

機械学習応用システムの要求工学に関する一考察: 要求 獲得上の課題特定のセンスメイキング論的アプローチ

中木 裕章 岸本 眞一郎 吉岡 信和

近年の ICT の高速化・大容量化とともに、機械学習の理論構築と技術開発に関する研究が精力的に報告されている。それに伴い、機械学習応用システム設計開発の需要と供給は近年大幅な増加の傾向にある。他方、従来の機械学習研究のパースペクティブは主に理学指向であった。近年の需給増加の時代においては工学的指向からシステム設計開発手法の体系的な分析と構築が急務である。本研究では、機械学習の要求獲得のためにユーザー、要求エンジニア、データサイエンティスト、ソフトウェアエンジニアの依存関係を文献調査により明らかにする。

With the recent acceleration and capacity-enlargement of ICT, science and technology of machine learning have been reported vigorously. As a result, the demand for and supply of machine learning system design and development in industry have been increasing significantly in recent years. On the other hand, the previous studies in the area has been carried out from the point of view of academic or basic-scientific perspectives. In the recent era of increasing supplies and demands, systematically analyze and construct system design and development methodologies from some engineering perspectives are required. In this study, a literature survey on the dependency among users, requirements engineers, data scientists, and software engineers is demonstrated for identifying the requirements acquisitions for design and developments of machine learning systems.

1 はじめに

20 世紀前半に端を発した人工知能の理論構築と技術開発は、はじめは数学者、論理学者、分析哲学者らによる純粋理学的な探究としてスタートした(第 1 次 AI ブーム)。20 世紀中頃から後半にかけての第 2 次 AI ブームでは、知識表現やセマンティックウェブに関する基礎研究、その応用としてのエキスパートシステムの設計開発が盛んにおこなわれた[3]。そして近年、いわゆる第 3 次 AI ブームと呼ばれる昨今においては、

近年の ICT の高速化・大容量化に下支えされつつ、機械学習理論にもとづく帰納的推論アルゴリズムの基礎研究が大きく進歩を遂げた[29]。たとえば第 3 次 AI ブームでは、これまでの技術理論や応用の革新性だけではなく、社会との関係性といった安全性・信頼性等の社会的要請を充足するための研究開発にも取り組みが広がっている[27]。

機械学習を応用した IT システム(以降「機械学習応用システム」と呼称)は、可読性の高い開発言語である R や Python の普及による技術の民主化を伴い、その需要と供給は近年大幅に増加するとみられる[26]。他方、従来の機械学習研究のパースペクティブは主に理学指向であった[23]。それゆえ近年の需給増大に向かう時代においては、工学的アプローチから課題を分析し、システム設計開発手法の構築を図る試みが急務といえる。

実学としてのソフトウェア工学は、現代社会において不可欠の経営基盤である情報技術を取り扱う性質上、その射程に経営課題の発見と解決を(少なくとも

Consideration on Requirements Engineering for Machine-Learning Systems: Sensemaking Approach for Identifying Issues of Requirements Acquisition.

Hiroaki Nakaki, Shinichiro Kishimoto, 北陸先端科学技術大学院大学 先端科学技術研究科 知識科学系, Dept. of Graduate School of Advanced Science and Technology, Science and Technology, Japan Advanced Institute.

Nobukazu Yoshioka, 早稲田大学 理工学術院総合研究所, Waseda Research Institute for Science and Engineering.

暗黙のうちに) 含んでいる。事実, たとえばわが国の AI 白書 (2020) では, 企業の AI 導入の動機となる課題意識について「技術課題としてではなく, 経営課題として捉える」必要が主張されている [8]。一方で, 機械学習を含んだ AI 導入に取り組んでいる企業はわずか 20 % に留まっているとする報告もある [5]。わが国の産業社会における AI 普及の現状からみて, 経営課題を意識した AI 導入の達成には越えるべき障壁が存在していることが示唆される。以降, この「越えるべき障壁」のひとつとして, われわれは要求工学上の問題に着目する。

1.1 機械学習要応用システムの要求工学

要求工学とは, ソフトウェアが持つべき機能や性能などの特性を定義したり管理したりするための技術や技法を意味する。ソフトウェア開発で避けるべきことは, 「ユーザーのニーズを満たしていない機能を作ること」である [6]。しかしながら, 「多くの企業は AI プロジェクトへの理解と期待が曖昧で, かつビジネス価値を実現する方法をよく知らないままさまざまな AI モデルを開発しているため, ビジネス目標を定義し, AI 開発をどの様に行うかを理解することが重要」である [18]。ユーザーのニーズを満たした機能を作るために, 要求獲得や定義の重要性が増している。

1.2 課題

機械学習応用システムではニーズを満たした機能を提供するための要求獲得や定義において, 「実現可能な要求の規定が難しい」 [1] という課題がある。この課題には 3 つの要因があると考えられる。すなわち, ステークホルダ間の「知識範囲の違い」, ビッグデータからの「知見の抽出」, 機能の予測不可能性と実装の不確実性による「機械学習の機能課題」である。

「知識範囲の違い」は, たとえば委託側と受託側で次のようなものがある。ユーザーは課題解決を考える際に, 機械学習を活用した応用問題に直接的に対応した解決目標を考える傾向がある。すなわち, 機械学習を応用問題に適用可能な範囲について考える。一方, 受託側は機械学習の工学的に定義された問題設定を行う傾向がある。すなわち, 機械学習の技術的に可能な範

