

# インフラ維持管理におけるデータサイエンスの実践

湧田雄基・磯部正幸・石川雄章

本稿では、土木インフラ維持管理分野を対象としたデータサイエンス試行における分析の流れや課題について紹介する。中でも、特にデータクレンジングやデータ調製についての課題について、これらの課題に対応しデータサイエンスを手軽に試行可能とするクラウド型の支援システムのプロトタイプ開発についての報告を行う。

キーワード：インフラ維持管理、データサイエンス、データ準備、業務支援

## 1. はじめに

社会インフラは、我々の安心・安全に暮らし、産業および商業活動を行うために不可欠である。この社会インフラに関して、戦後の経済成長期に建設された多数の建造物の高齢化が一気に進展し、維持管理の業務負担が急増することが予想されている。こうした状況に加え、ビッグデータ活用技術の発展を受け、本稿で対象とするインフラ維持管理の分野において、業務で蓄積されているデータにAIを適用し人を支援できる可能性に期待が高まっている。

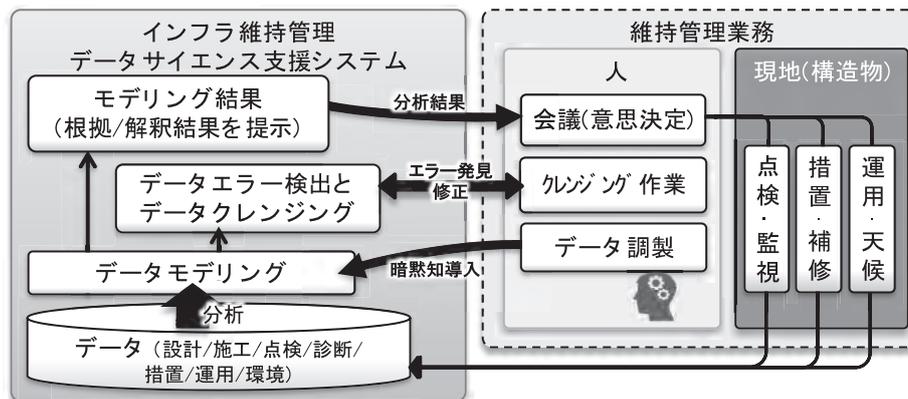
インフラ維持管理の業務では、点検記録等の資料に基づく意思決定が基本となるため、データを活用するAIやデータサイエンス技術<sup>1)</sup>との親和性は高いと考えられる。関連する既往研究では、点検記録に基づいた診断をAIで代替する応用<sup>2)~4)</sup>等が積極的に実施されており、コンクリートや舗装面の変状発見におけるAIの活用に関してはレビュー論文も発表され、多

数の研究が報告されている<sup>5), 6)</sup>。

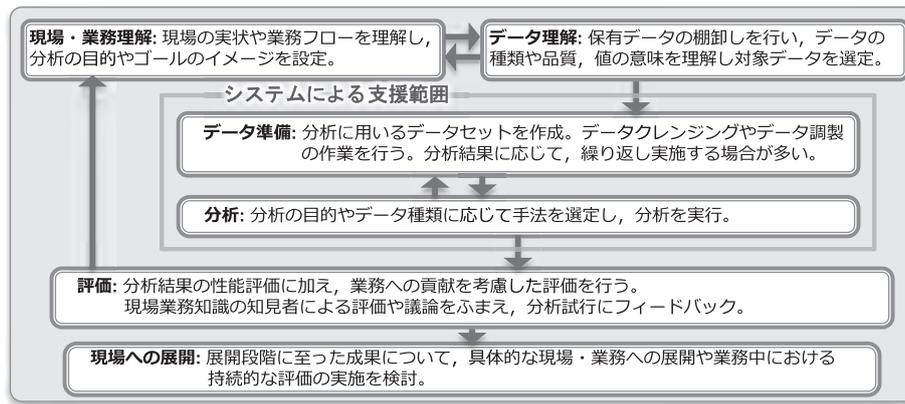
AI活用の試行としては、データ分析ツールにデータを投入すれば、分析結果を得ることはできるものの、結果を業務に有効に活用するためには、分析結果の正しい理解と業務上の経験や工学的知見に照らした議論が重要である。本研究では、コンピュータが得意とする処理を最大限に活用し「人」が行う意思決定に、より多くの時間を費やす事ができるよう支援を行う「データサイエンス試行支援システム（以降、支援システム）」について開発を実施している（図—1）。本稿では、これらの成果について紹介する。

