Information Retrieval Research 1 (2025) 339-343 Published 12/2025

Editorial

AI Resilience and Information Retrieval

Paul B. Kantor PAUL.KANTOR@RUTGERS.EDU
University of Wisconsin
Madison, Wisconsin, United States of America

"Tout est pour le mieux dans le meilleur des mondes"
Dr. Pangloss, in Voltaire's Candide

IR provides multiple versions and sources to confirm an assertion. AI provides (essentially statistical) compilation and weighting of all those inputs. And generates a coherent and syntactically appropriate statement. What could go wrong?

With the rise of artificial intelligence, we have all asked what role Information Retrieval (IR) is to play, if any, in this new technological revolution. I want to suggest that there are three dimensions to this, all bearing on the complex relation between the world of documents and the physical world in which we all live. I'll put my concerns under three distinct headings: consistency, confidence, and completeness.

As to consistency: all of us who've engaged in what one of my colleagues has called "the torturing of AI" can provide instances where a generative AI (genAI) has confidently provided contradictory answers to the same question. If the AI were a human assistant, we would only need to point out the contradiction for that human assistant to understand that when there is a contradiction, both answers cannot be correct. At most one of them is.

While instances of genAI will apologize quite politely, they do not seem to understand that as descriptions of the real world (as opposed to, say, the world of spiritual beliefs) it is not acceptable to make contradictory assertions. To this, the friends and relations of genAI respond that since computers will soon be able to discover, and then prove mathematical theorems, they will "of course" be able to recognize and correct contradictory assertions about the real world. I've been wrong about many things before, but I remain intensely skeptical about this hope. Aside from the profound difficulties of mapping the real world to mathematics, Gödel (1931, 1987) has shown that even in mathematics not every true statement must be provable.

As to confidence: for many years we in IR have worked to adorn the results of information retrieval with numbers indicating confidence. More than 50 years ago, Sherman Kent, working at the United States Central Intelligence Agency, noticed that reports and assessments were filled with words reflecting confidence, but no one had sought to normalize the meaning of those terms. In a famous note (Kent, 1964) he proposed that these terms be assigned to ranges of probability. While admirable, this presents the challenge of assigning

 WWW: https://irrj.org DOI: 10.54195/irrj.25004

probabilities to decisions that may not yet have been made, and will be unique events. What does it mean to say that "the probability that Madame X will win election in country Y is 75%"? (Menzel, 2025). Even if we resort to the quantum mechanical concept of a many worlds picture (Everett, 1957), we seem to live in only one of them, and so the answer is "yes" or "no." There are Brier scores (Wikipedia, 2025a, "Brier score"), and other ways to compute how well we are doing with such predictions. But the truth is that such scores represent aggregated accuracy, and are only meaningful when they are compiled over a large number of scored projections.

However, what I mean today by confidence is something less esoteric, but more troubling. Across the world, swindlers, politicians, and other kinds of "confidence men" have long known that the key to deceiving us is to look and sound like someone who can be trusted. In the generation of texts, we call this "the look and feel" of being trustworthy. Sadly, by exploiting the relatively low entropy of text and natural language, genAI is precisely able to achieve the correct "look and feel" absolutely without regard to whether what it says is or is not true. In other words, we have finally built a machine that is specifically capable of fooling us about whether it knows what it is talking about.

This leads to the third point: *completeness*. The problem here is what I call, with apologies to a man who is in some ways, both the father and the godfather of the scientific revolution, the "*Aristotelian fallacy.*" Aristotle was at heart a scientist, who gathered as many facts as he could. However, in organizing and understanding those facts, and seeking deeper understanding, he placed what we now believe is too much faith in deduction, and not enough in induction or in abduction, in its original sense (Douven, 2025). Although he used induction (the weighing of evidence on one side or another) Aristotle believed that any such conclusions also needed a deductive justification, such as "the efficacy of this form of government flows from the virtue of the citizens of the state." (This is a notion that was accepted without question at the founding of America, and is under substantial threat as I write). The modern scientific revolution began by rejecting a belief that all explanations are to be found in that "Aristotelian" way (Wikipedia, 2025c, "Roger Bacon").

