インフラ維持管理分野におけるデータサイエンスのプロセス

ついで理解を行う。必要に応じて、業務分析等を行い、各業務において作成している資料を調査し、意思決定や業務評価に必要となる指標の整理を行う。このとき、資料や指標の作成に必要となるデータの調査も合わせて行う。データサイエンスに活用可能なデータの棚卸しを行い、データを理解し、分析対象データの選定を行う。「データ整備」では、分析に利用するデータセットの作成として、データのクレンジングやデータ統合等による整形を行う。この結果として得られたデータセットを用いて、「分析」を行う。最後に、分析結果の「評価」と「現場への展開」を行うが、これらは、分析結果を正しく理解し、業務への活用を目指す為に、データを活用した生産性向上に関わる重要なプロセスと言える。

データサイエンスの結果を業務に活用する場合の業務改善効果としては、例えば「時間削減」、「コスト削減」、「リスク低減」などの定量的な指標による評価が考えられる<sup>18)</sup>。時間削減については、点検や診断の業務や計画、資料作成等に要する時間の短縮や業務の省略化に繋がるかどうかについて、シミュレーションや実験により時間を計測し評価する方法である。コスト削減については、時間短縮に繋がる部分もあるが、管理に必要な資材や人員や業務体制自体の見直しを費用に換算して評価する方法や保有資産の価値による評価などの研究がある。リスクについては、未把握の潜在的な劣化を把握することで見落とし防止に繋げるなどの評価を行うことが考えられる。これらの評価は、業務に直結した指標 (KPI: Key Performance Indicator) である。一方で、データサイエンスの結果としては、これらの指標を直接算出できることは希で、分析結果の性能指標を介し、間接的に解釈することができる。例えば、「人による構造物の診断業務を代替することで時間短縮に繋がる」ことや、「変状検出作業の精度が向上することでリスク低減に繋がる」などである。ここでは、データサイエンスを業務で活用するために、間接的な指標となる分析結果指標を解釈するための処理について述べる。

		判別の結果	
		正と判別(P:Positive)	負と判別(N:Negative)
真値	正例 (P:Positive)	TP: True Positive (正解=真陽性)	FN: False Negative (不正解)
	負例 (N:Negative)	FP: False Positive (不正解=偽陽性)	TN: True Negative (正解)

図-5 混同行列による分析結果の表現

### 5. 結果解釈の支援

インフラ維持管理におけるデータ分析では、上で述べたとおり「変状有無の判定」や「劣化判定」、「補修等の対策の必要性についての判別」など2クラスの判別を行う例が多く見られる。ここでは、このような2クラス判別を対象に、分析結果の解釈について述べる。

2クラス判別の結果評価においては、「混同行列 (Confusion Matrix)」による評価が一般的である。この混同行列は、2つのクラスに正しく判別できているかについて、判別結果となる件数配分と、判別結果の正否についての割合を2×2の行列で示した表である (図-5)。この結果より、精度や検出率などが把握でき、判別モデルの性能がわかる。行列は、TP, FP, FN, TNの4つの指標で構成される。判別結果として、例えば「変状有り」などを対象として正と判別された結果について、正解であった件数はTP, 不正解であった件数はFPである。

これらの値より、判別結果の性能を示す精度 (Precision) については、TPをTP+FPで除して把握することができる。また、検出率 (Recall) については、TPをTP+FNで除して把握できる。これらの2つの指標は、2クラス判別の性能を評価する重要な指標となるが、一般には、判別の「閾値」を調整することでこれらの性能は変化する。この変化を表現したグラフがP-R曲線となる。

データサイエンスに慣れた者であれば、上で挙げた混同行列から、判別モデルを用いた場合の効果について、解釈することができるが、これらの数値から結果を解釈することは、経験のない者には難しい。また、応用分野の者に対して、この結果を説明するためには、多くの場合は、日本語での言い換えを必要とする。

