

Expectations from Artificial Intelligence: What Changed During the Decades?

Nicola Orio

Department of Cultural Heritage, University of Padua, Italy
nicola.orio@unipd.it

Abstract. Artificial intelligence has been often perceived as a promising field that can provide methods and tools that can be extended to other research areas and exploited in industrial applications. Digital libraries and archives can thus have a twofold benefit. Since research and evaluation can hardly be carried out without the development of a working system, based on a real collection accessed by real users, it could be expected that artificial intelligence plays a central role in the design and development of the many components of a digital library or of a digital archive. This short communication starts from a review of what was expected from artificial intelligence in the research areas that converge in the interdisciplinary field of digital libraries, concluding with a number of considerations on what could be today's expectations from the variety of persons – researchers, stakeholders, librarians, users – involved in the design, development and access to organized digital content.

1 Introduction

Methods and tools developed in the field of Artificial Intelligence (AI) are pervasive in many other research areas. Advancements in very different fields, from robotics to multimedia indexing were possible thanks to AI research results. AI is an interdisciplinary field both in the scientific background of researchers who contribute to the field and in the applications of its results.

Among other fields, it is probably particularly relevant to investigate the impact of AI in the development of Digital Libraries and Archives (DLAs). This because DLAs are themselves multidisciplinary both when considered as existing institutions and when considered as an applicative research area. The issue can be faced by comparing the early expectations, about how AI could contribute to the advancement of disciplines related to DLAs, with nowadays expectations, about the directions in which user experience can be improved using AI.

2 Early Expectations from Artificial Intelligence

One of the first concepts taught to university students in computer science and computer engineering is the, apparently simple, equation $data + context = information$. The context in which this important piece of information is delivered

is usually a course on database systems, where its direct implications in software design are more evident. In this case the goal is mainly to provide the user with all the necessary background to contextualize and give meaning to the data, but since the early days of automatic information processing the idea of exploiting AI techniques to automatize this process has been considered as an important issue.

For instance, already in 1980, Schank published a short communication [7] in which he envisaged that Information Retrieval (IR) effectiveness could be improved by a “system able to analyze and understand the natural language of both new text inputs and queries to the data base”. The idea of applying AI to IR, and more in general to a number of aspects related to DLAs such as human-computer interfaces and automatic data entry, has been further elaborated in 1983 by DeJong [1] where a number of examples are given where the focus is again on understanding and deducing information.

The integration of database and AI research has been discussed in the literature since a long time, as witnessed by the introduction of the term *intelligent database* [6] already in 1989. An intelligent database should be able to go beyond the classical boolean searches and introduces the idea of a semantic representation of data. There is a close relationship between the two disciplines. For instance, decision support and data mining are research areas based on the integration of AI and database. The relationship is so tight that there has even been an effort to integrate university courses on the two fields in a single teaching offer [12].

Another research field that is usually perceived as tightly connected with AI is Human-Computer Interaction (HCI). Also in this case, the relationship is witnessed by the introduction of the term *intelligent user interfaces*, although the main focus was probably on the development of natural language processing techniques. Yet, it has been argued [3] that AI researchers usually have limited interest in interface design and that in general HCI are more interested in human factors ergonomic research.

As regards DLAs, the idea of exploring how AI can contribute to the field motivated the organization, already in 1997, of the international workshop “Artificial Intelligence and Digital Libraries”, which was part of the International Joint Conference on Artificial Intelligence (IJCAI). It is interesting to note that in the same year it was presented a paper on the integration of machine learning techniques directly targeted to the development of an *intelligent digital library* [9]. One of the outcomes of the workshop was a special issue of the International Journal on Digital Libraries, edited by Ferguson and Durfee in 1998 [2]. The main themes raised by the editorial of that special issue are valid still nowadays:

- Information discovery and retrieval;
- User interface design;
- Classification and indexing;
- Architectural design and issues.

As it can be seen, these four main points are all related to the research areas previously described, from HCI to IR and database research. It can be

noted that most of these expectations have been, at least partially, fulfilled. For instance, machine learning plays a central role in the management of multimedia content, which is increasingly present in DLAs. Recent approaches to automatic identification and classification of images, music and video largely relies on AI techniques [4]. The interaction with the user is increasingly based on a holistic approach, which takes into account user needs, behavior and sentiments [8].