At any specific moment in the evolution of human knowledge, more than induction and deduction are needed. Abduction supports the *creation of hypotheses* spurred by some consistency or inconsistency in the data about the world. Or, indeed, for the desire for more parsimonious explanations. It is an equal, perhaps even more critical component of scientific and technological advance. For some years, historians of science such as (Galison, 1997), have opposed the so-called Kuhnian notion of science as progressing by "change of [theoretical] paradigm" and have emphasized the importance of technologies in the advancement of science. Indeed, technology and theory fare best when they are united to form science. In 2025 Nobel laureate Joel Mokyr (Mokyr, 2005) was honored for pointing out that technological advance is hugely exponentiated by corresponding theoretical advances, which give us, its human masters, some usable understanding of both its strength and its limitations.

In this context, we must think carefully about the relation(s) between "the real world" and "the literature about that world." There is an axiom in chemistry "Why spend a day in the library when you can learn the same thing by working in the laboratory for a month?" (Corey, 2007). Indeed, the IR researcher Donald Swanson initiated the possibility

of finding an unexpected connection between a treatment and an illness, not in the clinic, but in the literature, a concept now called literature-based discovery (Wikipedia, 2025b, "Literature-based discovery"). While science, as the joke reminds us, wants to "look for a lost wallet under the streetlamp," unexpected connections revealed in the literature give us "new kinds of streetlamps." However, *it cannot be the case that whatever is to be discovered tomorrow is hidden in the accumulated knowledge of today.*

The literature of all mankind at any given moment, even if we hoovered up the contents of all nine billion brains, simply cannot contain predictions of all that will be found later. Nothing known or imagined before the invention of the microscope was anywhere near a true picture of what it would reveal. Before graphene was invented, using only scotch tape and pencil lead, there was no description in any text or brain of how it could be produced and of its astonishing physical properties.

So, taken together, these concerns about consistency, confidence, and completeness suggest that Retrieval-Assisted-genAI may slouch toward Bethlehem on three feet of clay. Can the techniques and principles of information retrieval help to remedy these weaknesses?

The danger involved, when machines will read what other machines have written, and the original "facts on the ground" will become ever more remote, was anticipated, in the magnificent dystopian novella called "The Machine Stops" (Forster, 1909). In some future world humans live in identical underground cells, and communicate their ideas without ever leaving those cells. At what we would now recognize as Zoom sessions, they hear from expert "lecturers" who, to become expert, had ventured up to the surface of the earth, to observe it. But the machine that maintains air, food, heat and light for all of the people had decided to eliminate the essential respirators, without which no one dared to venture to that real world. Let Forster, speaking from the grave, tell us how humanity responded:

> *. . . even the lecturers acquiesced when they found that a lecture on the sea was none the less stimulating when compiled out of other lectures that had already been delivered on the same subject. "Beware of first-hand ideas!" exclaimed one of the most advanced of them. "First-hand ideas do not really exist. They are but the physical impressions produced by love and fear, and on this gross foundation who could erect a philosophy? Let your ideas be second-hand, and if possible tenth-hand, for then they will be far removed from that disturbing element — direct observation. Do not learn anything about this subject of mine — the French Revolution. Learn instead what I think that Enicharmon thought Urizen thought Gutch thought Ho-Yung thought Chi-Bo-Sing thought Lafcadio Hearn thought Carlyle thought Mirabeau said about the French Revolution. Through the medium of these ten great minds, the blood that was shed at Paris and the windows that were broken at Versailles will be clarified to an idea which you may employ most profitably in your daily lives."*

Grim stuff. All three of the stated problems: consistency, confidence, and completeness seem, in one way or another, to circle back to these questions: are AI-generated statements true? Can that truth, in any way, be inferred from knowledge about how the statements were generated? Is there any way that our skill in retrieving and organizing what has been known and recorded can help us to know whether it and what we deduce from it is true?

There are some threads of hope related to the ideas of tracking provenance, authentication, and some kind of agreement about what is to be considered a faithful source of knowledge. Perhaps if the World Wide Web had more closely followed the road map of Ted Nelson's Xanadu (Nelson, 1999) we would be in better shape. Perhaps if search, and guideposts to related documents had become distributed rather than centralized (Kantor et al., 2000) we would be in better shape. But today the Internet lets any of us share our opinions. Any centralized robot can weigh and count them. The Internet itself can be locally shaped by nation-states wishing to advance one or another view of the real world. All in all, it is very hard to believe that we can define any scientific process that will tell us what to trust.