変状検出課題を対象とした2クラス判別結果の混同行列の具体例について、図-6に示す。

上図の混同行列結果についての解釈例については、例えば、次のような解釈が可能である。

「判別分析を用いて、5,178件に対して、変状有無の判別を行ったところ、3,871件が変状有りと判別され

		推定/判別結果		真値計	真値率
		Positive(P) 推定陽性	Negative(N) 推定陰性		
真値	Positive(P) 陽性	TP (True Positive) 1,083 [件]	FN (False Negative) 57 [件]	真P計 1,140 [件]	真P率 22.2[%]
	Negative(N) 陰性	FP (False Positive) 2,788 [件]	TN (True Negative) 1,250 [件]	真N計 4,038 [件]	真N率 77.98[%]
推定値計		推定P計 3,871 [件]	推定N計 1,307 [件]	総計 5,178 [件]	

図-6 混同行列の具体例

る。このうち、1,083 件に、実際に変状が存在していることから、この判別モデルの変状推定の「精度」は、28.0%となる。また、実際に変状がある箇所は、1,140 件であり、この判別モデルを用いることで、このうちの1,083 件を検出できたことになる。従って、この判別モデルによる変状の検出率は、95.0%であることがわかる。」

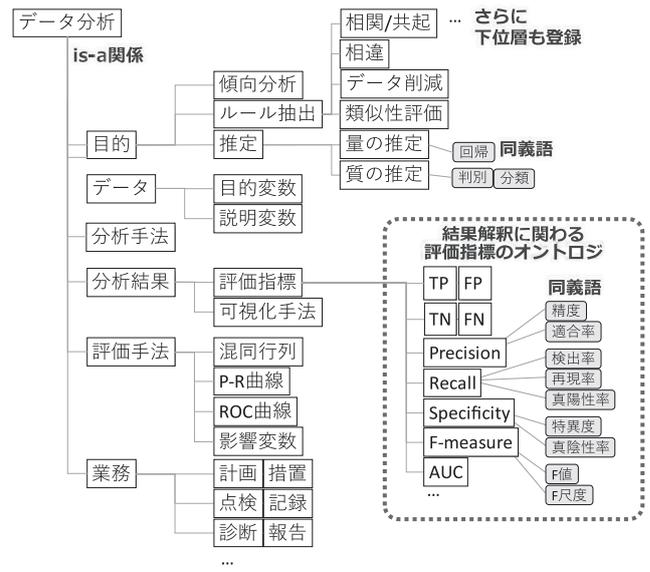
このような基礎的な解釈については、判別モデルの実用を仮定した場合の性能を仮説的に検証した場合の効果を推測するものである。混同行列の定量的な評価結果を実務に照らした評価を行う手段のひとつであり、指標による結果を日本語で説明したものである。

上で述べたように、分析結果について、混同行列で提示するより、文章により示した方が、不慣れな者にとっての解釈性は良い。このとき、データサイエンティストであれば、混同行列を見て、分析結果を日本語で説明することができる。このような、データサイエンティストが行っている、データ分析結果を解釈しその結果を言葉で説明する「翻訳作業」について、この変換操作を自動化し、分析結果を説明する機能を実現する。

混同行列と日本語文章表現との間を結ぶため、概念構造とこの構造に対応付く語彙をセットとしたオントロジを作成した。このオントロジは、データサイエンスにおいて扱う概念（クラス）とその関係性を定義したものであり、これに実体（インスタンス）として語彙を登録することで、語彙同士の関係を識別することが可能となる<sup>19), 20)</sup>。また、作成したオントロジを用い、分析内容や結果を日本語文章で表現する為の文章化テンプレートを作成した。

作成した「インフラ維持管理オントロジ」のプロトタイプの一部を図一七に示す。オントロジの編集には、古崎らにより開発されている法造<sup>21)</sup>を用いた。

