AIエージェントの理想と現実
期待される未来と技術の現在地を見つめ直す

国立情報学研究所 教授

佐藤一郎

自己紹介：佐藤一郎

- 国立情報学研究所・情報社会相関研究系・教授　/
  総合研究大学院大学・先端学術院・情報学コース・教授（併任）

- 学歴
  - 慶應義塾大学理工学部電気工学科卒、同大学理工学研究科大学院計算機科学専攻後期博士課程修了、博士(工学)

- 政府関連の委員など
  - デジタル庁「政策評価に関する有識者会議&行政事業レビュー」座長
  - 経産省「データ連携コミュニティとそのガバナンスに関する研究会」座長
  - 消費者庁「デジタル取引・特定商取引法等検討会」構成員
  - 内閣府知的財産戦略戦略本部「メタバース官民連携会議」構成員
  - 総務省「放送分野の視聴データの活用とプライバシー保護の在り方に関する検討会」構成員
  - OECD Research ethics working group, member
  - テレビ朝日系列番組「仮面ライダー（ゼロワン）」(2019年9月から2020年8月まで放映)のAI技術アドバイザー他

宣伝させてください

- 佐藤一郎著：「2030次世代AI」
- 2025年11月21日発売（日経BP）

講演概要

- 本講演では、近年注目を集めるAIエージェントについて、AIエージェントの活用の観点から、どのような効果が期待できるのかを具体例とともに示す。次に、AIエージェントに関わる品質を維持するために必要となるデータ整備を含む事前準備を論じる。さらにAIエージェント導入背景についても検討していく。

1．AIエージェントに関する典型的質問から
   - AIエージェントはどんな効果があるのか？（AIエージェントとは）
   - 自社にAIエージェントを導入するには？（AIエージェントの準備）
   - AIエージェントを導入しないといけないのか？（AIエージェント必然性）
2．まとめ