There is still room for optimism. My coeditor of the predecessor Information Retrieval Journal, Stephen Robertson, with his colleagues, devised a weight formula (Robertson et al., 1994) whose name, BM25, hints that it was better than perhaps some 24 other "Best Match" methods"[1]. In his Salton award lecture Robertson (2000) Steve argued quite well that IR "cannot have a theory." In this note I am arguing that establishing the "truth" of retrieval-augmented AI, as part of the IR enterprise, cannot become a science. Thus, it may fail to meet Mokyr's criterion for exponential growth – circular funding deals notwithstanding.

The advocates and builders of the new AI do recognize some risks. We surely lack a scientific understanding of its powers, and its limitations. But perhaps some clever *ad hoc* tool, some Truth Finder 25, will help us to hold our ground against the advancing flood of overconfident assertions. As we search for it, I hope that the innovators and inventors of such *Truth Finding* tools will share their ideas with us in the pages of this important new journal.

References

E. J. Corey. Frank H. Westheimer, major figure in 20th century chemistry, dies at 95. *Harvard Gazette*, April 19 2007. URL `https://news.harvard.edu/gazette/story/2007/04/frank-h-westheimer-major-figure-in-20th-century-chemistry-dies-at-95/`.

Igor Douven. *The Stanford Encyclopedia of Philosophy*, chapter: Abduction. Metaphysics Research Lab, Stanford University, 2025. URL `https://plato.stanford.edu/archives/sum2025/entries/abduction/`.

Hugh Everett. *'Relative State' Formulation of Quantum Mechanics*. PhD Thesis, Princeton University, 1957. URL `https://journals.aps.org/rmp/abstract/10.1103/RevModPhys.29.454`.

E. M. Forster. The machine stops. *The Oxford and Cambridge Review*, 1909. URL `https://www.cs.ucdavis.edu/~koehl/Teaching/ECS188/PDF_files/Machine_stops.pdf`.

1. I believe there were both fewer and more than 24. The name may refer to two components of the model. As their paper explains, the Okapi team explored many values of several model parameters to find the published model. This remains the basic principle by which Artificial Intelligence models are tuned and trained today.

Peter Galison. Image and logic: A material culture of microphysics. *University of Chicago Press*, 1997.

Kurt Gödel. Über formal unentscheidbare sätze der principia mathematica und verwandter systeme I. *Monatshefte für Mathematik und Physik*, 38:173–198, 1931. doi: 10.1007/BF01700692.

Kurt Gödel. *Collected Works, Volume I: Publications 1929–1936*, chapter: On Formally Undecidable Propositions of Principia Mathematica and Related Systems I. Oxford University Press, 1987.

Paul B. Kantor, Endre Boros, Benjamin Melamed, Vladimir Meñkov, Bracha Shapira, and David J. Neu. Capturing human intelligence in the net. *Communications of the ACM*, 43(8):112–115, 2000.

Sherman Kent. Words of estimative probability. *Studies in Intelligence*, 8(4):49–65, 1964. URL `https://www.cia.gov/resources/csi/static/Words-of-Estimative-Probability.pdf`.

Christopher Menzel. *The Stanford Encyclopedia of Philosophy*, chapter: Possible Worlds. Metaphysics Research Lab, Stanford University, 2025. URL `https://plato.stanford.edu/archives/fall2025/entries/possible-worlds/`.

Joel Mokyr. The intellectual origins of modern economic growth. *Journal of Economic History*, 65(2):285–351, 2005.

Theodor Holm Nelson. The unfinished revolution and Xanadu. *ACM Computing Surveys*, 31(4es), 1999.

Stepen E. Robertson, Steve Walker, Susan Jones, and Michelle Hancock-Beaulieu. Okapi at TREC-3. In *Proceedings of the Third Text REtrieval Conference (TREC)*. NIST, 1994. URL `https://trec.nist.gov/pubs/trec3/t3_proceedings.html`.

Stephen E. Robertson. Salton award lecture on theoretical argument in information retrieval. *ACM SIGIR Forum*, 34(1):1–10, 2000.

Wikipedia. Brier score, October 4, 2025a. URL `https://en.wikipedia.org/w/index.php?title=Brier_score&oldid=1315096299`.

Wikipedia. Literature-based discovery, September 30, 2025b. URL `https://en.wikipedia.org/w/index.php?title=Literature-based_discovery&oldid=1314260367`.

Wikipedia. Roger Bacon, September 15, 2025c. URL `https://en.wikipedia.org/w/index.php?title=Roger_Bacon&oldid=1311560035`.