囲で考える。実現可能な要求を規定するためには, 委託側と受託側のステークホルダ間との対話によるニーズとシーズのすり合わせが重要となる。しかしながら, ニーズとシーズのすり合わせ対話には, 「知識範囲の違い」による障害がある。すなわち, ユーザーは応用問題に対するニーズの知識やドメイン知識がある。一方, 受託側は技術実現性や機械学習といったシーズの知識がある [23]。ユーザーはこれらの知識と機械学習知識を元に課題解決の方法を提案し, 受託側はユーザーの問題について深く理解する努力が必要になる。

また, すり合わせの障害には, 知識範囲の違いに起因する「知見の抽出」の課題がある。要求抽出の際に, ビッグデータからビジネスに活用できる知見を引き出すには, 要求エンジニアはデータ分析知識が不足しているため, データサイエンティストとの協業によりビッグデータからの「知見の抽出」が不可欠となるからである [1][23]。

また, すり合わせの障害には「機械学習の機能課題」もある。機械学習応用システムでは, 要求獲得が最も難しい作業の 1 つである [9]。多くの機械学習ソフト開発は, 機械学習では何が導出されるか事前に予測することが困難であるからである。特に, モデルを仮定する前に現実的な立場で, データの示唆する情報を多面的に捉える探索的データ解析の初期調査レベルから始まる案件に多く, 要求エンジニアやデータサイエンティストが果たすべき役割は多い。主な要因は, 「実装の不確実性と機能の予測不可能性であり, 性能 (精度), 開発期間やコスト, 費用対効果などを事前に保証できないことである」 [9]。

このように, ステークホルダ間のすり合わせには, 「知識範囲の違い」「知見の抽出」「機械学習の機能課題」の 3 つの障害があり, ステークホルダの協業は必須である。「知識範囲の違い」については, 機械学習ソフトウェアだけではなく, 従来のソフトウェアやビッグデータソフトウェアにも共通する課題であり, 多くの研究が行われている。またビッグデータからの「知見の抽出」については, ビッグデータ分野にて既に多くの研究があるが, ビッグデータを扱う機械学習応用システムにおいても同様の課題があり, 実際, ステークホルダの協業についての研究も行われている [9]。「機

機械学習の機能課題」は、従来のソフトウェア開発には無い、機械学習応用システムの独自の課題である。そのため、要求エンジニアとデータサイエンティストの協業による要求獲得能力の向上を図る [1]。など、研究が盛んになってきている。しかしながら、「機械学習の機能課題」を克服し要求獲得を行うために必須となるステークホルダの依存関係や協業のために必要な要件は未だ明らかになっていないため、「実現可能な要求を規定すること」は依然難しい。そこで本研究では、「機械学習の機能課題」を克服し、要求獲得に必要なステークホルダの依存関係について、文献調査により明らかにする。

本論文の構成は、6 節で構成する。2 節は背景と関連研究の記述、3 節は既存研究調査に採用した方法の紹介、4 節は調査の結果と分析、5 節はこれらの結果の議論、最終の 6 節で本研究の結論について述べる。

2 関連研究

2.1 従来のソフトウェア開発における要求工学研究

従来のソフトウェアの要求獲得は、「正確な要求」を追求することであった。すなわち、「要求を定義することの第一義は、要求を漏れなく入手し、ソフトウェアエンジニアに正確に伝えること」であり、また「これまでにも多くの要求工学関係の技術が提案されてきたが、それらの多くが要求を正確に定義するための技術」であった [28]。また、従来のソフトウェア工学では、継続的実験やリスタートアップなどのアプローチにより、開発中の製品が市場や顧客のニーズと適合するか仮説を立て検証することで、「作業の無駄を省く」 [6] という正確な要求を追求する手法についても研究が行われてきた。そして「正確な要求」の追求は、従来のソフトウェアだけでなく、機械学習応用システムにおいても同様に研究が行われている [18]。

2.2 機械学習工学の要求工学における課題

知見の抽出：ビッグデータの要求工学プロセスにおけるステークホルダの協業について、近年、研究が行われている。例えば、Nguyen-Duc A らは”BIG DATA REQUIREMENT ENGINEERING PROCESS” [17] と

して、次のような要求獲得プロセスの提案を行っている Requirements elicitation, Data acquisition (New Process), Data analysis and value discovery (New Process), Use cases consolidation (New Process), Requirement negotiation, Requirement modelling, Requirement validation, Requirement Specification.

機械学習の機能課題：(1) 要求エンジニア・データサイエンティストの役割：Vogelsang によると、要求エンジニアとデータサイエンティストの協業は、ユーザーの要求獲得において不可欠である [23]。同研究では、要求エンジニアは (a) 機械学習の性能指標を理解する必要があり、(b) 説明可能性、差別のないこと、および特定の法的要求などの新しい品質要求を意識する必要があり、(c) 要求工学プロセスに機械学習の特殊性を統合する必要があることを提案している。(2) ステークホルダの協業：AI エンジニアの業務は、ユーザーのニーズや、AI ソフトウェアのビジネス面やドメイン知識をつなげるため、ユーザーと自社の間で反復的かつ継続的な要求獲得アプローチを行うことを提案している [17]。また、AI 開発のためには各ステークホルダの協業は必須であり、AI 開発の実態について 7 社のインタビュー調査研究が行われている [18]。しかしながら、これらの研究においてステークホルダの相互依存関係については、明らかになっていない。

従って、機械学習応用システムでは、「技術的に可能な範囲」と「応用問題に適用可能な範囲」のすり合わせのため、要求エンジニア、データサイエンティストだけではなく、ユーザーやソフトウェアエンジニアを含めたステークホルダの協業は必須である。しかしながら、協業に必要な要件は明らかになっていないため、実現可能性のある要求を規定することが難しい。そこで本研究では、機械学習の要求獲得に必要なステークホルダの依存関係について、文献調査により課題を明らかにする。

