

第20回図書館総合展 JMLA/JPLA フォーラム「図書館AIの最先端：AIの活用でリサーチが変わる」参加報告

村瀬由紀子*

朝日大学図書館

I. はじめに

2018年10月31日（水）に開催された第20回図書館総合展において、特定非営利活動法人日本医学図書館協会（以下、JMLAという）と日本薬学図書館協議会（以下、JPLAという）が共催するフォーラム「図書館AIの最先端：AIの活用でリサーチが変わる」が開催された。下記の講師の講演があった。

- ・村田輝氏（国立大学法人電気通信大学 学術国際部学術情報課課長）
- ・井手孝次郎氏（ユサコ株式会社アカデミア事業部プロダクト部プロダクトマネージャー）
- ・木戸邦彦氏（株式会社日立製作所 研究開発グループ 東京社会イノベーション協創センター 主任研究員）
- *司会進行：平紀子氏（JPLA理事）

講演に先立ち平氏より、英オックスフォード大学准教授マイケル・A・オズボーン氏の論文「雇用の未来」が全世界で話題になる等AIが取り沙汰される昨今、AIが導入されるとなごどのように変化するか皆様と考えたいとの提言があった。情報処理技術の進歩に伴う環境の変化を見据えた3つの講演内容について報告する。

II. 講演の内容

1. AI研究とのコラボによる次世代図書館の構築

講師：村田輝氏

電気通信大学において2017年4月に設置された新学修スペース「UEC Ambient Intelligence Agora」（以下、Agoraという）の事例報告があった。このスペースは、ディスカッションやプレゼンテーション等のアクティブラーニングを支援するとともに、学修環境に設置されたセンサーから得られたデータを人工知能や適応学習等の研究に活用することを目指している。

Agoraの設置は、アクティブラーニング空間の整備を求めていた図書館と、AI研究推進のためのビッグデータを必要としていた人工知能研究者双方の目指す方向が合致し、推し進められた。Agoraの収容人数は約270人で、対話型ロボット、ガラス製ホワイトボード、机上に映し出せるプロジェクタ、人間工学に基づいた機能性と洗練されたデザイン性を持った机・椅子等を設置している。学生の人気スポットとなっており、図書館利用者は改修前と比較して約15%の増となった。

Agoraには、温湿度・照度・人感・CO₂濃度を測定できるセンサーや、センサー機能を持ったCCDカメラ、指向性マイクが大量に設置され、18台のディープラーニングマシンにより取得データを解析できるようになっている。AIの解析結果からCO₂濃度の変化に応じて自動的に換気したり、照明の調光・調色を行ったり、AIにより自律的に制御が行われるアンビエント学修環境を作り出すことを目標としている。

また、もう一つの特徴として、対話型ロボット10台の導入が挙げられる。AIと人間が違和感なく共存していくためには、親しみやすい形や動作をもたせることで、人間が安心してコミュニケーションを取ることがで

*Yukiko MURASE：〒501-0296 岐阜県瑞穂市穂積1851-1.
hyukiko@alice.asahi-u.ac.jp (2019年2月1日 受理)

写真1. 村田氏による講演

きるロボットの導入が有効である。研究室の学生も交え、このような対話型ロボットを使って学修支援を行うシステムの開発をAgoraにおいて進めている。

電気通信大学では、「阿吽の呼吸」を理解し、「空気を読む」能力を持った人間のよき相棒としてのAIをイメージし、単なる道具ではなく、人間の相手となる汎用AIの実現を目指している。現在のAgoraのセンサーだけでは実験環境として十分ではないことから、ストレス度や集中力を測定できるセンサーを椅子に取り付けること等、センサーのさらなる充実を検討中である。将来的には、AI研究の成果を活用することで、知能化された学修環境が利用者の創造性を誘発し、イノベーションの創出を促す次世代型図書館を実現したい、とのことだった。

2. AIベースの文献検索プラットフォーム「Qinsight」のご紹介

講師：井手孝次郎氏

AIを活用したサービス「Qinsight」は、米国Quertle社が開発・提供している情報検索ツールである。海外では2008年から提供を開始し、国内では2017年10月からユサコ株式会社が販売を開始した。Qinsightはバイオメディカル・ライフサイエンス分野に特化した文献検索プラットフォームであり、検索は英語で行う。元研究者であるJeffrey社長とVicki副社長が、より簡単に関連の高い情報が効率よく検索できないかという思いから開発した商品である。