## 2. インフラ維持管理におけるデータサイエンスプロセス

インフラ維持管理においてデータサイエンスを活用するための流れについて、著者らの既往研究では、次の6つのプロセスにより定義を行っている：現場・業



図—1 インフラ維持管理におけるデータサイエンス試行支援システムの利用のイメージ



図一 2 インフラ維持管理におけるデータサイエンスプロセス

務理解, データ理解, データ準備 (データクレンジング, データ調製), 分析, 評価, 現場への展開 (図一 2)。これらのプロセスは, データマイニングのプロセスとして良く知られる CRISP-DM および KDD を元に, 本研究においてインフラ維持管理分野に適合するようカスタマイズしたものである<sup>7)~9)</sup>。

インフラ維持管理においては, 「維持管理の効率向上」や「維持管理の質向上のための知識の抽出」がデータサイエンスの目的となり得る。従って, 本研究では, CRISP-DM で定義されている「ビジネス理解」プロセスを「現場・業務理解」とした。

また, 「評価」プロセスにおいては, 分析性能が良いかどうかだけでなく, 現場への導入を想定した場合の「業務への貢献」についても評価の視点となる。このとき, 業務上の制約や業務で目指す KPI (業務負荷低減, リスク低減 等) を考慮した上で, 妥当な分析性能を選定するといった検討も必要である。従って, ここでは, データに対する理解だけでなく, 結果を応用するための業務を熟知しているチームと議論し, 活用に向けた具体策を組み立てる事が重要である。

本稿で報告を行う支援システムは, 上で述べたデータサイエンスプロセスのうち, 「データ準備」, 「分析」, 「評価」のプロセスを支援の対象としている。

### 3. インフラ維持管理分野におけるデータサイエンス実践の課題

著者らは, 東京大学大学院情報学環 社会連携講座「情報技術によるインフラ高度化」において, 鉄道や高速道路などを保有するインフラ管理会社と共同しデータサイエンス研究を実施してきた<sup>10)</sup>。この取組みを通じ, インフラ維持管理分野におけるデータサイエンスにおけるデータ処理・分析に関する課題について次の3つを挙げる。第一に「データクレンジングの

課題」, 第二に「データ調製の課題」, 第三に「分析結果の解釈に関する課題」である。ここでは, これらの課題について概説する。

#### (1) データクレンジングの課題

維持管理業務にて実際に利用されているデータには, 多くの場合には不具合が含まれている。業務におけるデータ収集は, 担当者, 委託先企業, 事務所, 支社など, 様々な人や機関が関わっている上, 長期間の運用においてデータや業務の仕様が変更される場合も有る。そのため, データ分析に際し, 蓄積されているデータを集約すると, 多くの場合には, データのバラツキや品質が課題となる。本研究では, データサイエンスに活用するために問題となるデータ上の不具合をデータエラーと呼び, このデータエラーの除去作業をデータクレンジングと呼ぶ。大規模なデータを対象とするデータクレンジングでは, クレンジングの精度を維持しながら, 処理をできるだけ自動化する事が望まれるが, データエラーの検出やクレンジング作業においては, データ値の意味を考慮する必要がある。作業の自動化を難しくする要因となっている。

#### (2) データ調製の課題

データ調製は, 分析に投入するデータの形式を整える作業である。分析結果は, 入力するデータにより大きく変化するため, 分析の目的に応じて, 入力データを適切に編集することで, 分析結果の精度や解釈性を向上させるなどのコントロールが可能である。このデータ調製の操作により, 分析に対して人が持つ知見を導入する事も可能となる。変数選択やデータ粒度の設定など, データ調製の方針は慎重に設計する必要がある。頻繁に用いるデータ調製の処理には, テーブル結合, 変数選択, データカラム統合等がある。テーブル統合においては, データレコードの粒度が異なる場

