

2024年10月23日
システムワークフロー検討作業部会

「次世代検索サービスの検討と実現について（懇談）」に関する検討状況

第15回システムワークフロー検討作業部会（10月7日開催）において、本件について懇談を行い、学術情報機関にどのような生成AIの活用が求められるのか、検討を行った。システムワークフロー検討作業部会では、AIを用いた次世代検索サービスについて、①求められる機能、②技術的な要素、③利用場面の3つの観点より、今後継続して検討を行う予定である。

■ 話題提供・意見交換

<記事紹介>

- Taylor & Francis 社、AI 活用に向けた提携の一環として同社のコンテンツやデータへのアクセスを Microsoft 社に提供

<https://current.ndl.go.jp/car/224458>

- 海外では、大手出版社である Taylor & Francis 社が Microsoft 社にコンテンツ（電子ブックや電子ジャーナル）を提供するといった、学術情報を AI に取り込み活用する取り組みが始まっている。日本国内においても、CiNii Research や IRDB 等に収録された OA のデータを用いて同様の取り組みができないか検討する必要があるのではないか。
- Microsoft 社は検索エンジンである「Bing」に対する AI 活用・開発を進めており、利用率向上につながっている。
- 海外大手出版社の Elsevier や Springer 等でも同様の取り組みを行っている。

<目録・メタデータ班における検討>

1) 著者名典拠の統合と充実

- 複数のデータソース（紙、電子、デジタルアーカイブ）から著者情報を統合する。
- ORCID ID や所属情報を NOTE フィールドに追加する。
 - NCR2018 適用を機に、典拠データを作成することの重みが増した。典拠データのうち、著者名典拠の統合と充実を目指し、紙の資料を中心としている NACSIS-CAT に対して、複数のデータソース（紙、電子、デジタルアーカイブ）からのデータを統合することを目指すべきと考える。また、NACSIS-CAT には ORCID ID や所属情報が入っていないが、研究者情報を集約するといった目的のもとで、researchmap よりも充実したものを、AI 技術を用いて作り上げられ

たらよいのではないか。

2) 日本での著作典拠の統合

- 各大学図書館と協力し、データの不足部分や必要な要素を特定する。
- NDL サーチのような統合的な検索環境を提供し、精度を向上させる。
 - 著作典拠は、RDA をはじめとした世界の目録規則の中で重要度を増しているが、日本ではこれまで著作典拠の管理について一体的な取り組みが行われてこなかった。今後は、各大学図書館と協力し、データの不足部分や必要な要素を特定し、NDL サーチのような系統的な検索環境を提供できるようになればよい。著作典拠は、機関を超えて国で一つの著作典拠を作成し、有することが理想であり、そのためにも情報交換を通して精度を向上することが重要である。

3) LRM (Library Reference Model) 思想

- コアとなるデータ項目を決定し、それ以外は自由にスキーマを発展させる。
- 統合する際には必須項目だけを共有する。
 - LRM (Library Reference Model) 思想は、FRBR の考え方の根底にあり、著作や著者に関するコアとなるデータ項目を決定し、それ以外は各機関で自由にスキーマを発展させることができ、統合する際には必須項目のみを共有するといったものである。ダブリンコアにも同様の発想があり、図書に関してもそれに類した考え方が発展するのではないかと考えられる。

4) ILL (Interlibrary Loan) への活用

- NDL サーチのように ILL にも活用可能な統合データベースを構築する。

5) 著者名典拠、著作典拠データの機械学習・AI の活用

- 典拠の自動判定、並立書誌の統合、名寄せなどを AI で実現する。
- 既存データを活かし、関連度の高い情報を提供する。

6) 電子ブックなどの主題分析

- 本文データを活用し、より精度の高い主題分析を行う。
 - NACSIS-CAT では主題の付与が必須ではなく、件名が付与されていないデータが多いため、件名の自動付与について AI を積極的に活用することができたらよい。

7) ヨミ・翻字の現状

- ヨミや翻字の必要性は低下しており、AI による自動化が可能である。

8) ファジー検索

- AI を使い、CiNii Research に検索動向データを取り入れ、サジェスト機能を強化する。

<電気通信大学の事例>

- 生成 AI に関しては、大学の授業で取り扱われており、学生の 9 割以上が利用してい

ることを踏まえ、図書館での活用方法を探っている状況である。一般利用が可能な生成 AI を活用した論文・情報検索エンジンとして、Perplexity AI や Consensus 等が台頭してきたが、それらに対する課題として、ターゲットが絞られていないことや検索精度のばらつき、最終的に人の目を入れるといったプロセスの重要性、学生が考えさせるプロセスを組み入れること等が挙げられた。

- ノルウェーのスタートアップが開発した研究論文を推薦するツール「keenious」を、2024年10月よりテスト導入予定（試行期間は1年3か月）である。授業において、論文検索に関する課題に活用する予定である。
 - <https://keenious.com/>
 - 概要：生成 AI から情報を引き出す際、Prompt のドキュメント/テキストを元に分析し、研究論文を推薦する。情報源は OpenAlex を用いている。アメリカやイギリスの大学で導入が進んでおり、日本ではテスト導入が始まるところである。クロスランゲージ機能を有しており、様々な言語対応をしている。

< 慶應義塾大学の事例 >

- 慶應 AI センター
 - <https://www.keio.ac.jp/ja/news/2024/9/27/27-161978/>
 - 概要：2024年9月設立。言語や画像を学んだ基盤モデルの開発や、基盤モデルを用いたロボット研究、自律型 AI と人との共生研究、科学的発見のための AI 研究などを進める。

< 日本図書館研究会 >

- 日本図書館研究会 セミナー
 - <https://nalseminar2024.peatix.com/>
 - <https://www.nal-lib.jp/2024seminar/>