Qinsightは、特許情報を含む4,300万件以上のドキュメント、1,300万件以上のフルテキストの内容を、概念ベースで検索することができる。検索する際の複雑な演算子は不要であり、最も重要と考える概念を入力すると、大文字・小文字の意味の違い、類義語・同義語・省略形をAIが自動的に判別し、まとめて検索する。入力されたキーワードそのままを検索するのではなく、そのキーワードが意味のある関連性で使用されている文章を発見できる。入力する概念は、自然語のフレーズでも検索可能である。

また、視覚分析（Visual Analytics）を提供しており、検索結果リストからは判別できない概念間の繋がりや関連性等を発見することができる。

Qinsightのもう一つの特徴として、概念をグループ化したキーワード「Power Term」の機能があり、そのグループに含まれる具体的な概念で検索することができる。現在は、60個程の「Power Term」の登録があり、リクエストがあれば検証し、妥当性が承認された後に追

写真2. 井手氏による講演

加される。

検索例として、他のデータベース「PubMed」や「Google Scholar」との比較、「Qinsight」の検索の仕組みと実際の画面の説明があった。検索の仕組みとして「概念ベースのナレッジデータベース」があり、見た目上全く異なる単語同士の関係性や論文の中で著者が述べている事実内容を理解する。その上で、キーワードを基に利用者が求めている回答を出す。実際の検索結果の表示には、AIによる検索結果「Focused Results」だけでなく、どちらかという従来のような検索結果「Broader Results」もあり、こちらから検索漏れを確認することができる。最も重要な部分として、ヒットした根拠となる部分を表示する「Relevant Statements」があり、否定的・肯定的な文脈で記載されている結果のみに絞り込むフィルタリング機能もある。

さらに、データの提供元サイトへのリンク、リンクリゾルバへのリンクがある。

また、カスタマイズすることで、機関独自で保有のデータの検索や、独自の視覚分析（Visual Analytics）の提供も可能である。

3. 治験計画立案支援のための情報収集AI技術

講師：木戸邦彦氏

治験計画立案支援のための情報収集におけるAI技術について、日立製作所の取り組みが紹介された。まず本題に先立ち、その背景にある技術についての説明があった。

情報源であるテキストからの情報を活用する作業として、テキストからデータを抜き出す「抽出」、抽出したデータを「構造化」し、AIに学習させる「知識化」の3つの段階がある。構造化の例として、抽出した単語の検

索項目と検索数値の一体化，単語への症状変化の付与等がある。知識化については，投薬と症状の関係性を抜き出しAIに学習させる等の例が挙げられる。

このうち「抽出」については，専門用語の抽出等は既に製品化されている。「構造化」「知識化」については研究段階であるが，分野によっては実用化段階にある。

構造化における医療系テキストを対象とした情報抽出技術については，欧州研究学会「CLEF」が主催する医療情報抽出技術コンペティション参加の取り組みと結果が報告された。

時間的關係，対象患者，体の部位の属性，症状（悪化，改善等の傾向），文章の内容，その他の属性（否定や不確かさ等）を表に埋め込むタスクの説明があった。AIに学習させ，判別判断処理を行った結果，全体平均は87%，部位は79.7%の精度であり，コンペティションにおいて総合1位を獲得した。このように，構造化も精緻にチューニングを施せば，90%程度の精度がでる技術段階にある。

知識化については，ディベート型判断支援AIの研究開発に関する取り組みが紹介された。従来のイエスまたはノーで回答するようなQ&Aシステムではなく，人々の価値観を考慮して判断する仕組み，技術を目指している。それは，従来のシステムでは対応が難しい経営的・社会的問題等を賛否両論の面からリサーチし，判断の根拠を提示することで，論理的な対応を支援するものである。

内部構造として，学習・根拠の収集・価値体系辞書の3つについて説明があった。

最初に，ディープラーニングの注目箇所推定機能等の技術を使用して，判断事例をAIに与えて学習させる。これにより，回答すべき問題に対しての根拠文を収集し，価値観（賛成意見か反対意見か等）を覚えさせる。その方法は，10個のトピックに対して，30程度の観点を与える事例を300件程度用意して学習させるというものである。ディープラーニングを使用すると，言語依存性がないというメリットがあり，意外性のある根拠を提示することが可能となる。