合には、値の統合や再集計等の処理が必要となる場合もある。これらの処理は煩雑であり、分析対象データが個別のデータテーブルに分割されて記録されている場合には、何度も実施する必要がある。

### (3) 分析結果の解釈に関する課題

インフラ維持管理において、分析結果を実業務に導入するためには、分析の手続きや処理の内容・結果を解釈し、過去の知見に照らした上で納得できるかどうか、重要視される場合が多い。分析結果が導かれた理由についての納得感を得られないままに、分析結果を業務で活用することは受け入れられ難い。即ち、分析処理が優先して利用している変数や値の把握、分析結果が得られた根拠の把握が重要となる。しかしながら、一般に、データ分析結果は、分析に慣れたデータサイエンティストが理解できる表現により出力される場合が多く、維持管理業務に詳しい者のために必ずしも分かりやすい表現にはなっていない。従って、分析に詳しい者と業務に詳しい者など、異なる専門家のチームにより議論しながら、分析結果の解釈と分析試行を繰り返す事が必要となる。

以上で述べた各課題により、インフラ維持管理におけるデータサイエンスは、試行錯誤に多くの時間を要する。そこで、データサイエンスの試行錯誤を加速し、誰でも試すことができるように「データクレンジング」や「データ調製」の処理をボタン化しシステムで支援する支援システムの開発を進めた。

## 4. インフラ維持管理におけるデータエラー

著者らのこれまでの研究において、インフラ維持管理において頻繁に見られるデータエラーについて、典

型例の整理を行った。データエラーの例を表—1に挙げる。

#1～#5は、値自体に不具合があり、データ調製を実施する前に可能な限り解決すべきデータエラーである。また、#4～#10のデータエラーは、データクレンジングが困難な場合も含むため、主にデータ調製により処理の検討を行う。このうち、#9と#10は人間の作業によるデータの作成を伴うため、データが持つ情報量への期待と、データ調製の負担とのバランスに基づいて、実施を検討することとなる。

## 5. データクレンジング

データクレンジング作業は、データ分析の信頼性や精度向上のために、データエラーを修正する操作である。データエラーの検出においては、データの基本統計の計算により値のバラツキや分布の確認を行う事や、グラフや地図を用いた可視化も有効である。

データエラーには、「値の有無」、「値の正しさ」、「値の表現形式」等により、段階がある事に注意が必要である。図—3は、データエラーの段階とクレンジング可否との対応のイメージを図化したもので、Olsonらによる整理を元に、インフラ維持管理での具体的な例を図下段に記載した。

図中のデータエラーの具体例として、橋梁の名称や座標値を例として示した。

例えば、値の有効性に関して、値が無く空欄である場合には、値は欠損という扱いとなる。また、値が入っていたとしても、橋梁名が入るべき箇所に数字が入っている場合や、あるいは、経度が入るべき箇所に0が入っている場合や、正しくはA橋梁Bランプというデータが入るべき箇所に、C橋梁Dランプと記載さ

表—1 インフラ維持管理におけるデータエラーの例

#	データエラー	説明及び具体例
1	表現の不整合	全角と半角の差異。ローカル表現による表記の揺れ。
2	変換ミス、タイポ	漢字の変換ミスや、タイプ入力ミス。
3	外れ値	他の値に比べて大きく異なる値。
4	値の欠損	データ値が空欄の場合、値が登録されていないのか、点検の結果により異常無と判定されたのかは不明。
5	デフォルト値	実際の値とは異なり、年度に0や、センサ値に計測上限値などの値が登録されている。
6	データの意味が混在	判定基準の仕様変更等により、値の表現が混在（B判定とA3判定等）。
7	複数の値が共存	カンマ区切り（1.23,4.56）や、範囲指定（0078-0090）等、一つのセルに複数の値が存在。
8	記録時期の差異	値の取得時期がレコード毎に異なる。5年で一巡する点検方式の場合に、構造物毎に計測時期が異なっている。
9	データの機械可読性	点検記録や図面が、紙の台帳や紙をスキャンした画像やPDF形式で管理されている。
10	自由文での記録	現場での所感等の補足情報や、詳細情報が自由文で入力されている。