時間も限られていることもあり、話題を絞って講演します

AIエージェントに関わる典型質問

- AIエージェントに関して尋ねられること

  Q: AIエージェントはどんな効果があるのか？

  Q: 自社にAIエージェントを導入するには？

  Q: AIエージェントを導入しないといけないのか？

佐藤一郎（国立情報学研究所）

AIエージェントに関わる典型質問

- AIエージェントに関して尋ねられること

  Q: AIエージェントはどんな効果があるのか？

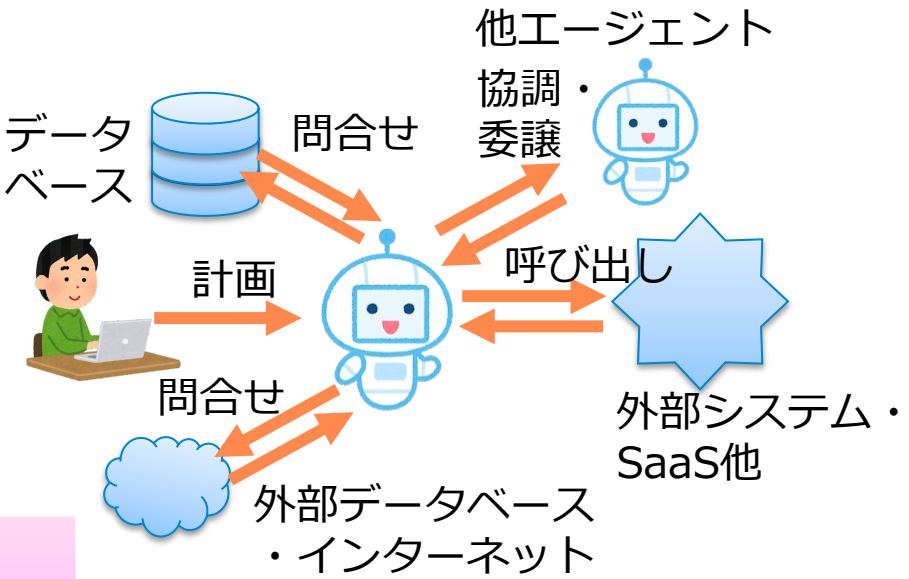
  Q: 自社にAIエージェントを導入するには？

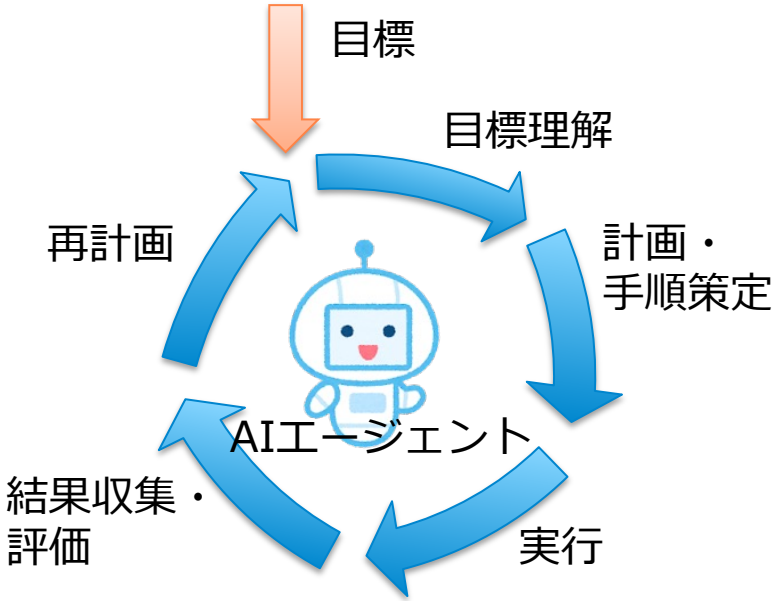
  Q: AIエージェントを導入しないといけないのか？

佐藤一郎（国立情報学研究所）

AIエージェントとは何か

- AIエージェントに定義はない
  - 各社が都合よく、 AIエージェントという用語を使っている状態
  - 強いて定義するならば
    - 人間が与えた目標に基づき、環境を認識し、計画・手順を立案し、計画・手順に従って外部システムを呼び出し・実行し、結果を評価・改善する自律的仕組み





学術研究におけるエージェント

- コンピュータサイエンス分野で、エージェントと呼ぶ対象そのものが研究対象となったのは1990年前後
  - Software agents
  - Intelligent agents
  - Multiple agents
  - Autonomous agents
  - Interface agents
  - Mobile agents
  - Belief-Desire-Intention (BDI) agents
  - ・・・・
- 2000年以降
  - Multiple agents分野が大きく伸びる
  - 機械学習を含む統計的学習メカニズムの導入

エージェント定義は混乱
〇〇エージェントと呼ぶことで収拾

AIエージェントの実状

- エージェント学術研究の実状
  - 1980年代から目標から計画策定(Planning)は主要研究テーマ
    - AIが一意に解釈可能な形式で目標が与えられ、計画の実行対象は少数の要素から構成された閉じた系
      - 論理的導出、探索問題、ヒューリスティック解法、強化学習他
    - 実環境において計画策定ができるとは言い難いのではないか

- AIエージェントでは
  - 目標は自然言語で与えられ、LLMを活かした言語的推論により計画策定
  - 汎用性は高いが、目標の正確な抽出は困難、検証が困難といえる
    - 少なくても短期的に汎用自律エージェントが完成する可能性は低く、当面は領域特化型・限定自律型に留まるのではないか

技術的限界からAIエージェントは局所的な自動実行手段に矮小化している状況

AIエージェントは日本企業に向くのか

- AIエージェントはトップダウン型で業務改革を行うための技術
  - 現場裁量が強く、ボトムアップ型改革を指向する日本企業に向くのか
    - 日本企業でAIエージェントを活かすには意図的にボトムアップ型とすべき

- AIエージェントの技術的＆組織的限界からトップダウン型改革は難しい
  - 目標理解と計画抽出は難しい
  - 多くの組織においてAI向けのデータ整備が不十分

- 技術的限界から、現状、AIエージェントによる全体的な計画策定は困難であり、局所的な計画策定に留まる
  - 局所的な自動化手段として矮小化されたAIエージェントは日本企業でも有用

