

労働災害知識モデルに基づく安全 AI システムの開発

高 藤 淳・鈴木 理 史

建設業で発生する様々な労働災害は、主にワープロや表計算ソフトによる報告書の形態で記録されている。しかし、原記録のままでは、高度にデジタル活用することは難しい。そこで、筆者らは、データ科学の手法により労働災害知識を抽出し、労働災害モデルの形態に当該知識を構造化することで、労働災害防止に寄与する安全 AI システム（以下、本システムも同義）を開発した。本稿では、安全 AI システムの技術的特徴と労働災害防止への貢献性について報告し、現場での PoC（概念実証）の展望にも言及する。

キーワード：AI、データ科学、知識構造化、ダイナミック・チェックリスト、労働災害防止

1. はじめに

近年、建設業においても他産業同様、デジタル・トランスフォーメーション（DX）が進展している。建設業 DX では、現場業務を事務所に持ち帰らず現場で完結するなど、主に業務効率化の観点から実施されることが多い。一方、安全衛生管理は、その重要性は共有されつつも、効率化・省力化・低コスト化の視座では DX の対象となり難い側面が見受けられる。実際、筆者らは、複数の企業と、労働災害データの活用を協議する機会を得たが、いずれも当該データがデジタル活用に向けた整備途上の水準であることを確認している。ただ、多数の企業から、労働災害データをデジタル活用したい、特に、AI 的手法で活用したいという意向が示されていることから、安全 DX の推進機運が高まっているものと推察される。

安全 DX の最初のステップとしては、労働災害データベースの構築や、労働災害報告書を対象とした情報検索システムの構築などが取り組み易い。ただ、これらの手法は、過去の情報へのアクセスを容易にする利点はあるが、労働災害データに潜在する真のリスク因子、リスク因子と発現リスクの関係性、リスク要因の類同性、などを明らかにしない。

ここで、建設業の労働災害データの基本的特性として、まず、労働災害が高頻度で発生するものでは無く、企業や事業所の単位では、累積的にも数百件から数千件の規模である量的特性があげられる。次に、コード体系や記述の精細度が人に依存して恣意的であるなどの質的特性が見られる。前記量的特性を踏まえると、

大規模なデータの存在を前提とした機械学習的手法は適さず、データ科学の手法を適用して潜在するリスク因子や関係性を顕在化することが有効である。また、前記質的特性を鑑みて、データの統制を図るとともに、記述の恣意性を均しつつリスク知識を構造化するモデリング手法が適している。

本システムでは、このような労働災害データの特性を踏まえて、データ分析のフェーズを設けた上で、労働安全衛生総合研究所（以下、安衛研）により提唱されている IMTOC 法¹⁾を用いた知識モデリングを行う手法を確立し、システム化を図った。

以下、2章で、労働災害データの分析概要を示し、3章で、分析結果に基づく労働災害知識のモデル構造化について述べる。最後に、安全 AI システムの構築について述べるとともに、本システムの社会実装の展望に関して言及する。

2. 労働災害データの分析

本稿が対象とする労働災害データは、実際に企業内で蓄積された12年分2,155件の労働災害報告書である。原データは表計算ソフトで作成されており、欠損値や内容エラー等が存在したため、データクレンジングを行った。さらに、分析用に構造化データ形式（XML）に変換した後、データ分析を実施した。

労働災害データは、一般に構造化データ領域と非構造化データ領域で構成される。構造化データ領域は、日時・場所等のメタ情報および被災者のプロフィール情報を含み、非構造化データ領域は、災害発生状況（以

下,「状況」), 災害発生原因 (以下,「原因」), 災害防止対策 (以下,「対策」)を含む。本分析では, 構造化データに対して, 定量的分析手法を適用し, 非構造化データに対して, 定性的分析手法を適用する。

(1) 定量的分析

定量的分析の目的は, リスク因子とリスクとの関係性, 傾向等を定量的に明らかにすることである。本分析の結果は, 後述のようにシステムに組み込まれて用いられる。

(a) 「事故の型」と単一属性項目のクロス分析

「事故の型」に対する属性項目は, 企業固有のコード体系により定義される「職種」「作業種」「起因物」と, カテゴリカル変数として取り扱う「年齢」「経験年数」「時間帯」である。本分析により, 個々の属性項目が, どのような事故に関連性が強いかが明確化される。一例として, 図-1に「年代」毎の「事故の型」の基本分布を示す。20代では他より相対的に「切れ・こすれ」が多く, 「転倒」に着目すると40代以降漸増するなどの傾向が読み取れる。本稿では割愛した「年齢」以外の属性についても, 「事故の型」との関連性に関する傾向が明確化されている。

