



White Paper

INSIGHT AS A SERVICE: INFO PROS AND ARTIFICIAL INTELLIGENCE

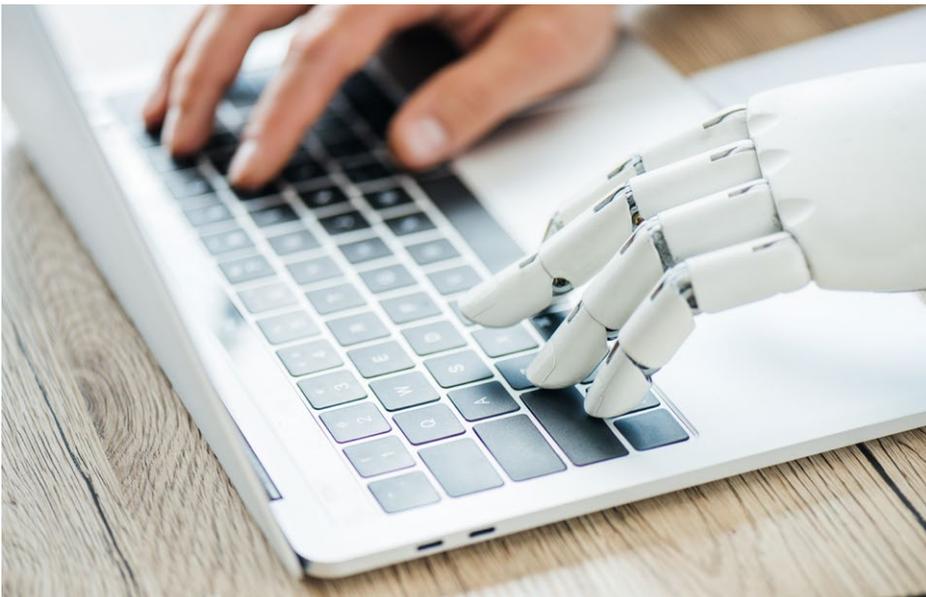
ADVANCING
DISCOVERY

Mary Ellen Bates

Introduction

“Thinking machines” and artificial intelligence started out with some bad press. Victor Frankenstein’s monster and the HAL 9000 computer onboard the spacecraft of 2001: A Space Odyssey were neither good citizens or user-friendly technology. These two examples point out both the power and peril of attempts to mimic human thought processes and decision-making that resulted in creations with tremendous power but no soul.

Today, advances in machine learning, the emergence of big data, and the exponential growth of raw computing power are bringing artificial intelligence into both the workplace and the home. Collaborative robots, or “cobots,” interact with humans on the factory floor, delivering parts or performing repetitive or dangerous tasks. Netflix, Amazon, and Pandora monitor user behavior in order to make personalized recommendations to their customers. Smart home devices can vacuum the rug, monitor the moisture levels in your garden, order more laundry detergent when you run low, and alert you to expired food in your refrigerator.



Libraries and AI

So far, most of the AI applications that have gone mainstream simply do what a smart human could do, only cheaper or more efficiently. But ever since information professionals and game show aficionados saw Watson—IBM's question-answering technology that can handle natural-language queries—trounce two Jeopardy! champions in 2011, we have wondered when our computer overlords would take over the reference desk.

IBM Watson Health has gone beyond mastery of trivia to create deeper insights into genomics research content from Springer Nature. This collaboration, in which the full text of five of Springer Nature's premier publications is incorporated into Watson Health, improves discoverability of relevant material. Watson learns from each query and from each user's navigation through the retrieved material, building its understanding of key concepts and relationships with each interaction. Springer Nature is also collaborating with BenchSci.com, a discoverability platform using AI to extract antibody usage data from the published literature. As content providers and publishers gain familiarity with tools such as natural language processing and machine learning, info pros can expect more of these search-enhancing collaborations.

Libraries are beginning to see the impact of artificial intelligence in both client-facing and back-stage operations. A chatbot on a library web site can handle directional questions, tell a patron when a book is due, point a user to relevant library resources, or even answer simple informational requests. While a chatbot interaction cannot take the place of an in-depth reference interview, these intelligent assistants can reduce the need for library staff to respond to routine questions while routing any unusual or difficult requests directly to an information professional.

Machine learning and sound- and image-recognition technologies are being used to analyze digital collections—text, images, audio, and datasets—to identify topics and entities, assign metadata, and enable non-textual search and discovery. Info pros are collaborating with machine learning technologies to enhance classification schemes to improve both search recall and precision. Some of the early projects in this area involve applying AI technologies to historical image collections—making subject headings consistent; extracting names, dates, and locations from descriptions; and even using visual pattern recognition to identify similar images.