佐藤一郎（国立情報学研究所）

AIエージェントと品質

- 言語生成AIの利用は汎用性を高めるが、言語生成AIに関わる品質問題を継承
  - 目標理解及び計画・手順策定の正確性
  - 一貫性・安定性の欠如
  - 説明可能性・透明性の制約
  - 業務適用性の制限（ドメイン・企業知見の欠如）
  - モデル更新・改善コストの増大

- AIエージェントの品質指標が整備されていない
  - 品質を測れないと改善は困難
    - AIエージェントの品質は対象とデータに依存するため、言語生成AIそのものより品質指標は難しい

- 当面は
  - AIエージェントは局所的な自動化に留めるとともに、人間による評価・制御は不可欠ではないか
  - AIエージェントの間違いを許容する分野から活用

AIエージェントの品質指標

- 目標理解・目的整合性：意図・経営目標・業務目的を正しく理解しているか
- 計画・推論能力：目標を実行可能な手順に落とせるか
- 出力の正確性・信頼性：誤情報や誤判断がどれだけ少ないか
- 一貫性・安定性：状況が同じなら同じ判断をするか
- 説明可能性・透明性：なぜその判断をしたか説明できるか
- 業務適合性・専門性：自社・業界の状況・制約に合っているか
- データ品質・知識基盤：学習・参照データの質
- セキュリティ・倫理性：安全・公平・法令順守
- 継続改善力・運用成熟度：長期的に改善し続けられる体制があるか

備考：AIエージェントは言語生成AIを利用するが、言語生成AIの品質指標がAIエージェントの品質指標の一部でしかない

佐藤一郎（国立情報学研究所）

AIエージェントに関わる典型質問

- AIエージェントに関して尋ねられること

  Q: AIエージェントはどんな効果があるのか？

  Q: 自社にAIエージェントを導入するには？

  Q: AIエージェントを導入しないといけないのか？

佐藤一郎（国立情報学研究所）

自社にAIエージェントを導入するには？

- AI-readyの組織構造、業務、データなのか

- 何のためにAIエージェントを導入するのか
  - AIエージェントはデジタル技術の組織の意思決定・業務統治の再設計
    - 既存業務の理解
    - 業務をどのように変えたいのか
    - 責任分解の明確化
    - 改善の能力

- データ整備などのAIエージェント導入の準備
  - AIエージェントの品質は参照するデータ次第
    - 部署別サイロ化の解消（部署横断的な検索可能性）
    - メタデータの整備
    - AIが参照するデータの選別
    - データの真正性と最新性

AIエージェントの品質はデータ次第

- AIエージェントの品質は、モデルだけでなく、企業内データの影響を受ける
  - データの正確性・信頼性
  - データの標準化（意味的な一貫性を含む）
  - データの網羅性・偏り
  - データの鮮度・更新性
  - データのメタ情報（取得経緯・文脈）
  - データガバナンス及び責任分解点

- 部署から最新かつ正確なデータを収集・変換・メタ情報付加を集めるのではなく、データの取得段階からAIエージェントに対応したデータとしない限りは円滑なAIエージェント利用は困難ではないか

データサイロ化の解消

- AIエージェントは部署横断的にデータを参照し、それに基づいて判断、業務システムを実行できることが前提
  - 発見可能性（Discoverability）：どこに何があるか分かる
    - 異なる部署やシステムごとにデータを管理
      - 例：独自DB、ERP、CRM、ファイルサーバ他
  - 整合可能性（Reconciliation）：文脈把握、同一概念を同一として扱える
    - データの文脈情報の欠如、用語の不一致
      - 例：「顧客」＝請求先／利用者／契約主体…が部署で異なる
  - 到達可能性（Accessibility）：データを適切な範囲で参照
    - データへのアクセス権限の相違や管理不足
      - 利用者識別子、アクセス権限設定が部署により相違