(b) アソシエーション分析

一方, 複数の属性項目のどの項目組が特定の「事故の型」に固有の関連性を有するかを分析するために,

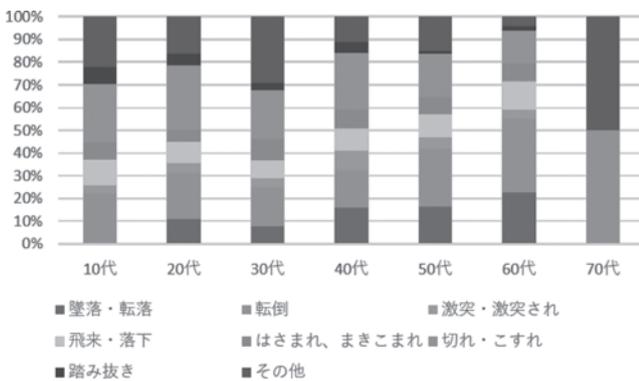


図-1 「事故の型」-「年代」の基本分布

アソシエーション分析の手法を適用した。アソシエーション分析には, 統計解析ソフトウェア R²⁾ (以下, R) の arules パッケージ³⁾を用いた。労働災害データに対するアソシエーション分析の結果の一部を表-1に示す。本分析の結果, 例えば, 『「30代」の「鉄骨工」が, 「S 躯体」に関連する作業を行う際, 「転落・墜落」事故に遭いやすい』等の複合的な要因による事故との相関性を定量的に得ることができる。

(2) 定性的分析

定量的分析に対して, 定性的分析の目的は, リスクの多様な特徴や性質を明らかにすることであり, 1) 「状況」中のリスクにつながる状況や条件の同定, 2) 「原因」及び「対策」の類型化が該当する。前者は, 被災者の不適格行動, 機械類の不適切な作用等を明らかにし, 気象の影響や環境要因を同定することを含む。また, 後者は, 労働災害データの原記述が本質的に同義の内容を異なる表現形態で表出化していることに対して, その本質的同義性を捉え集約することを意味する。

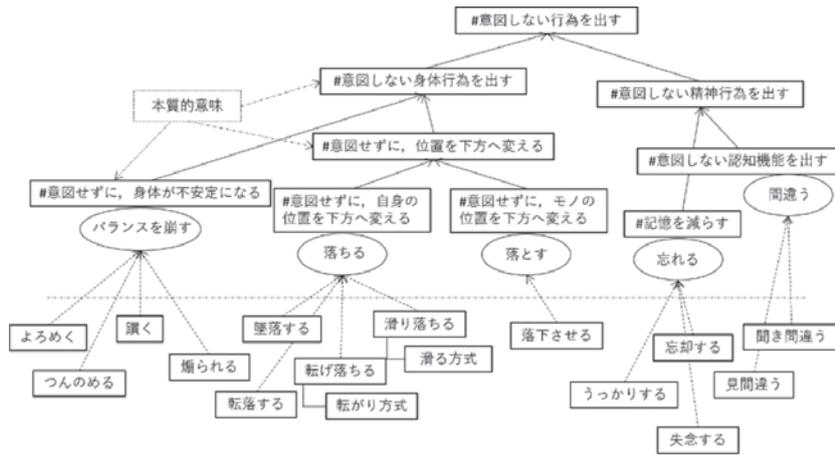
(a) 「状況」分析

「状況」記述に対して, テキスト解析を行うことにより, 災害に至る作業者の行為, 建設機械の不適切な機能発揮, 被災環境・条件など多彩な情報が得られる。本テキスト解析では, 係り受け解析器として CaboCha⁴⁾を用いた。CaboCha の解析結果を基に, 述語に対する格構造を同定することで, 労働災害に関連性が高い表現を抽出した。さらに, 抽出した語彙情報を基に, オントロジー工学の手法⁵⁾により本質的意味を同定することで, 「意図しない行為/機能」に関する基礎的な体系を構築した (図-2)。