オントロジ作成においては、インフラ維持管理分野

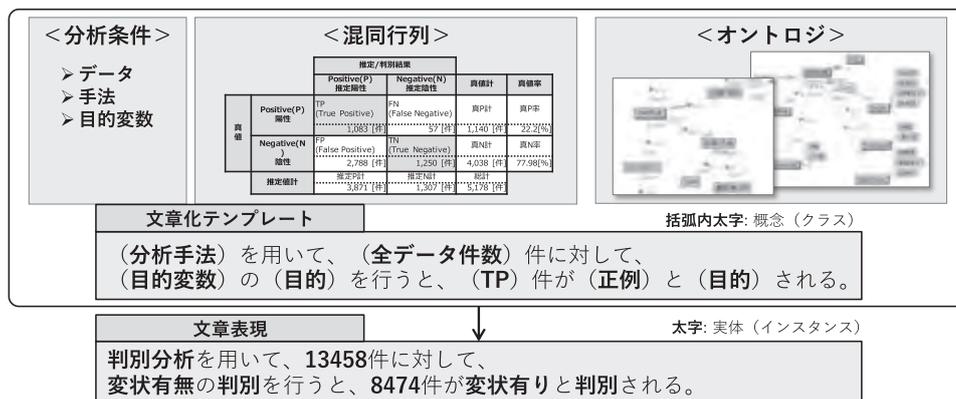


図一七 インフラ維持管理オントロジ

におけるデータサイエンスで用いる概念を列挙し<sup>18)</sup>、これらの関係を定義し、構造化した(図一七)。例えば、データ分析の構成要素として「目的」、「データ」、「分析手法」、「分析結果」、「評価手法」、「業務」などを is-a の関係により定義した。さらに、これらの要素の子となる要素についても is-a の関係で定義し、樹状の構造とした。

例えば、分析結果を説明するための「評価指標」について見てみると、TP, FP などの混同行列の要素があるほか、Precision や Recall などの指標がある(図一七)。Precision は、精度や適合率といった同義語も存在しており、これらも同義語の語彙として登録している。

このオントロジでは、先に述べた混同行列の構成要素となる TP, FP, FN, TN などの指標やこれらにより定義される精度、検出率等の指標、また、判別対象としている正例値や分析目的、分析手法等の情報をインスタンスとして登録し、扱うことを可能としてい



図一八 オントロジを用いた分析結果の文章表現の試行例

る。

これらの値については、データ分析によって得られ、オントロジによりこれらの関係性を把握できる。この値を用いた日本語による文章化を行う文章テンプレートを準備することで、分析結果の日本語表現を機械処理により行うことが可能となった(図-8)。

## 6. 支援システム

著者らは、これまでの研究により得られたインフラ維持管理分野におけるデータサイエンスの事例に基づき、事例で頻繁に用いるデータサイエンスに必要な処理を実装し、インフラ維持管理のデータサイエンスの試行を支援する機能を提供するシステムを開発している<sup>22)</sup>。その中で、上で述べた分析結果の解釈のための日本語文章表現についてもオントロジや文章テンプレートを扱える構成により実装を行っている。

本システムのコンセプト(特長と狙い)は、次の3点である。

- (i) データサイエンスの流れを把握できる: 処理の流れ、分析手法の選択、結果の見方について支援してくれるテンプレートを提供
- (ii) 有効な分析を簡単に試行できる: 万能で複雑な事ができるツール構築は目指さない
- (iii) インフラ維持管理分野に適した処理を備える: 研究事例を基にした「データクレンジング」のための「辞書」、「データ調製」のための「ルール」、再利用可能な典型的な「処理手順(レシピ)」のテンプレートに沿ったハンズオン・システム

これまでに、データ処理手法の研究開発および実装のほか、ユーザが利用するためのGUIの設計と実装を進めた。GUIについては、データサイエンス試行

の流れに沿って、「データ確認」「クレンジング」「データ調製」「分析+評価」の4つのフェイズ毎に画面およびサーバ側の処理を構成した(図-9)。このGUIの中で、利用者へのメッセージを表示するウィンドウを設けている。このメッセージでは、データサイエンスのプロセスに沿って、現在の画面で何をすべきなのかについて、ユーザに促すことができるほか、先に述べたデータ分析結果の解釈についての日本語表現による説明を表示する役割を担う。

本システムでは、前述の研究開発の成果にあたるモデリング手法やデータクレンジング手法を、定型的な処理であるモジュールとして実装し、サーバ内に実装するソフトウェア開発を進めている。

本システムでは、情報を専門とする者ではなく、土木や現場知見を専門とする者がデータサイエンスに挑戦する事を支援するものであり、そのため、過去の知見・データに基づき、インフラ維持管理分野で活用できるデータクレンジングの辞書を備え、同分野で有効なデータサイエンスの流れを提供するツールをガイドとセットで提供し、初学者でも迷わずデータサイエンスを試行できる成果となるよう開発を進めている。