佐藤一郎（国立情報学研究所）

AIエージェントに関わる典型質問

- AIエージェントに関して尋ねられること

  Q: AIエージェントはどんな効果があるのか？

  Q: 自社にAIエージェントを導入するには？

  Q: AIエージェントを導入しないといけないのか？

佐藤一郎（国立情報学研究所）

AIエージェント対応

- 企業の多くはAIエージェントを一人称で考えがちですが
  - ライバルとなる企業、取引先、消費者もAIエージェントを活用
  - 大半の企業は、取引先や消費者のAIエージェントとの取引の都合上、自社にAIエージェントの導入を強いられるのではないか

今後、企業にとっての関心事　自社がAIエージェントを活かすこと　＜　取引先や消費者のAIエージェントに選んでもらうこと

- AIエージェントの技術的な課題があるにしても、取引先や消費者のAIエージェントに選んでもらうことが重要となる
  - 取引先や消費者のAIエージェントに対策のAIエージェント導入でも、データ整備や組織構造などの準備は必要

AIエージェントを介した取引(B2B)

- 企業の多くはAIエージェントを一人称で考えがちだが、取引先、消費者もAIエージェントを活用する

- 発注側企業がAIエージェントを利用する場合
  - AIエージェントが商品の価格・品質・納期から合理的に判断・取引を代行
    - 従来の担当者同士の関係性や信頼感に基づく取引は通用しない
  - 人間はAIエージェントの取引頻度や迅速性に追いつけないことから、受注側もAIエージェントによる自動処理は不可避となる
    - 取引そのものの高速化、取引単位の細粒度化、スポット市場化していく

取引先のAIエージェントに選ばれる企業となることが目的化

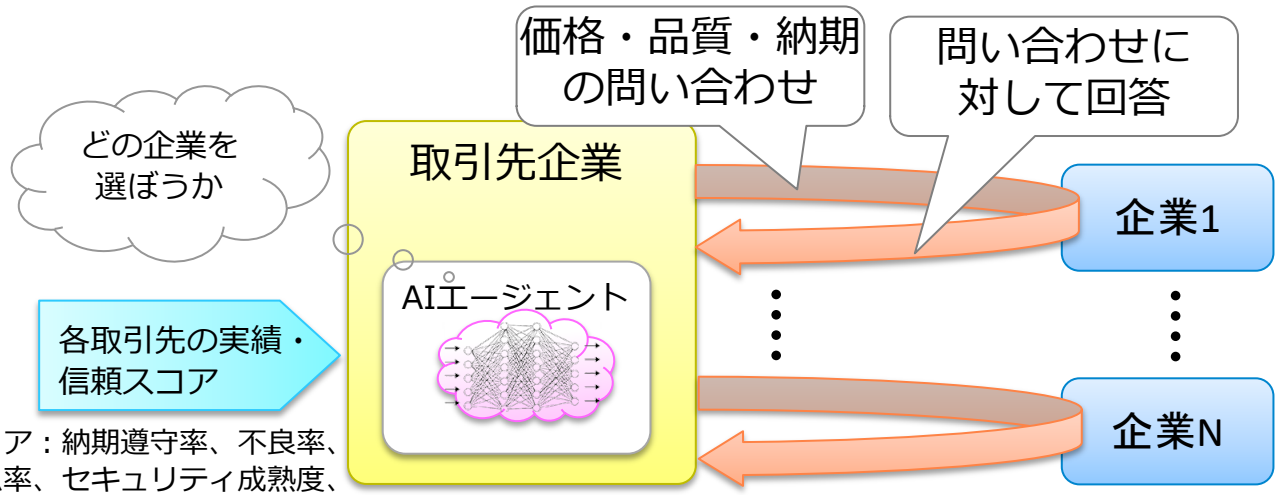
多くの企業にとってAIエージェントの導入理由は、取引先や消費者のAIエージェントへの対応するための自動化ツールが大半になるのではないか

佐藤一郎（国立情報学研究所）

企業に求められるAIエージェント対応

- AIエージェントを活かす発注側の流れ
  - エージェント要件定義 → 仕様・見積依頼 → 仕様・見積取得 → 比較 → 発注

- 受注側企業の関心事は、取引先のAIエージェントに選んでもらうことに変わる
  - 選んでもらう対策は、取引先のAIエージェントが求める情報を、AIが扱いやすい方法・形式で提供
    - ウェブ検索のSEO的な小手先対策が通じるとは限らない