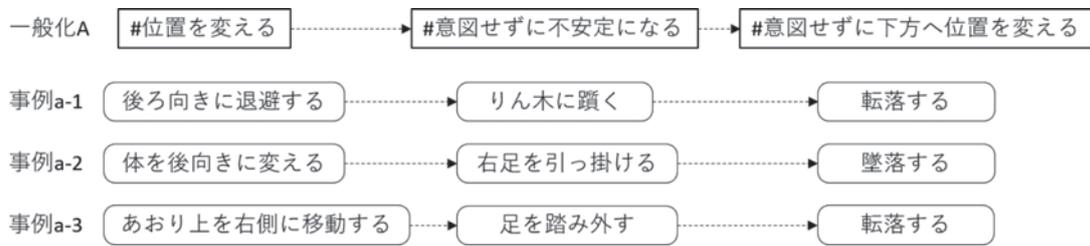
本体系は, 表層的な記述の形式的な相違を捨象し, 労働災害事例の本質的同義性の判断に寄与する。例えば, 図-3に示すように, 事例 a-1, 2, 3は各々具体的な事象としては異なるが, 一般化することで本質的等価性を確認することができる。

表-1 アソシエーション分析の結果 (部分)

IF X	THEN Y	lift
{職種 = 鉄骨工, 起因物 = S躯体, 年代 = 30s}	{転落・墜落}	8.654618
{作業種 = 塗装工事, 年代 = 60s, 時間帯 = 午前: 早}	{転倒}	7.751799
{作業種 = 解体工事, 起因物 = 工事機械等, 時間帯 = 午後: 遅}	{激突・激突され}	16.576923
{作業種 = ALC版取付け, 起因物 = 資機材, 工機具等, 入場後 = 100日}	{飛来・落下}	9.248927
{作業種 = 揚重機械組立解体, 起因物 = 定置式クレーン, 年代 = 20s}	{はさまれ・まきこまれ}	5.374065
{職種 = 土工, 手伝い, 作業種 = 運搬作業, 起因物 = 足場・建築物, 資機材等}	{崩壊・倒壊}	32.164179
{作業種 = 金物取付け, 起因物 = 資機材, 工機具等, 年代 = 40s}	{切れ・こすれ}	9.248927
{職種 = 鉄筋工, 起因物 = 釘, 金属片, スレート, 経験年数 = 1年, 時間帯 = 午前: 早}	{踏み抜き}	40.66038



図一 2 意図外行為の体系化例 (部分)



図一 3 個別事例の一般化

表一 2 「原因」類型

物的要因	人的要因	管理的要因
建設機械の不具合・劣化	建設機械の誤操作・不適切な稼働	安全衛生管理不十分
建設資機材の不良・劣化	建設資機材の不適切使用・誤操作	安全衛生教育不十分
建造物・部材の劣化	安全帯の不使用・不適切使用	作業責任者不在・指示不適切
保護具・服装の欠陥	保護具・服装の不備	誘導員／監視員の不在・不備
不安全・不衛生な作業環境	不安全行動	安全衛生措置不十分・不備
気象・自然環境の影響	単独作業	墜落防止措置不十分・不備
不意の危険遭遇	作業手順・指示の非順守	リスクアセスメント不十分・不備
	安全確認不十分・不備	作業計画不十分・不備
	経験／知識不足・無資格	不適切な作業手順
	危険認識不十分	不適切な作業方法
	認知的誤謬(不注意・錯覚・不可知)	関係者間の連絡・調整・連携不足
	心理的緊張緩和(慣れ・油断)	
	心理的不安定化(焦り・パニック)	
	体調不良・心身不備	
	作業者間の意思疎通不足・不備	
	危険予知不十分・不備	

(b) 「原因」分析

「原因」の類型化を図るために、クラスター分析の手法を適用した。本クラスター分析では、まず、全ての「原因」文に対して、潜在意味解析法(LSA)により、言語特徴量を計量化する(潜在意味解析はR²のsvdパッケージ⁶⁾を用いた)。次に、前記特徴量を基に、文同士のコサイン距離を測度として類似性を判定し、クラスターを形成する。最後に、分析者が、意味内容

の相違を斟酌し、調整を図ることで適正なクラスター群を得る。

本手法により得られたクラスター個々に、内容を現すラベルを付与することで、表一2に示す34個の「原因」類型に集約される分類結果を得た。ここで、本「原因」類型は、大分類的位置づけであり、必要に応じて任意の類型内でさらに詳細な細分類に区分することが可能である。