ある社会において構成員個人の持つ価値意識の総体が，その社会の文化のもとで体系化されているような状態を価値体系というが，データベース化されたものを価値体系辞書という。相関関係データベースは，大量の記事中の文章を読み込み，ある価値に対して促進か抑制か解析し，判断根拠の事例として管理するものである。価値体系辞書と相関関係データベースを組み合わせることで，様々な観点・複数の価値を考慮した判断根拠を提示できる。この判断支援AI技術は，まだ試作品の段階であり，現在研究中である。

続いて，これらの技術をもとに構築される治験計画立案支援のための情報システムについて紹介があった。なお，この新薬開発における臨床試験の効率化に向けた取り組みは，自社に限らず田辺三菱製薬株式会社との広く

写真3. 木戸氏による講演

異なる視点を融合する顧客協創により、実施されている。

治験実施計画を作成するには、目的・デザイン・用法用量・対象とする被験者等を定め、治験計画の主要項目を決定する。新薬開発の成功率の向上、コストの適正を摺り合わせながら最適な治験計画を作成する。最適な治験計画をAIがつくることを最終目標としているが、まずは、AIが思考を補助することを第一段階として提案している。

過去の治験情報を集めるデータソースとして、PubMedのアブストラクト、各国規制当局（FDA, EMA, PMDA）の審査報告の情報、ClinicalTrials.govの情報、メジャーな医学雑誌の情報等を取り込む。データソースと治験の主要項目がどのような関係にあるか整理する際に、データソースを構造化することが必須となる。

この情報システムは、全体の仕組みとして情報収集AIと判断支援AIの二つの構成要素から成り立っている。判断支援AIについては、2018年度から取り組みを始めたところである。

情報収集AIは、情報ソースから取り込まれた情報を構造化して収集し、治験のフェーズを判別する。さらに、文章のどこに治験関連の情報があるかを同定し、用法用量等の細かな情報を抜き取りデータベースに貯めていき、判断支援AIが使用できる情報に加工するという大きな役割を担っている。

最後に、論文のアブストラクトから抜き出した具体的な処理内容の例が提示された。薬や作用機序等を入力すると、試験デザインの有効性評価ができ、横断的且つ多様的に可視化できる。治験計画立案支援のための情報システムの構築については試作品の段階で、実用化に向けて評価中であるが、一部は2019年2月に製品化予定であるとのことだった。

Ⅲ. おわりに

1990年代の検索エンジンは、2019年現在の検索エンジンとは比較にならない程お粗末なものだった。しかし、この20年間で蓄積されたビッグデータと技術の進歩により、検索エンジンの性能は飛躍的に向上しており、情報収集のスピードは確実に速くなった。その結果、ネットで調べられることも増え、学生や教員が図書館で辞書を引く姿を見なくなった。参考調査を所蔵館に問い合わせる機会が激減する等、図書館業務も20年前と大きく形を変えている。しかし、図書館の隅にひっそりと置かれている辞書類を利用する姿を見なくなっても、辞書の考え方が検索の基本的考え方であることに変わりはない。

10年後20年後にAI技術がさらに進化し、一部の図書館業務がAIに取って代わられることがあるかもしれないが、時代にあわせて形を変えても図書館の仕事が生き残れるように、今を生きる我々が成果を残していく必要があると考える。

参考文献

- 1) Life/style TECHNOLOGY 自閉症児向けセラピーに人型ロボットが活躍：技術. Newsweek. 2018;33(48):58.
- 2) (株)日立製作所. 電子カルテから病態や病気の部位等の指定された情報を高精度に抽出する技術を開発 [internet]. <https://www.hitachi.co.jp/New/cnews/month/2014/09/0917b.html> [accessed 2019-01-21]
- 3) (株)日立製作所. 田辺三菱製薬と日立が、AI技術を活用し新薬開発における臨床試験の効率化に向けて協創を開始 [internet]. <http://www.hitachi.co.jp/New/cnews/month/2018/03/0326a.html> [accessed 2019-01-21]