Info pros are also using data visualization tools such as Springer Nature SciGraph Explorer (scigraph.springernature.com) to identify unexpected connections among concepts, researchers, and institutions. Imagine applying this level of textual analysis to an AI-enhanced library catalog; libraries and information centers could support their organization's innovation and discovery when they provide enhanced access to both the content and open-source tools that enable new insights.



AI Behind the Scenes at Springer Nature

In addition to high-profile initiatives such as IBM Watson Health and Springer Nature SciGraph Explorer, AI is embedded in Springer Nature's infrastructure. SN uses bibliographic reference analysis tools to create structured data from cited references and enhance discovery of publications through cross-referencing. Natural language processing tools add artificially generated keywords to publications in SpringerLink based on an analysis of its entire corpus of published material, improving retrieval of relevant information. SN partnered with The Open University in the UK to create the Computer Science Ontology—a taxonomy of computer science topics that was automatically generated with data mining technologies and without human intervention, and which can be updated regularly. These tools enable better trend detection from a global perspective than any individual could offer, and ensure more objective indexing and retrieval of information.

Info Pros' New Roles

In the early days of search engines, info pros had to play the role of interpreter and guide, helping their users navigate the wilds of the World Wide Web. They designed custom search engines to enhance and focus patrons' queries, built research guides for frequently requested topics, and developed search tools to access internal resources.

Today, info pros continue to participate in designing new knowledge discovery tools, built around the library ethos of finding the best answer, not just the most popular web site. Advertising-supported search engines not only have a financial incentive to track and monetize users' activities, but they also tend to de-contextualize information, surfacing a single page from deep within a web site isolated from clues with which to judge its veracity or reliability. Info pros can become involved in designing the next AI-based knowledge discovery tools and, just as they built customized search engines and created LibGuides, info pros can embed their focus on enabling the best information into these new tools.

While AI is often seen in high-end settings such as algorithmic stock trading or computer-aided interpretation of medical images, information centers of any size can find a role to play within their organizations. Public libraries can expand makerspaces to include a virtual reality lab or to offer workshops on machine learning tools.

Information centers within organizations can work with their IT departments to identify and curate open-source datasets and text and data mining tools that would be of use to specific user groups.

The University of Rhode Island recently established a new AI Lab within the university library; the lab offers students, faculty, and researchers the opportunity to learn about data mining, robotics, machine learning and related technologies, as well the ethical implications of these technologies. Similar models, scaled to the appropriate size based on the organization, could be implemented in a wide range of library settings.

Recommendation algorithms can be built into online catalogs or integrated library systems to enhance discovery of library resources. Even small information centers and solo librarians can bring AI concepts and applications to their users. They can start by creating an internal blog or web page with links to AI-related articles and videos relevant to their clients' interests; there are almost 300 TED Talks alone on artificial intelligence, as well as podcasts, online courses, and other resources for the non-expert. They can host brown bag lunches and virtual meetings on topics that encourage further exploration—"AI For Everyone" or "Machine Learning 101". They can use a free or low-cost version of [IBM Watson Assistant](#) to create a library chatbot as a demonstration of a relatively simple implementation of AI.



Evaluating information sources is a key skill of info pros. Just as they select resources for the library collection and subscribe to the digital content that offers the best ROI for their users, info pros can also work with project groups to select the most appropriate datasets for a machine learning or big data initiative. Their familiarity with information sources and techniques for evaluating the quality of a resource means they can ensure that the data being used accurately reflects the real world it is intended to represent. Info pros look at how the dataset was created and the data collected, where an implicit bias may have crept in, and whether more current datasets are available.

Info pros can see the strategic potential in bringing AI technology to a digital collection. Historical content—census data, historical manuscripts, news articles, and public records, for example—can be analyzed in new ways with text and data mining tools to reveal unrecognized cultural, social or historical trends. Info pros can identify opportunities to enhance existing collections with discovery tools that can add geolocation, identify named entities, and establish dates to identify temporal correlations.

Info pros can also lead the discussion within their organizations regarding user privacy and security issues when implementing an AI project. Now that recordings from household smart speakers such as Amazon Echo are being sought by investigators to help solve murder cases, concern about access to and use of personal data is not merely theoretical. The European Union's General Data Protection Regulation (GDPR) and privacy regulations being planned and enacted in other regions have raised the awareness of personal data—what is being collected and what is done with the information. Info pros have a unique perspective on the information flow within organizations, and can advocate for the protection of identifiable information within datasets being used in AI applications.