本「原因」類型を用いることで、類型単位の詳細な分析が可能となる。本類型に基づく「墜落・転落」では、図-4aに示すように人的要因が71%、管理的要因が23%を占め両者で90%を超える。また、前記人的要因は、図-4bに示すように上位の「不安全行動」「建設資機材の不適切使用・誤操作」「安全確認不十分・不備」「危険予知不十分・不備」が過半を占める。さらに、人的要因で1位の「不安全行動」と2位の「建設資機材の不適切使用・誤操作」を各々ドリルダウンすると、図-4cに示す結果が得られる。

(c) 「対策」分析

「対策」は、「原因」と対を成すため、「原因」類型毎に「対策」を集約することができる。その上で、本質的同義性に基づく統一化を図ることができる。ここで「対策」の記述に着目すると、日本語の特性に起因する主体の省略や対象の不明確さ等が見受けられる。つまり、「対策」すべき主体や「対策」の対象が曖昧

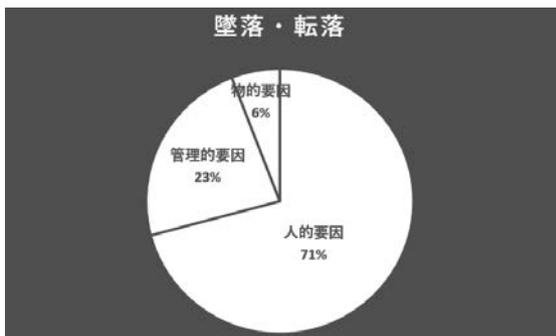


図-4a 「墜落・転落」の要因分布

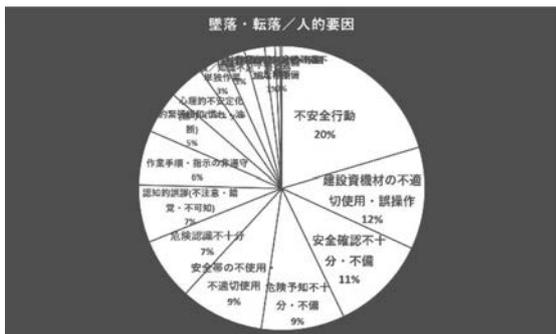


図-4b 「墜落・転落／人的要因」の詳細分布

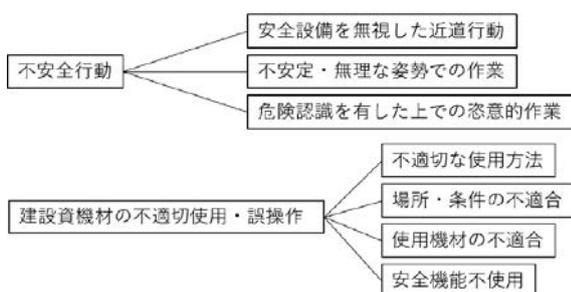


図-4c 「墜落・転落／人的要因」のドリルダウン例

であり、「誰」が「誰／何」に「どうする」かのメッセージ性が不明確であることを意味する。

そこで、筆者らは、「対策」のメッセージ性を明確化するため、下記のような統制表記を提案している。実施主体として、「事業者」「作業責任者」「作業員」の3区分を設けた上での統制表記の具体例を例1～3に示す。本統制表記を用いることで、対策主体に応じた実施方を動的に提示することが可能となる。

- ・「対策」統制表記:[対策主体] (が／は) [対策客体] (に／を) [行為 | 作用]
- ・例1. 事業者は、バックホウ運転者に、揚重作業ではクレーンモードを使用するよう周知する。
- ・例2. 作業責任者は、作業員に、荷下ろし時のルールを遵守させる。
- ・例3. 作業員は、搬出場所で、積み荷の荷姿・安定状態を確認する。

3. 労働災害知識のモデル構造化

任意の分野において、蓄積したデータの利活用の為には、内在する知識を構造化することの必要性が指摘されている^{7),8)}。労働災害データに関しても、輻輳する諸要件を整理し、多角的な利活用を図るため、労働災害知識のモデル構造化を行う必要がある。

(1) 拡張 IMTOC モデル構造

原 IMTOC モデルは、労働災害知識の構成要素を「I:産業」「M:起因物」「T:事故の型」「O:各種条件」「C:原因と対策」と規定する。本システムでは、「O:各種条件」と「C:原因と対策」に関して独自の詳細化を図り、拡張 IMTOC モデルとして再定義した。拡張 IMTOC モデルは、計算機内部でナレッジグラフとして表現される(図-5)。