3 Nowadays Expectations from Artificial Intelligence

Given the centrality of users in a DL system, the integration of AI research results in working digital libraries has to take into account nowadays expectation of all the types of users, from stakeholders and librarians to the general public. These expectations are not supposed to be the main driver of individual research projects, but can highlight the general direction to which projects might converge. Even if there is a time span of 45 years, nowadays expectations can probably be expressed by a famous quote by Marvin Minsky, in a 1970 interview to *Life* magazine:

In from three to eight years we will have a machine with the general intelligence of an average human being. I mean a machine that will be able to read Shakespeare, grease a car, play office politics, tell a joke, have a fight. At that point the machine will begin to educate itself with fantastic speed. In a few months it will be at genius level and a few months after that its powers will be incalculable

This sentence actually summarizes the abilities of the main character of the movie *Her* (Spike Jonze, 2013), a hand-free operative system voiced by Scarlett Johansson. Apart from being able to compose music and create jokes, as envisaged by Minsky, the operative system in *Her* seems to understand the content of documents and actively use this ability to filter out irrelevant mails, assign priorities to messages, and even spontaneously select texts based on their artistic quality in order to publish a book. Of course, science fiction always looks ahead of the actual time, but at the same it is an important *litmus test* on which computer abilities are considered crucial for an effective improvement of the user experience.

Thus, today's expectations are probably not very different from the ones envisaged in the eighties and early nineties: AI should provide results that help computer programs to automatically infer a context in order to transform structured and unstructured data into understandable information.

To this general idea, it could probably be added another facet on the concept of understandability. Users, in particular scholars, are willing to understand why a system recommends, suggests, filters, a particular set of documents [11]. That is, not only an AI-powered system should be able to understand documents content, but the process has to be understood by the users. Understanding the underling process, even if they might not understand the actual formal models and algorithms, allows users to predict the system behavior and to direct the

interaction towards the expected results. In a sense, this can be considered as another variant of the well-known *Turing test*. Instead of asking “*Can machines think?*”, users of an AI-powered DL might ask “*Can I understand, and thus trust, the results of machine thinking?*”.

4 Conclusions

The need for integrating intelligent techniques in all the many components of DLAs, from search tools to data management and user interfaces, emerged decades ago and resulted in the development of many *intelligent* tools that are used to manage large digital multimedia collections. Yet, very few digital library systems dare to call them selves intelligent [5]. This is probably due to the complexity of a DLA system, where the existence of a number of tools that implement AI results is not sufficient to perceive a complete digital library or digital archive as an intelligent actor to interact with. The design of a system centered on the final user experience can be an important goal for future research.

References

1. G. DeJong. Artificial Intelligence Implications for Information Retrieval. In *Proc. of ACM SIGIR*, pages 10–17, 1983.
2. I.A.. Ferguson and E.H. Durfee. Artificial Intelligence in Digital Libraries: Moving from Chaos to (More) Order. *International Journal on Digital Libraries*, 2(1):1–2, 1998.
3. J. Grudin. Turing Maturing: The Separation of Artificial Intelligence and Human-Computer Interaction. *Interactions*, 13(5):54–57, 2006.
4. A.-E. Hassanien, A. Abraham and J. Kacprzyk. *Computational Intelligence in Multimedia Processing: Recent Advances*. Springer Publishing Company Inc., 2008.
5. J.P. Leidig and E.A. Fox. Intelligent Digital Libraries and Tailored Services. *Journal of Intelligent Information Systems*, 43(3):463–480, 2014.
6. K. Parsaye et al. *Intelligent Databases: Object-Oriented, Deductive Hypermedia Technologies*. John Wiley & Sons Inc., 1989.
7. R.C. Schank. Artificial Intelligence and Information Retrieval. In *Proc. of ACM Annual Conference*, pages 157–17, 1980.
8. B.W. Schuller. Modeling User Affect and Sentiment in Intelligent User Interfaces: A Tutorial Overview. In *Proc. of International Conference on Intelligent User Interfaces*, pages 443–446, 2015.
9. G. Semeraro et al. IDL: A Prototypical Intelligent Digital Library Service. In *Proc. of the Congress of the Italian Association for Artificial Intelligence on Advances in Artificial Intelligence*, pages 447–450, 1997.
10. L.C. Smith. Artificial Intelligence in Information Retrieval Systems. *Information Processing and Management*, 12(3):189–222, 1976.
11. C.M. Steiner et al. Evaluating a Digital Humanities Research Environment: The CULTURA Approach. *International Journal on Digital Libraries*, 15(1):53–70, 2014.
12. P.A. Taylor and D.L. Wyatt. Database and Artificial Intelligence Integration: A Challenge to Academia. *ACM SIGCSE Bulletin*, 24(4):35–42, 1992.