3 既存研究調査

3.1 スコープ

機械学習応用システムの要求獲得においては、要求エンジニアによるユーザーからの要求獲得だけでは

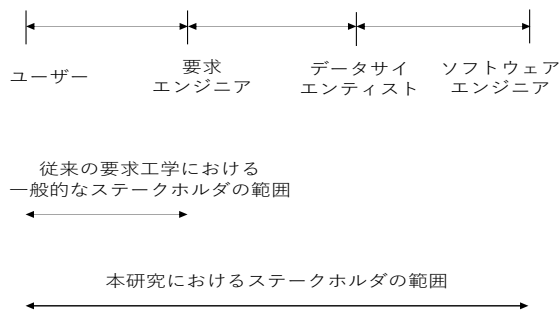


図1 ステークホルダの対象範囲

なく、元来それ自体では意味を持たないデータの意味づけやユーザーに対してビッグデータから得たインサイトや機能アイデアについて、ユーザーのニーズとの妥当性確認が必要となる。また、要求エンジニア、データサイエンティストやソフトウェアエンジニアなどの専門的知見を持った各ステークホルダの協業も必要となる。このため本調査では、ユーザー、要求エンジニア、データサイエンティスト、ソフトウェアエンジニアの各ステークホルダをスコープに含める。以下に本研究でスコープに含めるステークホルダの対象範囲を図1に記す。

3.2 リサーチクエスチョン

データで構成された機械学習の活用には、利点と難しさがある。機械学習を活用することの主な利点は、(a) ソフト開発プロセスやシミュレーションの高速化、(b) 大量で複雑なデータの可視化、(c) 新しい知見の発見がある。一方、機械学習を活用することの難しさは、(c) 何が導出されるか不明確といった機能の「予測不可能性」があり、これによってユーザーの期待やエンジニアの期待とシステム構築結果のギャップが生じる、(d) 訓練済みモデルを正しく評価できないといった実装の「不確実性」、(e) 知識やノウハウの再利用が難しい、といった点が考えられる。

しかし将来、ステークホルダが協業を行うための要求獲得プロセスや手法の研究が一層必要であり、研究も行われ始めている。だが、これらのステークホルダの協業に必要な要件は、未だ明らかになっていな

い。そこで次のリサーチクエスチョン (RQ) を設定した。

RQ: 機械学習応用システムの要求工学において、要求エンジニア、データサイエンティスト、ソフトウェアエンジニアが相互に連携し、要求獲得を行うために必要な要件は何か。

3.3 調査方法

本調査では、ソフトウェア工学分野の文献リサーチを対象とする” Systematic literature review process ” [2] を調査方法として採用する。調査は次の3つのフェーズで構成される。フェーズ1: 計画レビュー、フェーズ2: 実行レビュー、フェーズ3: 書類レビュー。

まず、フェーズ1の計画レビューでは、(1) レビュープロトコルの設定 ((a) 解釈フレームワーク、(b) 解釈対象データセット、(c) 解釈プロセス)、及び(2) プロトコルの妥当性確認の方法について次のように定義する。フェーズ1: 計画 (1) レビュープロトコルの設定: (a) 解釈フレームワーク: 多くのソフトウェア開発プロジェクトが要求定義段階で直面している課題について、一般的な要求工学が取り扱うプロセスと活動概念を表1の通り整理し、文献調査における解釈フレームワークとした。尚、解釈フレームワークの作成にあたっては、software engineering, 9th edition [22] を参考にプロセス分類を行い、さらに詳細な検討を行うため各プロセスに含まれる概念に分けた。

(b) 解釈データセット: 既存研究の文献抽出は、網羅的に文献が収録され、かつグローバルで広く活用されている Elsevier 社の Scopus システムより行う。また、次の検索キーワードを用いて同システムより文献検索を行う。検索キーワード: requirement engineering, machine learning, big data, AI. (c) 解釈プロセス: 抽出した文献データは、以下の手順で解釈を行い分類する。1. 各文献の抄録より、解釈フレームワークの定義に基づいてどのプロセスについて行われた研究かを内容より判定する。2. 次に、抄録の内容より各プロセスのどの概要に相当するかについて、解釈フレームワークに記載した定義より判定する。3. 上記1及び2項の判定の結果を解釈フレームワークに基づいて一覧化する。(2) プロトコルの妥当性確認レビュープロ

プロセス	概念	
要求獲得	対象領域が抱えている問題の理解 ・現状認識 ・問題理解(分析)	ビジネス活動を分析者が正確に理解するための作業 個々の当事者が抱えている問題を分析し特定する作業
要求分析	問題を解決するための方法の発見 ・問題定義 ・問題解決 ・方法策定	解決すべき問題の定義 問題の解決法の決定 システムによる問題解決の実現方法の策定 要求分析技術*を含む
要求仕様化	ソフトウェアによる実現方法の定義 ・モデリング ・要求記述	要求をシステム設計者に正確に伝える作業
要求確認	・妥当性確認	仕様化された要求が適切な要求であることを保証する
要求管理	・要求変更時の要求追跡 ・要求変更記録	要求の再利用

*要求分析技術とは、自然言語処理、オントロジーなどの自然言語としての仕様書を扱う技術と、顧客との交渉、要求の優先順位付けや要求選択などの合意形成に関するコミュニケーション技術、リスク管理や影響波及分析などのマネジメント技術を指す。