「O:各種条件」のグラフパートには、前章の「状況」

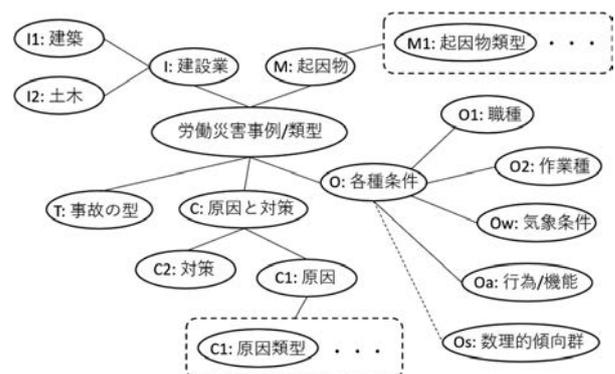


図-5 拡張 IMTOC ナレッジグラフ

の分析結果が反映される。労働災害において本質的な要件と考えられる人の「意図外行為」や機械類の「意図外作用」,「気象の影響」などが含まれる。また,「C:原因と対策」のグラフパートには,前章の「原因」と「対策」の分析結果が反映される。

ここで,本システムでは,予測すべきリスクを「リスク原因類型+事故の型」で定義する。「原因」は「対策」の源泉であり,災害の「何故起きたか」を明示することは,喚起すべき注意の核心的情報と考えられるからである。ナレッジグラフを用いることで,予測されるリスクの表現系を導出することができる。

(2) ダイナミック・チェックリスト

一般に,安全管理システムでは,想定されるリスクに対して,「被害の程度」×「発生頻度」の尺度を用いてリスク群が等級付けされる。しかし,このような一律の尺度では,リスク情報の需要者個々に,真に必要としている情報を提供することは困難である。

そこで,筆者らは,リスク情報をニーズに合わせて動的に提供するため,推薦システム⁹⁾の考え方を援用し,ダイナミック・チェックリスト機構を発案した。同機構の実現には,2章で行った定量的データ分析の結果が用いられる。

データ分析により,単一又は属性組と「事故の型」

との相関性が定量化されている。従って,予測されるリスク群に対して,各「事故の型」の予測リスク毎に式-1を適用することで,所与の任意の条件に対する関与性スコアを得ることができ,条件に応じた動的なレーティングが可能となる。式-1において,bは「原因」類型の発現頻度を基礎とするスコア, x_i は当該属性と当該「事故の型」との相関性を基礎とするスコア, a_i は x_i の重みを表す係数, c_i は人的補正定数である。bを基本スコア, Σ 項をペナルティスコアと見做すことができる。

簡単のため a_i を1, c_i を0として,式-1を適用した結果例を図-6に示す。本結果例の場合,20代では経験の無さによる危険への感度が低い様子が窺われ,50代では慣れからくる油断が生じている様子が窺われるなど,式-1による一定の有意性を読み取ることができる。

4. 安全 AI システムの構築

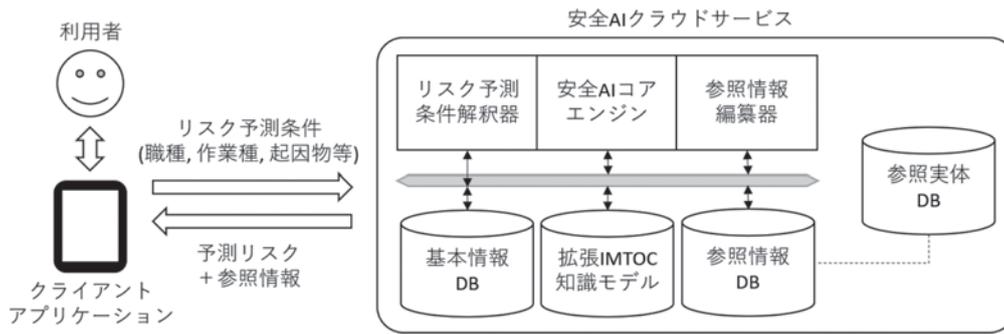
狭義の安全 AI システムは,労働災害データの分析結果を知識源とし,拡張 IMTOC モデルを基礎とするモデルベース AI システムである。広義の安全 AI システムは,クラウドサービスであり,クライアントアプリケーションを包含する。当該システムの全体構成

$$Score = b + \sum_{i=1}^n (a_i x_i + c_i), \quad b: \text{base value}, a_i \text{ and } c_i: \text{constant}, x_i: \text{value of an attribute}$$