値の有無	値が存在				値が欠損
値の正しさ	値が正しい		値が誤り		
値の表現形式	正しい表現形式	誤った表現形式			
値の解釈	正しく有効な値が正式な表現で入っている	正しく有効な値が入っているが表現形式が誤り	値があるが誤っている(値が無効)		値が無い
具体的例1:橋梁名 (正:A橋梁Bランプ)	A橋梁Bランプ	a橋梁bランプ	C橋梁Dランプ	0.28, 0, 2015	(空欄)
具体的例2:経度 (正:139.1234567)	139.1234567	139度7分40.74秒	35.1234567	0	(空欄)
	OK	クレンジング可	クレンジング困難(データ調製により解決を検討)		

図-3 データエラーの段階 (Olson らの報告<sup>11)</sup>に基づく)

れている場合、値は登録されているものの、値は誤りでありそのままではこれらのデータは利用できない。

最後に、値の表現形式について、A 橋梁 B ランプと記載されているべきところが、a 橋梁 b ランプと記載されている場合には、意味は、誤りではなさそうに思われるものの、表現形式が誤っている。

以上に例示したデータエラーを発見し、修正していくという操作が、データクレンジングであるが、検出したすべてのエラーを修正できるとは限らないという事に注意が必要である。

## 6. データ調製

データ調製とは、分析の目的に応じて、データの形式を変更する操作である。例えば、表形式データの場合には、①変数選択、②変数操作、③データテーブルの統合、④データレコードの選択の4つの手順がある。データ調製は、分析結果の評価をふまえて複数回実行するが多い。

- ①変数選択：データ列毎に解析に用いるデータを選択する操作である。分析対象として準備したデータより、「分析への貢献の可能性」、「有効値の割合」、「データエラーの量」等を加味し、変数の取捨選択を行う。また、対象とするデータサイズを絞り込むために、最初に検討することも有効である。
- ②変数操作：「変数の生成」、「データ値の変換」、「変数の統合」、質的データを定量化データに変換する「値の数量化」、多数の値を単一のカテゴリに統合する「値のカテゴリ化」等がある。
- ③データテーブルの統合：インフラ維持管理では、業務に応じたデータが別々の台帳で管理されている事より、電子データも別々のデータテーブルで管理されている場合が多く、データテーブルの統合処理は頻繁に行う。例えば、構造諸元、検査台帳、補修台帳などである。これらのデータテーブルの統合にお

いては、テーブル間で対応関係のあるデータレコードを探す必要がある。そのために、ユニークキー等を用いる。統合の際のユニークキーには、例えば、道路や鉄道等の長尺構造物では、距離基準値を参照する。また、データテーブル毎に、データレコードの集計単位が異なる場合には、テーブル統合のために、データレコードの再集計が必要となる場合もある。

- ④データレコードの選択：データ調製の最後に実施する場合が多い。これは、以上のデータ調製の結果に応じて、値の充填率等のデータクオリティを確認した上で、使えるデータレコードの選択を行う場合が多いためである。

## 7. データサイエンス試行支援システム

### (1) 背景

本章では、前述のデータサイエンスのプロセスに従って、データクレンジングとデータ調整、分析と評価を行う支援システムについて紹介する。

データサイエンスに活用できるソフトウェアツールについて、近年、多くの研究開発が報告されている。例えば、汎用的なデータサイエンスプラットフォームに関する研究<sup>12)</sup>や、放射線腫瘍学のデータサイエンスプラットフォーム<sup>13)</sup>、スマートグリッド向け<sup>14)</sup>や、スマートシティ向け<sup>15)</sup>のデータ解析プラットフォーム等がある。また、データクレンジングに特化したツール<sup>16), 17)</sup>や、オントロジーを用いた自然言語処理に基づくデータクレンジングツール<sup>18)</sup>も報告されている。