As Daniel Lee, Director of Enterprise Information Solutions at ARC Business Solutions noted, *"The biggest impact info pros can have concerning AI is to influence its development, specifically the people doing the development, with our ethics and philosophy—two areas that are irrelevant to a line of code or algorithm."*

What's on the Horizon?

As with any truly transformative technology, asking “How might AI help us do our work better?” misses the point. When the web first emerged, very few people anticipated what impact e-commerce would have on white-collar professions. The disintermediation of buyer and seller that the web enabled has disrupted professions such as travel agent, real estate agent, and insurance agent; even routine accounting and legal work can be done with a cloud-based service. The proliferation of e-books and other digital content has dramatically changed the appearance and functions of libraries large and small, public and private.

Simple AI tools can help with existing library processes without displacing information professionals. A chatbot may be able to handle ready-reference questions, freeing up staff to focus on more complex queries that need an info pro to understand the full context of the request and to assist or direct the user to the most appropriate resources for the situation.

The real impact of AI will not be in making existing library processes better but in enabling entirely new capabilities that address information needs that library users may not realize could be met. Info pros understand how their clients think about, search for and use information. They can recognize opportunities to introduce AI into their organization. They may bring a text and data mining tool to an internal dataset, helping a project team find new insight from existing data. They might acquire a data visualization tool to help users find unexpected connections in the published literature.

Just as libraries have been transformed from housing physical collections of books to providing online access to information, libraries will use AI tools to provide not just information but deep intelligence—offering “Insight As A Service.”

Info pros can prepare for providing IAAS by scanning the AI horizon, to get a sense of what kinds of tools may be relevant to their user groups. Sources such as *Nature Machine Intelligence* and the Springer Nature eBook *Intelligent Technologies and Robotics Collection* offer a wide perspective on trends in artificial intelligence and related technologies. Courses directed to non-engineers are available through online learning platforms such as Coursera and LinkedIn Learning (formerly Lynda). Even Andrew Ng, one of the pioneers in machine learning and AI, is offering a non-technical course in Coursera titled “AI for Everyone.” (coursera.org/learn/ai-for-everyone)

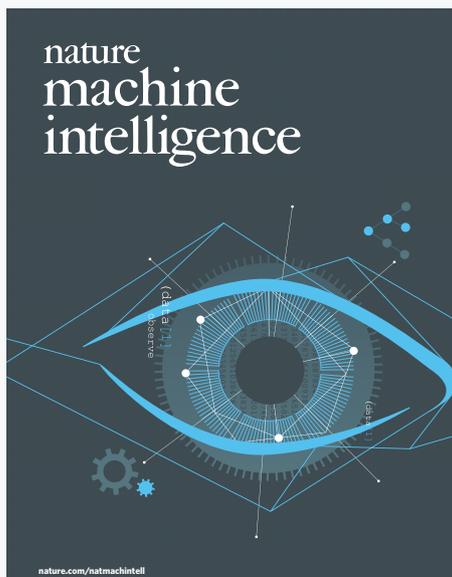
Info pros can make a practice of observing and exploring how information flows within their organization and identifying points at which information could be transformed or used in entirely new ways. Even if the library does not have the budget for advanced AI tools or resources, info pros can bring partnership opportunities to project teams or groups that could invest in the necessary technology or datasets. An AI lab maintained by the library could include open source AI platforms, a curated collection of open access datasets, and links to groups within the organization already involved in AI initiatives.



Just as, in the past, info pros developed skills in constructing complex Boolean search queries and at least a passing knowledge of HTML coding in order to manage information sources, so info pros now are become familiar with scripting languages such as Python, data visualization tools such as Tableau, and chatbot development platforms such as IBM Watson Assistant.

While to many, artificial intelligence may seem to be removed from daily library operations, being aware of its potential impact can help info pros recognize opportunities to drive the conversation about AI and enhance the value of information within their organization.

Content from Springer Nature



nature.com/natmachintell

SPRINGER NATURE springernature.com

eBooks

INTELLIGENT TECHNOLOGIES AND ROBOTICS COLLECTION

375 titles in the 2019 collection

ADVANCING DISCOVERY

• Best in Books
• Complete
• Always connected

springernature.com/ebooks

Springer

springernature.com/ebooks

Join the Conversation

As with all Springer Nature white papers, we aim to maintain an open channel of communication and knowledge sharing across our scholarly and research communities. We welcome your thoughts and reflections on this white paper and encourage you to reach out.

You can find content and further resources to support your activities from our dedicated website springernature.com/rd, or receive these directly in to your inbox by signing up for our alerts at springernature.com/alerts.

Springer Nature advances discovery by publishing robust and insightful research, supporting the development of new areas of knowledge and making ideas and information accessible around the world.