表 1 ステークホルダの範囲

トコルの妥当性を確認するため、上述した調査方法について第 3 者によるチェックを実施する。

フェーズ 2：実行レビュー
フェーズ 2 の実行レビューは、次の手順で作業を行う。
a) 関連研究の特定、b) 主要研究の選定、c) 調査品質の評価、d) データ抽出、e) データ統合

フェーズ 3：フェーズ 3 の書類レビューは、次の手順で作業を行う。
a) 分類・集計結果、b) 分類・集計が上述したフェーズ 1 及びフェーズ 2 で規定した手順通り行われているかについて、第 3 者による確認を行う。

上述した分類基準により調査を行うにあたり、ステークホルダの協業を行うための分類仮説を設定した。「機械学習が分かる人が、要求獲得プロセスに入り要求獲得の作業を行う」などの協業や要求獲得ノウハウの蓄積を必要とする記述があるかどうかについて確認を行う。

4 結果と考察

4.1 分析結果

上述した調査方法に基づいて作業を行った結果、Elsevier 社の Scopus システムより 194 件抽出され、その他 Scopus 最で公開されていない新の

文献を 40 件加え、調査対象は合計 234 件となった (<https://docs.google.com/spreadsheets/d/1R7So2BtLTcZ3y1Fd6/edit?us=sharingouid=100625203759564252109rtpof=truesd=true>)。そのうち 15 件が初期仮説に該当し、内訳は以下の通りであった。
(内訳)

(1) 機械学習エンジニアが、要求獲得作業に参画する必要がある:[該当数:4件]。・”A Multiple Case Study of Artificial Intelligent System Development in Industry” [18] ・”A Requirement Engineering Model for Big Data Software” [1] ・”Continuous experimentation on artificial intelligence software: A research agenda” [17] ・”An MDE Method for Improving Deep Learning Dataset Requirements Engineering using Alloy and UML” [25]

(2) 要求獲得を行う者が、機械学習エンジニアリングについて理解する必要がある:[該当数:1件]。・”Requirements engineering for machine learning: Perspectives from data scientists” [23]

(3) 機械学習だけではなく、形式知が必要である(要求獲得に関する工夫やノウハウの提案) :[該当数:10件]。

(3) の内訳:(a) モデリングフレームワークを用いて要求獲得を行うべき:[該当数:7件]。・”Business-driven data analytics: A conceptual modeling framework” [13] ・”Conceptual modeling for business analytics: A framework and potential benefits” [12] ・”Towards accountability for machine learning datasets: Practices from software engineering and infrastructure” [7] ・”Modeling machine learning requirements from three perspectives: a case report from the healthcare domain” [16] ・”A Big Data Conceptual Model to Improve Quality of Business Analytics” [19] ・”From Business Goals to Analytics and Machine Learning Solutions: A Conceptual Modeling Framework” [14] ・”Engineering the Black-Box Meta Model of Data Exploration” [25]

(b) リファレンスモデルを用いて要求獲得を行うべき:[該当数:1件]。・”A Reference Model for Big Data Analytics” [20]

(c) ソフトウェア開発の各フェーズにタスクセット

を設定し活用すべき: [該当数: 1 件]. ・”Emerging and changing tasks in the development process for machine learning systems” [10]

(d) デザインパターンを作成し、機械学習を他のケースに転用しやすくする: [該当数: 1 件]. ・”Identifying and Analyzing Data Model Requirements and Technology Potentials of Machine Learning Systems in the Manufacturing Industry of the Future” [21]

上記の分類結果より、次の考察を行った。

(1) 「機械学習が分かる人」が要件獲得に参画するだけでなく、要求獲得を行う者が機械学習エンジニアリングについて理解する必要がある。

(2) 機械学習応用システムにおける要求獲得に関する「手法」(フレームワーク)を活用して要求獲得を支援する方法もある。

(3) 要求獲得者と機械学習エンジニア相互の要求獲得に関する「手法」に関する研究は、今回の調査では確認できなかった。

4.2 主体と客体についての考察

上述した結果について、異なるステークホルダが協業すること、及び要求獲得のための手法は、要求獲得の正確さや効率化のために必要な方法と考える。異なる専門知識を持つステークホルダが多視点で要求獲得に参画することで、網羅的に検討できると思われるからである。だが、実装の不確実性や機能の予測不可能性のより高い機械学習応用システムでは、要求の見方や認識は、ユーザーを含めたステークホルダの相互依存関係の上で成立するという面もあると思われる。[従来] 万物には共有の真理があるとする考え方が主流であった。例えば、「自分がいま直面しているビジネス環境は、周囲の誰にも同じように見える。従って、事業環境を正確に分析すれば、普遍的な真実・真理が得られる」とする、実体的なパースペクティブに基づく考え方である。

[現状] IoTの普及に伴い、カスタマージャーニーなど、モノの売り切りモデルから売った後もユーザー行動を把握してサービスを売る等、循環的なビジネスモデルが出現し始めている。自社は、ユーザーと切り離せず、むしろユーザーや更にその外部にあるデジタル

環境全てに囲まれた一部とも言える。今後、このような考え方やビジネスモデルが増えていくことが推察される。

[将来] 「絶対的なビジネス環境の真理や要求」の探索ではなく、「要求の見方や認識は、ユーザーと自社、および自社内の複数のステークホルダーの相互依存関係の上で成立する」という視点による研究が必要ではないか。しかしながら、機械学習応用システムの要求工学で求められるユーザーや要求エンジニア、データサイエンティスト、ソフトウェアエンジニアのステークホルダの相互依存関係或いは協業の要件については、未だ明らかになっていない。