データサイエンスを実行するためのツールとしては、Microsoft Excel を用いた可視化や、統計量の算出、簡単な多変量解析も有効である。また、統計、多変量解析、機械学習には、R, Python, MATLAB 等のプログラミング環境も活用される。分析ソフトウェアパッケージには、Weka, SPSS がある。さらに、

Azure ML や Amazon Machine Learning などのパブリッククラウドサービスも提供されている。ただし、これらのツールは、さまざまな分析方法に対応でき汎用性が高く、データの扱いに手慣れた情報系技術者や各ドメインの専門的な研究者によるデータサイエンス試行をサポートするものとなっている。

一般に、データ分析試行においては、ほとんどがデータ準備（データクレンジング、データ調製等）に時間を要するとされる。このような課題に対し、データ準備を支援するツールも多くリリースされている。Trifacta<sup>19)</sup> と Google Cloud Dataprep<sup>20)</sup> は、データ前処理のための専門サービスとして使用される。データエラーの発見とデータ準備の結果可視化においては、Tableau<sup>21)</sup> や BI (Business Intelligence) ツールも有効である。

以上で示した既往研究およびツールは、汎用的なものであり、どちらかという、手慣れたデータサイエンティスト向けである。我々の提案システムでは、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したいと考えており、インフラ維持管理分野におけるデータサイエンスに有効な手続きを備え、初学者でも迷わずデータサイエンスを試行できるツールとして開発を進めている。

## (2) 提案システムの特長

提案システムでは、これまでの研究内容に基づき、インフラ維持管理分野のデータを予め解析し、クレンジングに有用な辞書やデータ調製処理を備えている事が大きな特長である。汎用性を提供するのではなく、インフラ整備管理分野を中心としたデータを対象とし、データサイエンスの試行を提案することで、ユーザーが迷わずに操作できるツールを目指している。

また、複数のユーザーが同時に分散データを準備し、データと結果を共有できるように、クラウドベースのプラットフォームとして実装している。データサイエンスを実行するには、データエラーの発見、エラーのクレンジング、データの準備など、さまざまなデータにさまざまな手順を適用するノウハウが必要であるが、これらのノウハウをテンプレート化し、共有が可能となっている。ドメイン知識、データ誤りの辞書、およびデータ準備手順に基づく様々な知識を保持し、知識データベースとして保持する事が可能である。また、データの種類や分析の目的に応じて分析方法を選択し、進めるデータサイエンス全体の具体的な手順を「レシピ」と呼び、このテンプレートを保持す

る事ができる。これらのノウハウ「レシピ」は過去の研究事例から作成され、クラウドを使用して知識データベースとして蓄積される。この「レシピ」をコア技術として、ノウハウを蓄積し共有することが可能である。

以上の特長により、ユーザーは「多機能」なツールに苦しむことなく、有効な手順を選び、分析を簡単に試すことが可能となる。一連のデータサイエンス手順を一度試すことで、ユーザーにデータサイエンスの経験を与える事ができる。

## (3) システムの構成

システムは、サーバーホストとワーカーホストで構成する（図-4）。サーバホストは、ユーザとのデータの入出力や対話を行い、データ解析レシピ（知識ベース）を管理する。データクレンジングのための分析レシピおよび辞書は、知識データベースとして RDF に格納される。RDF 形式で管理することで、様々な分析方法を他のデータに適用するなどの汎用的な管理が可能である。ワーカーホストでは、データのクレンジング、データ構成、データ解析、評価処理を実行する。処理前後のデータテーブルは、作業側側のデータ格納部に格納される。また、データ解析を行う計算エンジンでは、解析処理の動作状況によって負荷が変動するため、ワーカーホスト上で実行されている複数の Docker コンテナに分割し、計算エンジンを実行する仕組みとなっている。



図-4 提案システムの構成

## (4) 利用手順と機能

本研究では、データサイエンスの一連のデータ操作手続きを自動的に適用する機能をレシピと呼ぶ。なお、一連のデータサイエンス手順のうち、例えば、前処理等の一部分を定義したレシピも設定可能である。支援システム利用の具体例として、インフラ検査結果のデータセット（複数のテーブルを含む）に対して判別分析を行う場合について、以下に紹介する。