## 7. おわりに

本稿では、インフラ維持管理分野において、データ活用による生産性向上や業務効率化を目指すデータサイエンスについて、結果の解釈を行う為の技術開発について述べた。特に、2クラス判別問題を題材とした場合の判別結果について、その解釈結果を日本語で表現するためのオントロジを活用した文章表現に関する事例について示し、また、この技術を用いたデータサイエンス支援システムについての紹介を行った。データサイエンスは、近年特に大きく注目されているが、

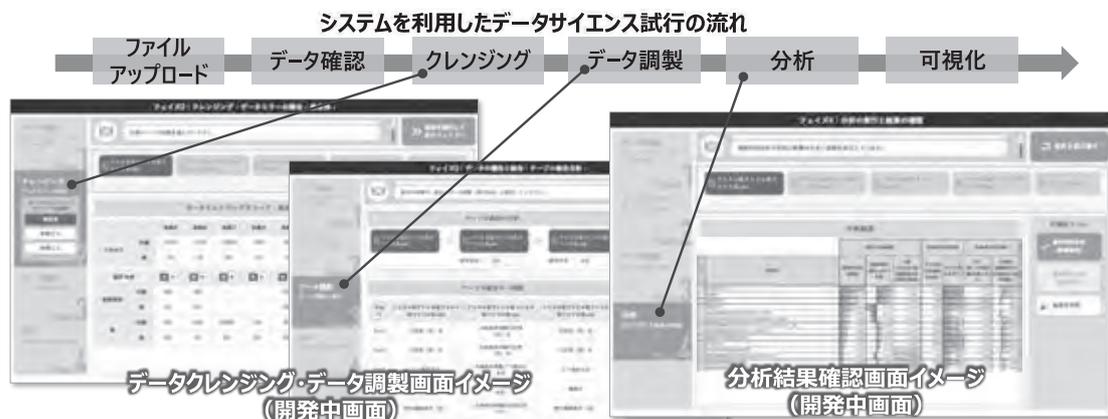


図-9 プロトタイプシステム利用の流れと開発中画面イメージ

冒頭でも述べたとおり，データサイエンティストの不足が課題となっている。データサイエンス実践の流れをより多くの人に理解してもらい，また，その分析結果の解釈までを理解し，業務とのつながりとして理解されることで，インフラ分野においてデータサイエンスを活用する機会が増えることを目指していきたい。

## 謝 辞

本研究開発の一部は，総務省 SCOPE #162103120, #181503004 の委託および，JSPS 科研費 JP00377847 の助成を受けたものである。