4.3 センスメイキング理論 (理論と概要)

そこで、相互依存的(関係論的)な視点を前提として、主体(自組織の各ステークホルダー)と客体(ユーザーおよび周囲の環境)の関連性について、ダイナミックに循環するプロセスとして捉えるマイクロ組織心理学の分野におけるセンスメイキング論[24]に着目した。プロセスは主に予測や期待していないことについて(1) 感知/Scanning (情報を感知する)した後、(2) 解釈・意味づけ/Interpretation し、(3) 行動・行為/Enactment (行動して環境に働きかける)するプロセスである。

解釈・意味づけのプロセスでは、異なるステークホルダ(多視点)によって捕らえられた要求の認識をすり合わせていくことが肝要となる。すなわち、同じ環境でも感知されたユーザーおよびユーザーを取り巻く周囲の環境をどう解釈するかによって、意味合いはステークホルダによって異なる。特に、市場環境の変化などにより先が見通しにくい状況や、前例が無いコンテキストなどにおいては、要求の「正解」を見つけることは不可能である。そこで多義的な解釈を収束させるプロセスは非常に重要となる。つまり、多様な解釈から特定の要求を選別し、意味づけを行い、異なるステークホルダが納得(sensemaking)し、ユーザーを含めたステークホルダでの要求の解釈の方向性を決めることである。センスメイキング理論の提唱者である Weick によれば、「納得性」(plausibility)について次のように見解を述べている。

・ Divers as these situations may seem, efforts are made to construct a plausible sense of what is happening, and this sense of plausibility normalization the breach, restores the expectation, and enables projects to continue.(Weick et al.,2005,p419)

・ the concept of sensemaking suggestions that plausibility rather than accuracy is the ongoing standard that guides learning.(Weick et al.,2005,p419)

そして、主体は客体と分離できないため、組織は行動して客体（あるいは環境）に働きかけて環境への認識を変える。また、環境に働きかけることで、新しい要求を感知する。このように行動・行為によって、認識された環境への解釈について納得性を高めることができる。また、Weick は論文の中でセンスメイキングの特性について表2の通り7つの要素を説明している。

No.	特性	概要
1	アイデンティティ (Identity)	常に環境の中での立ち位置を知る
2	振り返り (Retrospect)	過去のデータの見直しや事後的な振り返りによってセンスメイキングできる
3	行動・行為 (Enactive of environments)	行動することで環境に働きかけることができる
4	社会性 (Social context)	主体と客体は切り離せず、常に客体（他者）との関連性の中で起きる
5	継続性 (Ongoing)	センスメイキングのプロセスは、繰り返される循環プロセスである
6	環境情報の部分的感知 (Salient cues)	認識・解釈されたものは常に全体の一部ではない
7	最もらしさ (Plausibility)	「正確性」ではなく、「説得性」を持って主体（自組織）や客体（ユーザー）をセンスメイキングできる

表2 センスメイキングの7つの特性

メンバーが主観的に事象を見つても、変化する状況の中で共同化された主観がアップデートされながら組織として整合性の取れた振る舞いや原因の所在を想定することで、事象を客観視する他の説明的プロセスとは大きく異なる。

一般に、要求分析で重視されるのは、客観的な情報とそれに基づいた正確な分析である。しかしこれは「正解」があることが前提となる。一方、市場環境が大き

く変化し、経験や前例の無い状況状況や、見通しの難しい世界で、組織がどのように柔軟に意思決定し、新しいものを生み出していけるかに、センスメイキング理論は大きな示唆を与えている。

センスメイキングの5段階プロセス：センスメイキング理論は、組織研究を行っていた Weick によって唱えられ、後に主にはマイクロ組織心理学の分野に影響を及ぼした。Madsbjerg らは、センスメイキング理論を要求獲得や分析に適用し、同理論の実践について5つのプロセスを提案した[11]。(1)問題を捉え直す(問題を人の体験として見る。顧客視点で見る)。(2)データを集める(仮説検証や正確さの追求ではなく、先入観の無い方法で集める)。(3)データのパターンを探す。(4)鍵となるインサイトを生み出す(データから真実を発見する)。(5)事業にインパクトをもたらす。

4.4 センスメイキング論にもとづく解釈フレームワークの提案

本研究では以降、Madsbjerg のセンスメイキング論の視点からみた要求工学のあるべき姿についての考察を示す。この考察のためには、センスメイキング論に従って機械学習応用システムの要求工学のあり方を分析・評価するための解釈フレームワークが必要である。ゆえに、本研究ではこの必要とされる解釈フレームワークについても構成・提案する。

解釈フレームワークの構成は以下の手順によりおこなった。すなわち、Madsbjerg の5段階プロセスにおける各工程の目的に応じ、3-3 レビュープロトコルで示した機械学習応用システムの要求工学のプロセスと概念を適用した。そして、センスメイキングの5段階の各プロセスに対して、機械学習応用システムの要求獲得に必要な表3に示した要求工学プロセスと概念を記述した。次に、センスメイキングの5段階の各プロセスに必要な要件を機械学習応用システムの要求獲得作業と比較し、対応する項目として「多視点」「双方向性」「持続性」「共創性」「戦略性」を列挙した。これらの考察により、上記5つの項目からなる解釈フレームワークを設計した。