式-1 任意の「事故の型」における予測リスクのスコア計算式



図-6 ダイナミック・チェックリストの導出例



図一七 安全 AI システム構成



図一八 安全 AI クライアントのリスク予測実施例

を図一七に示す。

リスク予測条件解釈器は、利用者からのリスク予測条件に対しリスク予測に適した内部情報を生成して安全 AI コアエンジンに引き渡す。安全 AI コアエンジンは、拡張 IMTOC モデルを探索して予測されるリストを抽出する。この際、条件に応じた動的なスコアリングがなされる。最後に、参照情報編纂器が災害報告書実体へのポインターや参考画像などの情報を紐づけた後に、得られた注意喚起情報群をユーザーに返す。

本システムの前記推論プロセスの実施例を図一八に示す。本例は、KY 活動時に、関係者が KY 活動表を共有してリスクアセスメントを行う状況を表している。画面右側に、「職種」「作業種」「起因物」を所与の条件として、導出された予測リスクが列挙されるとともに、「対策」及び関連情報が表示されている。右上の「転記」ボタンを押下することで、対策候補が画面左側の「KY 活動表」に転記されている。「対策」がブランクの箇所は、KY 活動参加者が記載する。つまり、システムが提示する「対策」はあくまでも示唆のレベルであり、「対策」の主体は人の思考であることを意味する。

5. おわりに

労働災害データに対して、データ科学の手法を用いて分析を行った。その分析結果を基に労働災害知識のモデル構造化を図るとともに、ダイナミック・チェックリスト機構を実現することで、予測されるリスクを適応的・網羅的に提示する安全 AI システムを構築した。本システムは、クラウドサービスに包摂され、KY 活動に係る帳票等を提示するクライアント (eYACHO¹⁰) を通じて、予測されるリスク及び参照情報類を利用者に提示する。

また、本システムは、顧客企業のデータに立脚するため、安全 AI ソリューションとして個々に提供され、データ分析を嚆矢として個々のデータ特性に応じてチューニングされる。現在、複数の企業において、PoC (概念実証) が進められており、全社導入に移行したケースもある。

今後の課題として、生成 AI との連携が挙げられる。本システムの知識 AI は、決定論的な特性を有することで、確率論的な生成 AI と併せて、相補的にシナジーを発揮できることが期待される。また、フロントエンドに

認識 AI を設け、リアルな状況を知識 AI に伝達して潜在リスクの検知を行ったり、画像生成 AI により注意喚起のイメージ化を行ったりする等のマルチモード的發展も期待される。

最後に、実データの提供を受けた(株)大林組の関係者各位、幅広く安全衛生に関する知見を授けて頂いた労働安全衛生総合研究所の専門家各位、本研究を支援して頂いた(株) MetaMoji の全ての関係者に、深く感謝します。

J|C|M|A

《参考文献》

- 1) 梅崎重夫, 他: 厚生労働省科学研究費補助金 労働安全衛生総合研究事業「機械設備に係る簡易リスクアセスメント手法の開発に関する調査研究」平成 28 ~ 30 年度 総合研究報告書, 平成 31 年 (2019 年) 3 月
- 2) R ウェブページ: <https://cran.r-project.org/>
- 3) arules パッケージ: <https://cran.r-project.org/web/packages/arules/index.html>
- 4) CaboCha ウェブページ: <https://taku910.github.io/cabocha/>
- 5) 高藤淳, 來村徳信, 溝口理一郎: 人工物の機能概念・方式知識編纂技術と設計支援への応用, 日本機械学会 第 18 回設計工学・システム部門講演会, 2008 年 9 月 25 日
- 6) svd パッケージ: <https://cran.r-project.org/web/packages/svd/index.html>
- 7) 小宮山宏: 知識の構造化, オープンナレッジ, 2004 年 12 月 24 日
- 8) 齋藤雄志: 知識の構造化と知の戦略, 専修大学出版会, 2005 年 8 月 10 日
- 9) Deepak K. Agarwal, Bee-Chung Chen, 島田直希 (訳): 「推薦システム: 統計的機械学習の理論と実践」, 共立出版, 2018 年 4 月 21 日
- 10) eYACHO ウェブページ: <https://product.metamoji.com/gemba/eyacho/>

【筆者紹介】

高藤 淳 (たかふじ すなお)
(株) MetaMoji
セマンティック技術開発部
ディレクター



鈴木 理史 (すずき まさし)
(株)大林組
技術本部本部長室
担当課長