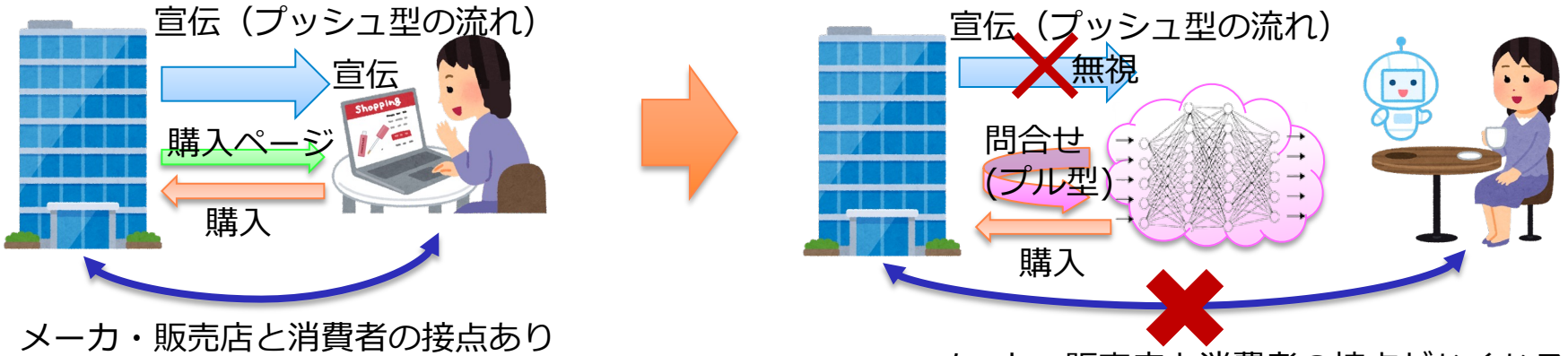
信頼スコア：納期遵守率、不良率、クレーム率、セキュリティ成熟度、ESG指標

AIエージェントの問い合わせに対して、想定外の回答を行う企業や、回答が遅い企業は生き残れないのではないか

- AIエージェントは問い合わせた回答以外の情報は処理してくれない

AIエージェントが変えるB2C

- 消費者は購入先の選定・購入をAIエージェントに任せる
  - 嗜好性が低い商品などは、商品の選択もAIエージェントに任せが増える
  - 商品選択・購入段階で消費者と企業（メーカ・販売店）との接点は消失
  - 人間からみたオンラインサイトの使い勝手は差別化要素ではなくなる
- 企業の関心事は(消費者の)AIエージェントに選んでもらうことに変わる
  - 選ばれるには、価格・品質・評判・納期に加えて、AIエージェントから問い合わせに対して、正確かつAIが扱い易い形式・方法で回答することになる