解釈フレームワーク：センスメイキングの5つのプロセスより、機械学習応用システムの要求獲得には

プロセス / 概念	必要要件	内容
① 問題を捉え直す (問題を人の体験として見る。顧客視点で見る) 【要求獲得】 【現状認識】	多視点	要求エンジニア、データサイエンティスト、ソフトウェアエンジニアなど、専門知識や知見を組み合わせて問題を再定義することが必要。
② データを集める (先入観の無い方法で集める (X: 仮説検証、正確さ) 【要求獲得】 【現状認識】	双方向性	要求エンジニア、データサイエンティスト、ソフトウェアエンジニアなどが双方向のコミュニケーションに基づく協業が必要。
③ パターンを探す 【要求獲得】 【問題理解】	持続性	・ユーザー/顧客の要求を収集するだけでは埋もれている真相を見抜くことはできない観察によってパターンが浮かびあがることでその真相が見えてくる。 ・パターンを探す知見の蓄積が必要。
④ 鍵となるインサイトを生み出す (データから「真実」を発見する) 【要求獲得】 【問題理解】	共創性	要求エンジニア、データサイエンティスト、ソフトウェアエンジニアなどが共創し、データからインサイトを見出す。
⑤ 事業にインパクトをもたらす 【要求獲得】 【問題理解】	戦略性	見出した「インサイト」はイノベティブな戦略の立案につながる要素があることが必要。

表 3 5つの要件

「多視点」「双方向性」「持続性」「共創性」「戦略性」の5つの要件が必要である。なかでも、データ自体には元来意味がないため「意味づけや解釈」は必須である点から、「多視点」と「共創性」は優先度の高い要件である。上述内容の一覧を表3に示す。

既存研究の分類4-1で示した初期仮説に該当した既存研究15件について、本解釈フレームワークに基づいて分類し整理した。分類については、次の手順で行う。

分類(1): 本解釈フレームワークの5つの各要件について、各研究における要求の捉え方または認識が実体的か、あるいは関係論的かによって分類する。本論文にて述べる「実体的」および「関係論的」とは、次の考え方を指す。

「実体的」: 要求には絶対的な真実や真理があるとする立場で、主体は客体を正確に観察・分析することにより、その真実や真理を知りステークホルダと共有できるとする考え方。

「関係論的」: 要求の見方や認識は、主体と客体の関係の上で成立すると考える立場。主体と客体は互いに影響・依存し合い、かつ要求はステークホルダの捉え方や認識によって異なるため、要求は認識主体の感じるコンテキストを通じてのみ認識されるとする。

分類(2): 分類(1)の結果、ステークホルダの協業

が必須となる「多視点」および「共創性」に該当した文献について、文献内容を基に「分析のみ」、「仮説抽出」に留まっているもの、「解決策の提案」まで行っているもの、および「実体的」と「関係論的」に分類する。

4.5 分類結果

分類(1)の結果は次の通りであった。

・「多視点」に該当した研究は5件あり、うち4件が「実体的」、1件が「関係論的」であった。「実体的」と判定された4件のうち、「現状認識」に関するものが3件あり、AI開発プロセス[18]、データサイエンティストとソフトウェアエンジニアの協業モデル[1]、要求エンジニアに必要な条件[23]であった。「問題理解」に関するものは1件あり、要求エンジニア・ソフトウェアエンジニアの協業によるデータ探索モデルの構築[25]についてであった。「相対的視点」1件の内訳は、ステークホルダによる反復・継続的な要求獲得する協業モデル[17]についてであった。・「双方向性」に該当した研究は2件あり、「実体的」および「問題理解」についてであった。内訳は、ステークホルダが協業しビジネス分析を行えるビッグデータ分析・解析フレームワーク[20][4]についてであった。・「持続性」に該当した研究は6件あり、全て「実体的」であった。内訳は、「現状認識」に関するもので、持続的な分析のためのフレームワークに関する内容3件[12][16][15]、要求変更管理1件[10]であった。また「問題理解」に関するものは、複数機能を維持するモデリングフレームワーク1件[13]、デザインパターン・知見の転用1件[21]であった。・「共創性」に該当した研究は1件あり、「実体的」および「現状認識」についてであった。内訳は、データサイエンティストとソフトウェアエンジニアの協業によるデータセット改善によるモデル駆動工学手法[25]についてであった。・「戦略性」に該当した研究は1件あり、「実体的」および「現状認識」についてであった。内容は、データ戦略として、データセット開発の透明性を維持するフレームワーク[7]についてであった。分類(2)の結果は次の通りであった。

・「多視点」および「関係論的」に該当した1件につ

要求の捉え方・認識

必要要件	実体的		関係論的	
	現状認識	問題理解	現状認識	問題理解
多視点	3	1	1	0
双方向性	0	2	0	0
持続性	4	2	0	0
共創性	1	0	0	0
戦略性	1	0	0	0

表 4 分類結果 (1)

いては、ビジネス目標または要求獲得にいて、産業用 AI ソフトウェアを開発するためにステークホルダが反復的に協業を行うことの必要性を主張しているが、分析のみに留まっている。

以上の分類の結果、関係論的な研究は 1 件のみであり、仮説抽出や解決策の提案は明らかになっていないことが分かった。分類 (1) の結果を表 4 に示す。

分類 (2) の結果は次の通りであった。

- ・「分析のみ」のものは 2 件あり、「実体的」「関係論的」各 1 件ずつであった。「実体的」によるものは、AI 開発用プロセス [18] についてであった。「関係論的」によるものは、ステークホルダによる反復・継続的な要求獲得する協業モデル [17] についてであった。
- ・「仮説抽出」まで行って入りものは 1 件あり、ステークホルダが協業しビジネス分析を行えるビッグデータ分収集・析フレームワーク [25] についてであった。
- ・「解決策の提案」まで行っているものは 3 件あり、要求の「現状認識」に関して、データサイエンティストとソフトウェアエンジニアの協業モデル [1]、要求エンジニアに必要な条件 [23]、データサイエンティストとソフトウェアエンジニアの協業によるデータセット改善によるモデル駆動工学手法 [25] についてであった。