## 《参考文献》

- 1) The World Bank : World Development Report 1994 : Infrastructure for Development, Oxford University Press, 1994.
- 2) 杉崎光一, 家入正隆, 北原武嗣, 長山智則, 河村圭, 松田浩 : 維持管理のイノベーションのためのモニタリング実装方法に関する研究, 土木学会論文集 F3 (土木情報学), 第 73 巻, 第 2 号, pp. II\_17-II\_32, 2017.
- 3) 貝戸 清之, 起塚 亮輔, 伊藤 哲男, 橋爪 謙治, 出口 宗浩 : 床版かぶりコンクリートの剥離・剥落発生リスクと最適点検政策, 土木学会論文集 F4 (建設マネジメント), 第 68 巻, 第 1 号, pp. 11-27, 2012.
- 4) 全 邦釘, 嶋本 ゆり, 大窪 和明, 三輪 知寛, 大賀 水田生 : ディープラーニングおよび Random Forest によるコンクリートのひび割れ自動検出手法, 土木学会論文集 F3 (土木情報学), 第 73 巻, 第 2 号, pp. I\_297-I\_307, 2017.
- 5) Maeda Keisuke, Takahashi Sho, Ogawa Takahiro, Haseyama Miki : Distress classification of road structures via decision level fusion, 2016 IEEE International Conference on Digital Signal Processing (DSP2016), pp. 589-593, 2016.
- 6) 野村 泰稔, 村尾 彩希, 阪口 幸広, 古田 均 : 深層畳み込みニューラルネットワークに基づくコンクリート表面のひび割れ検出システム, 土木学会論文集 F6 (安全問題), 第 73 巻, 第 2 号, pp. I\_189-I\_198, 2017.
- 7) Koch Christian, Georgieva Kristina, Kasireddy Varun, Akinci Burcu, Fieguth Paul : A review on computer vision based defect detection and condition assessment of concrete and asphalt civil infrastructure, Advanced Engineering Informatics, 第 29 巻, 第 2 号, pp. 196-210, 2015.
- 8) Paar Gerhard, Caballo-Perucha Maria d P, Kontrus Heiner, Sidla Oliver : Optical crack following on tunnel surfaces, Optics East 2006, 第 6382 巻, 第 638207 号, pp. 1-8, 2006.
- 9) Yamaguchi Tomoyuki, Hashimoto Shuji : Fast crack detection method for large-size concrete surface images using percolation-based image processing, Machine Vision and Applications, 第 21 巻, 第 5 号, pp. 797-809, 2010.
- 10) Liu Zhiwei, Suandi Shahrel A, Ohashi Takeshi, Ejima Toshiaki : Tunnel crack detection and classification system based on image processing, Electronic Imaging 2002, pp. 145-152, 2002.
- 11) 皆川 勝, 佐藤 茂, 上谷 丈和 : 事例ベースを援用した知識洗練機能付診断エキスパートシステムの開発, 土木学会論文集, 第 595 号, pp. 67-76, 1998.
- 12) 有賀貴志, 矢吹信喜, 新井泰 : 変状データを含む開削トンネルのプロダクトモデルの構築, Journal of Japan Society of Civil Engineers, Ser. F3 (Civil Engineering Informatics), 第 68 巻, 第 1 号, pp. 58-70, 2012.
- 13) 吉田 郁政, 本城 勇介, 大竹 雄 : EM アルゴリズムを用いた劣化曲線群の同定法, 土木学会論文集 A1 (構造・地震工学), 第 69 巻, 第 2 号, pp. 174-185, 2013.
- 14) 中津 功一郎, 古田 均, 野村 泰稔, 高橋 亨輔, 石橋 健, 三好 紀晶 : GA による実用化を目指した長期的な橋梁維持管理計画策定, 知能と情報, 第 23 巻, 第 4 号, pp. 469-479, 2011.
- 15) Conway Drew : THE DATA SCIENCE VENN DIAGRAM, 2010.
- 16) Fayyad Usama, Piatetsky-Shapiro Gregory, Smyth Padhraic : The KDD process for extracting useful knowledge from volumes of data, Communications of the ACM, 第 39 巻, 第 11 号, pp. 27-34, 1996.
- 17) Shearer Colin : The CRISP-DM model: the new blueprint for data mining, Journal of data warehousing, 第 5 巻, 第 4 号, pp. 13-22, 2000.
- 18) 東京大学大学院情報学環「情報技術によるインフラ高度化」社会連携講座 (第 2 期) : 社会連携講座の活動成果の概要 2014-2016, 2019.
- 19) Chuprina Svetlana, Alexandrov Vassil, Alexandrov Nia : Using ontology engineering methods to improve computer science and data science skills, Procedia Computer Science, 第 80 巻, pp. 1780-1790, 2016.
- 20) Panov Panče, Soldatova Larisa, Džeroski Sašo : Ontology of core data mining entities, Data Mining and Knowledge Discovery, 第 28 巻, 第 5-6 号, pp. 1222-1265, 2014.
- 21) 古崎晃司, 來村徳信, 佐野年伸, 本松慎一郎, 石川誠一, 溝口理一郎 : オントロジー構築・利用環境「法造」の開発と利用, 人工知能学会論文誌, 第 17 巻, 第 4 号, pp. 407-419, 2002.
- 22) Wakuda Yuki, Isobe Masayuki, Ishikawa Yusho : Data Science Support System for Infrastructure Maintenance on Cloud, 17th International Conference on Computing in Civil and Building Engineering ICCCB 2018, 2018.

## 【筆者紹介】

湧田 雄基 (わくだ ゆうき)  
北海道大学  
数理・データサイエンス教育研究センター  
特任准教授



阿部 真育 (あべ まいく)  
北海道大学  
数理・データサイエンス教育研究センター  
特任助教