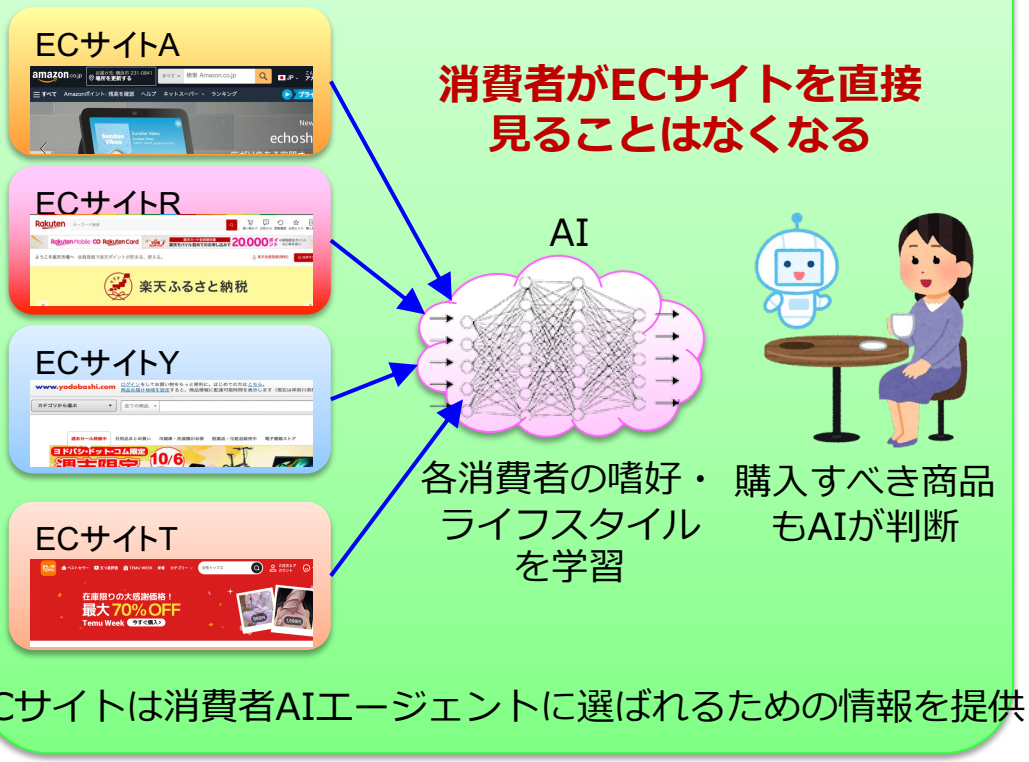
- AIエージェントが扱えるのは、プル型(問い合わせ型)情報だけ
  - AIエージェントは、企業からのプッシュ型情報を処理しきれないことから、プッシュ型情報（従来の宣伝）の効果は大きく下がる
  - AIエージェントの失敗を考えると、返品・返金可能事業者が優位になる

消費者によるAIエージェント利用

- 消費者もAIエージェントを利用する
  - AIエージェントが何を買うべきなのか、どこから買うべきかを決める

現在：各ECサイトに囲い込み
（使い勝手、慣れほか）

将来：利用者側AIエージェントが購入すべき商品を選び、
価格、配送、評判、納期からECサイトを選択・購入



ECサイトは消費者AIエージェントに選ばれるための情報を提供

- 大手ECサイトの優位性(使い勝手、消費者の使い慣れ、商品取り揃え他)は下がる
- 消費者は日常的に買う商品などは、商品選択もAIエージェントに任せになるはず

佐藤一郎（国立情報学研究所）

品質問題の回避

- AIエージェントの計画手順ミスの解決をAIエージェント以外に求める方法もありえる
  - 例：AIエージェントによる誤発注の損失を外部化
    - 発注先事業者、発注品、発注量、納入期限他を間違える可能性
      - 対策：発注先事業者による被害最小化

- ECサイトの場合、商品返品・返金を低コストで行う事業者もある

- 返品・返金コストが低い事業者が優位になる
  （なお、返品・返金コストが低い事業者は大手が多く、寡占化の恐れもある）

  - AIエージェントおよびその品質が業界構造を変える可能性もある

まとめ

- AIエージェントの効果
  - 理想は自然言語の目標から計画・手順を作成・実行し、結果を評価・改善する自律的仕組みとなるが、現状は局所的自動化手段に矮小化
    - その範囲でも有用性がある（特に日本では）

- 自社にAIエージェントの導入に向けた準備
  - AIエージェントはデジタル技術の組織の意思決定・業務統治の再設計
    - データ整備として部署別サイロ化の解消、メタデータの整備、AIが参照するデータの選別、データの真正性と最新性

- 取引先・消費者のAIエージェントへの対応のためのAIエージェント導入は不可避
  - AIエージェントから問合せ、取引に対応するためのAIエージェント導入
    - AIエージェントが取引先・消費者との関係性を大きく変える

AIエージェントは企業経営のあり方を問い直す戦略課題

佐藤一郎（国立情報学研究所）

もう一度、宣伝させてください

- 佐藤一郎著：「2030次世代AI」
- 2025年11月21日発売（日経BP）