要求の捉え方・認識

内容	実体的		関係論的	
	現状認識	問題理解	現状認識	問題理解
分析のみ	1	0	1	0
仮説抽出	0	1	0	0
解決策提案	3	0	0	0

表 5 分類結果 (2)

以上の分類の結果、「多視点」および「関係論的」に該当した 1 件については、ビジネス目標または要求獲得にいて、産業用 AI ソフトウェアを開発するためにステークホルダが反復的に協業を行うことの必要性を主張しているが、「分析のみ」に留まっており、「仮説抽出」や「解決策の提案」は明らかになっていないことが分かった。分類 (2) の結果を表 5 に示す。

5 本研究の可能性と限界、将来展望

機械学習応用システムの要求獲得の重要な目標は、「技術的に可能な範囲」と「応用問題に適用可能な範囲」について、ユーザーと対話できるようになることであり、実現のためにはステークホルダの協業は必須である。従って、ステークホルダの協業の要件を明らかにすることが、本研究の目的である。

研究の意義：従来のソフトウェア開発における要求工学においては、絶対的なビジネスの環境の真理や要求の探索を基本とし、事業環境を正確に分析すれば、普遍的な真実・真理が得られるとする「実体論」に基づいていた。実際、要求工学分野の多くの研究は、いかにユーザーの行動観察をベースに、「正確な要求」を追求するかに着目していた。だが、IoT 化の進展に伴い、ユーザーおよび自社のステークホルダは基本的に常時オンラインに接続された状態にある。このような

デジタル化されたビッグデータ環境における要求獲得の手法は、従来より行われてきた行動観察と相違点がある。たとえば、次のような点が考えられる。

[行動観察]

- ・サンプル数は限られていても、ユーザーのニーズに関する情報、利用文脈、ユーザーの意識、ユーザーの属性情報を入手できる。市場構造とニーズの見極めには行動観察がかなり優位。
- ・ポジショニングやサービス・スペックの設計には、行動観察（ユーザビリティ・テスト）で大筋の方向やオプションの広がりを含め、ビッグデータで微調整することを繰り返すのが一般的。
- ・取り組みの効果検証の場合、意識の変容を知るには行動観察が適している。
- ・長期的なトレンドの大局観を得るには行動観察が適している。

[ビッグデータ]

- ・全数を解析することで、頻度の少ないテール部分の情報も得られる。
- ・ビッグデータはリアルタイムの情報も得られる。
- ・ニーズ発生場面での対策の立案は、ビッグデータでは容易だが、行動観察では不可能。
- ・取り組みの効果検証の場合、利用行動の変容を知るにはビッグデータが適している。
- ・長期的なトレンドから見えてきたパターンの延長による近未来予測は、ビッグデータが適している。

従って、機械学習もビッグデータと同様に「機械学習をどう使うか」が課題となる。しかし、これまでの議論で見てきたように「機械学習の機能課題」、すなわち、何が導出されるか分からない予測不可能性と実装の「不確実性」がある。機械学習応用システムにおける要求獲得は、「機械学習をどう使うか」という視点ではなく、ユーザーや自組織のデジタル環境を前提にした、相対的な視点による要求の捉え方と手法が必要となる。従って本研究の意義は、機械学習応用システムの要求獲得において、関係論的な視点に基づいたステークホルダの協業の課題について提案した点にある。

今後の課題：本研究は既存の文献調査を基にステークホルダの協業要件について検討を行ってきたが、実務環境への適用可能性についての検討の余地がある。さらには、実務環境への適用を視野に入れ、今後は構成した解釈フレームワークの適用可能性や妥当性

についてのより深い検証と考察が必要と考えられる。

6 結論

機械学習応用システムの要求獲得は、従来のソフトウェア開発における要求獲得と著しく異なる。機械学習応用システムの要求獲得は、循環的な視点で要求を捉え、各ステークホルダの協業が必須となる。この協業において、センスメイキング理論に基づいて分析を行うことにより、ユーザー、要求エンジニア、データサイエンティスト、ソフトウェアエンジニアの各ステークホルダの協業に必要な要件が明らかになった。

7 謝辞

本研究は、北陸先端科学技術大学院大学 東京社会人コース 講義「機械学習工学」、および当該講義にて紹介された、文部科学省 平成 29 年度「成長分野を支える情報技術人材の育成拠点の形成 (enPiT)」enPiT-Pro 採択「スマートエスイー：スマートシステム&サービス技術の産学連携イノベティブ人材育成」を契機としてスタートいたしました。本研究のきっかけを下さった内平直志教授（北陸先端大知識、当該講義を企画・担当）に感謝いたします。

参考文献

- [1] Altarturi, Hamza Hussein, Ng, Keng-Yap, Ninggal, Mohd Izuan Hafez, Nazri, Azree Shahrel Ahmad, Ghani, and Abdul Azim Abd. A requirement engineering model for big data software. In *2017 IEEE Conference on Big Data and Analytics (ICBDA)*, pp. 111–117, 2017.
- [2] Pearl Brereton, Barbara A Kitchenham, David Budgen, Mark Turner, and Mohamed Khalil. Lessons from applying the systematic literature review process within the software engineering domain. *Journal of systems and software*, Vol. 80, No. 4, pp. 571–583, 2007.
- [3] B. G. Buchanan. A brief history of ai, 2021. <https://aitopics.org/misc/brief-history>.
- [4] Ranit Chatterjee, Abdul Ahmed, and Preethu Rose Anish. Identification and classification of architecturally significant functional requirements. In *2020 IEEE Seventh International Workshop on Artificial Intelligence for Requirements Engineering (AIRE)*, pp. 9–17, 2020.
- [5] Ben Dickson. Why most machine learning strategies fail, 2021. <https://bdtechtalks.com/2021/02/25/machine-learning-strategy-barriers/>.