ユーザの最初の操作として、データのクラウドへの

図一五 提案システムの開発中画面

アップロードを行うと、クラウド側はデータの変数毎の値の種類（定性的データ／定量データ）に基づき、基本統計を算出し、値のユニークリストと、出現頻度（度数分布）、値の充填率等を計算する。図一五に、基本統計の算出結果確認の画面（開発中）のイメージを示す。値の度数分布表示は、外れ値や値の揺れを把握するために有効である。このビューにおいて、値が一致しない値を選択して変更することができ、この画面でデータのクレンジングが可能である。ここで行われるクレンジング処理は、処理のレシピとして登録され、名寄せルールとして蓄積される。

次に、システムが保有する辞書やルールに基づいたデータクレンジングを実行する。本システムでサポートするデータクレンジングの具体例を示す。データクレンジングには、辞書型とルール型の2通りを用いる。また、変数によらず汎用的に適用が可能なクレンジングと、変数に固有のクレンジングを準備し、変数認識処理を行い、使い分ける（図一六）。汎用的なデータクレンジングについては、半角・全角変換や、不要文字除去等の機械的に実行が可能な処理を備える。

次に、データ調製プロセスとして、変数の数値化を実行し、文字列で登録されている質的データの数値化を行う。続いて、複数のデータテーブルの統合処理を行う。目的変数と説明変数選択を行う。説明変数選択は、変数同士の相関や、VIF 値（Variance Inflation

Factor）算出による多重共線性の確認などに基づいて実施する。

判別分析を実行し、分析結果を混同行列や P-R 曲線で評価を行う。

以上で挙げた一連の処理は、一般的なパターンであり、同様の手順による他のデータセット（例えば、他のオフィス、鉄道線、道路橋のデータ）に適用することも可能である。本システムでは、既存のデータサイエンス試行に基づいて定義されたレシピにより、データサイエンスプログラムの進捗状況に応じて、実行可能な次のステップの提案が可能である。

### 8. おわりに

本稿では、インフラ維持管理分野におけるデータサイエンスの流れを定義し、その試行における課題の整理を行った。データサイエンスにおけるデータ準備（データクレンジング、データ調製）や分析を支援し、データサイエンス試行を簡便に行うための基本機能を備えるシステムのプロトタイプ開発についての報告を行った。本研究の提案システムは、現段階では、限られたユーザによる試用を実施しているが、今後は、分析手法の拡充や運用の安定性の向上を経て、実用化を図りたい。

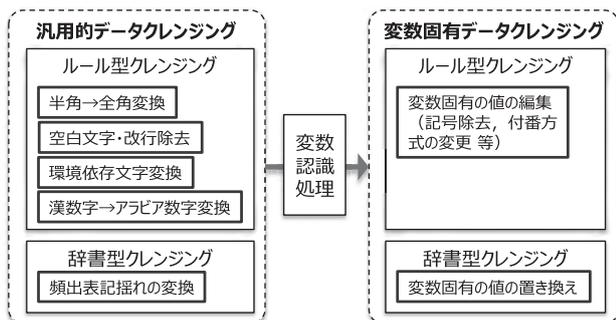
### 謝 辞

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。

JCMA

### 《参考文献》

- 1) Cao, L. Data science: Challenges and directions, Communications of the ACM, 60(8), pp. 59-68, 2017.
- 2) 本城 勇介, 大竹 雄, 佐藤 敦, “岐阜県橋梁点検データベースの統計解析に基づく簡易橋梁健全度評価指標”, 構造工学論文集, A 日本学術会議土木工学・建築学委員会 [監修], vol. 60, pp. 462-474,