- [6] Fabian Fagerholm, Alejandro Sanchez Guinea, Hanna Mäenpää, and Jürgen Münch. The right model for continuous experimentation. *Journal of Systems and Software*, Vol. 123, pp. 292–305, 2017.
- [7] Ben Hutchinson, Andrew Smart, Alex Hanna, Emily Denton, Christina Greer, Oddur Kjartansson, Parker Barnes, and Margaret Mitchell. Towards accountability for machine learning datasets: Practices from software engineering and infrastructure. FAccT '21, p. 560–575, New York, NY, USA, 2021. Association for Computing Machinery.
- [8] Japan Information-technology Promotion Agency. *Artificial Interigence White Paper*. p300 ebook.
- [9] Fuyuki Ishikawa and Yutaka Matsuno. Evidence-driven requirements engineering for uncertainty of machine learning-based systems. In *2020 IEEE 28th International Requirements Engineering Conference (RE)*, pp. 346–351. IEEE, 2020.
- [10] Hanyan Liu, Samuel Eksmo, Johan Risberg, and Regina Hebig. Emerging and changing tasks in the development process for machine learning systems. New York, NY, USA, 2020. Association for Computing Machinery.
- [11] Christian Madsbjerg and Mikkel B. Rasmussen. An anthropologist walks into a bar..., 2014. <https://hbr.org/2014/03/an-anthropologist-walks-into-a-bar>.
- [12] Soroosh Nalchigar and Eric Yu. Conceptual modeling for business analytics: A framework and potential benefits. In *2017 IEEE 19th Conference on Business Informatics (CBI)*, Vol. 01, pp. 369–378, 2017.
- [13] Soroosh Nalchigar and Eric Yu. Business-driven data analytics: A conceptual modeling framework. *Data Knowledge Engineering*, Vol. 117, , 04 2018.
- [14] Soroosh Nalchigar and Eric Yu. Dfrom business goals to analytics and machine learning solutions: A conceptual modeling framework, 2020. <https://tspace.library.utoronto.ca/handle/1807/103348>.
- [15] Soroosh Nalchigar and Eric Yu. From business goals to analytics and machine learning solutions: A conceptual modeling framework. 11 2020.
- [16] Soroosh Nalchigar, Eric Yu, and Karim Keshavjee. Modeling machine learning requirements from three perspectives: a case report from the health-care domain. *Requirements Engineering*, Vol. 26, pp. 1–18, 06 2021.
- [17] A. Nguyen-Duc and P. Abrahamsson. Continuous experimentation on artificial intelligence software: A research agenda. pp. 1513–1516, 2020.
- [18] A. Nguyen-Duc, I. Sundbø, E. Nascimento, T. Conte, I. Ahmed, and P. Abrahamsson. A multiple case study of artificial intelligent system development in industry. In *ACM International Conference Proceeding Series*, pp. 1–10, 2020.
- [19] Grace Park, Lawrence Chung, Haan Johng, Vijayan Sugumaran, Sooyong Park, Liping Zhao, and Sam Supakkul. A big data conceptual model to improve quality of business analytics. In Fabiano Dalpiaz, Jelena Zdravkovic, and Pericles Loucopoulos, editors, *Research Challenges in Information Science*, pp. 20–37, Cham, 2020. Springer International Publishing.
- [20] Grace Park, Vijayan Sugumaran, and Sooyong Park. A reference model for big data analytics. In *2018 9th IEEE Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)*, pp. 382–391, 2018.
- [21] Gunter Schuh, Paul Scholz, Thomas Leich, and Richard May. Identifying and analyzing data model requirements and technology potentials of machine learning systems in the manufacturing industry of the future. In *2020 61st International Scientific Conference on Information Technology and Management Science of Riga Technical University (ITMS)*, pp. 1–10, 2020.
- [22] Ian Sommerville. *Software Engineering (9th Edition)*. <http://index-of.co.uk/Engineering/Software>
- [23] A. Vogelsang and M. Borg. Requirements engineering for machine learning: Perspectives from data scientists. pp. 245–251, 2019.
- [24] Karl E Weick. *Sensemaking in organizations*, Vol. 3. Sage, 1995.
- [25] Robert Winter and Li Yang. Engineering the black-box meta model of data exploration. 11 2019.
- [26] 丸山宏ほか. 機械学習工学の必要性と現状. ウィンターワークショップ 2018・イン・宮島 論文集, Vol. 2018, pp. 14–15, 2018.
- [27] 国立研究開発法人科学技術振興機構研究開発戦略センター. 人工知能研究の新潮流, 2021. <https://www.jst.go.jp/crds/pdf/2021/RR/CRDS-FY2021-RR-01.pdf>.
- [28] 妻木俊彦. 要求工学: 現実と仮想をつなぐために. コンピュータ ソフトウェア, Vol. 29, No. 2, pp. 2.43–2.64, 2012.
- [29] 人工知能学界. 人工知能の歴史, 2021. <https://www.ai-gakkai.or.jp/whatsai/AIhistory.html>.