図一六 データクレンジング処理の概略フロー

- 2014/03, 2014.
- 3) 貝戸 清之, 福田 泰樹, 起塚 亮輔, 橋爪 謙治, 出口 宗浩, 横山 和昭, "遊離石灰法に基づく RC 床版の劣化予測および補修優先順位の決定法". 土木学会論文集 F4 (建設マネジメント), 68 (3), 123-140, 2012.
  - 4) 山崎崇央, 石田哲也, 生存時間解析を用いた東北地方における橋梁コンクリート部材の劣化定量分析, 土木学会論文集 F4 (建設マネジメント), 71 (4), pp. I\_11-I\_22, 2015.
  - 5) 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, vol. 29, no. 2, pp. 196-210, 2015.
  - 6) Zakeri H, Nejad Fereidoon Moghadas, Fahimifar Ahmad, "Image Based Techniques for Crack Detection, Classification and Quantification in Asphalt Pavement: A Review," Archives of Computational Methods in Engineering, pp. 1-43, 2016.
  - 7) Fayyad, U., Piatetsky-Shapiro, G., and Smyth, P. (1996). "The KDD process for extracting useful knowledge from volumes of data." Communications of the ACM, 39 (11), 27-34.
  - 8) Shearer, C. (2000). "The CRISP-DM model: the new blueprint for data mining." Journal of data warehousing, 5 (4), 13-22.
  - 9) Kurgan, L. A., and Musilek, P. (2006). "A survey of Knowledge Discovery and Data Mining process models." The Knowledge Engineering Review, 21 (1), 1-24.
  - 10) 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, <http://www.advanced-infra.org/english.html>, accessed on January 29, 2018.
  - 11) Olson, J. E. (2003). Data Quality: The Accuracy Dimension, Elsevier Science.
  - 12) Patterson, E., McBurney, R., Schmidt, H., Baldini, I., Mojsilović, A., and Varshney, K. (2017). "Dataflow representation of data analyses: Toward a platform for collaborative data science." IBM Journal of Research and Development, 61 (6), 9: 1-9: 13.
  - 13) Marungo, F., Robertson, S., Quon, H., Rhee, J., Paisley, H., Taylor, R. H., and McNutt, T. (2015). "Creating a data science platform for developing complication risk models for personalized treatment planning in radiation oncology." Proc., System Sciences (HICSS), 2015 48th Hawaii International Conference on, IEEE, 3132-3140.
  - 14) Simmhan, Y., Aman, S., Kumbhare, A., Liu, R., Stevens, S., Zhou, Q., and Prasanna, V. (2013). "Cloud-Based Software Platform for Big Data Analytics in Smart Grids." Computing in Science & Engineering, 15 (4), 38-47.
  - 15) Khan, Z., Anjum, A., and Kiani, S. L. (2013). "Cloud Based Big Data Analytics for Smart Future Cities." Proceedings of the 2013 IEEE/ACM 6th International Conference on Utility and Cloud Computing, IEEE Computer Society, 381-386.
  - 16) Zhang, F., Xue, H.-F., Xu, D.-S., Zhang, Y.-H., and You, F. (2013). "Big data cleaning algorithms in cloud computing." International Journal of Online Engineering (IJOE), 9 (3), 77-81.
  - 17) Faruquie, T. A., Prasad, K. H., Subramaniam, L. V., Mohania, M., Venkatachaliah, G., Kulkarni, S., and Basu, P. (2010). "Data cleansing as a transient service." Proc., 2010 IEEE 26th International Conference on Data Engineering (ICDE 2010), 1025-1036.
  - 18) Kontokostas, D., Brümmer, M., Hellmann, S., Lehmann, J., and Ioannidis, L. (2014). "NLP data cleansing based on linguistic ontology constraints." Proc., European Semantic Web Conference, Springer, 224-239.
  - 19) Trifacta (2018), Trifacta: Data Wrangling Tools & Software, <https://www.trifacta.com/>, accessed on January 29, 2018.
  - 20) Google Inc. (2018), Google Cloud Dataprep, <https://cloud.google.com/dataprep/>, accessed on January 29, 2018.
  - 21) Tableau Software (2018), Tableau, <https://www.tableau.com/>, accessed on January 29, 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



磯部 正幸 (いそべ まさゆき)  
ファイブソリューションズ株  
CEO



石川 雄章 (いしかわ ゆうしょう)  
東京大学 大学院情報学環  
特任教授

