



AI research assistants, intrinsic values, and the science we want

Ariel Guersenzvaig¹ · Javier Sánchez-Monedero²

Received: 3 August 2023 / Accepted: 22 December 2023

© The Author(s), under exclusive licence to Springer-Verlag London Ltd., part of Springer Nature 2024

Since their mass introduction in late 2022, AI chatbots like ChatGPT have garnered considerable attention due to the promise of widespread applications. Their purported advanced writing capacity has made it difficult for experts to differentiate between machine-generated and human-generated paper abstracts, as reported in *Nature* (Else 2023). However, many scholars emphasize that these systems should be seen as ‘stochastic parrots’ due to their lack of true understanding (Bender et al. 2021). Furthermore, these systems have been prone to produce ‘hallucinations’ (i.e., falsehoods), among other highlighted issues.

This is not the venue for an exhaustive critique; our purpose is to comment on a rather specific topic: the use of chatbots for the automation of research and bibliographical review that tends to precede all academic research. As an example, consider Elicit, a tool that aims to optimize the flow of academic research. According to its developing company, ‘If you ask a question, Elicit will show relevant papers and summaries of key information about those papers in an easy-to-use table’ (<https://elicit.org>, [faq](#). xxxx). It apparently does this by finding the most important information from the eight most ‘relevant’ articles among a selection of 400 articles that are related to the question. Alternatively, think of Perplexity Copilot (<https://blog.perplexity.ai>, [faq](#), [what-is-copilot](#). xxxx), which offers a ‘tailored list of sources and even summarized papers’ to students and academics.

We often teach our students that through bibliographical research, we find out what has been said about a topic, what other related views or theories exist, what gaps are still to be filled, and so on. Importantly, we emphasize that it serves to establish the foundations of our own research. But is the review just a mere instrument that we could optimize

using tools like Elicit and Perplexity Copilot? To answer this question, we must first take a detour to address a more general issue related to the way science can be carried out.

1 Intensive science and extensive science

Hopefully, without oversimplifying the matter too much, we can draw a parallel with agriculture and distinguish between ‘intensive’ and ‘extensive’ science. (To clarify: intensive agriculture seeks to maximize the yield and profit from a crop using industrial machinery, fertilizers or abundant irrigation. Extensive agriculture, on the other hand, is more respectful towards the environment and seeks a more sustainable use of the land. Something similar occurs in the cattle industry, although with the addition of the important matter of animal wellbeing.)

Intensive science is one that is ‘successful’ in terms of quantitative results, such as the mass publication of papers and the maximization of valuable scores by scientific quality agencies. Intensive practices allow researchers to survive and succeed in their academic careers. Furthermore, it embodies a model of vertical specialization focused on quantity rather than on less tangible or intangible aspects. We find instances of intensive scientific practices in all fields of science; also in the field of AI, where, for example, there are myriad articles on systems presented as able to ‘detect emotions,’ even though such systems are in reality incapable of detecting emotions without reducing the enormous complexity of human emotions to that which can be measured with an AI system, even at the cost of having to establish false universal categories and eliminate from the algorithmic model all references to corporality, context and culture.

So, if the main interest of a researcher is to ‘produce articles primarily for the purpose of obtaining citations and improving other ‘quality’ indicators, even if the articles do not possess much intrinsic value, it is probable that prior bibliographical research serves as nothing more than an instrument rather than an end in itself. In this case,

✉ Ariel Guersenzvaig
aguersenzvaig@elisava.net

¹ Elisava Barcelona School of Design and Engineering,
University of Vic-UCC, Barcelona, Spain

² Department of Computer Science and Numerical Analysis,
Universidad de Córdoba, Córdoba, Spain

researchers may seek to compile articles and catalog them, at most reading keywords, titles, and abstracts, thus increasing productivity. Can this task be automated?

Yes, and in this sense, tools like Elicit and Perplexity Copilot, excepting some technical difficulties that may be improved in the future, can be useful for this type of intensive science, which is reductionist and centered on volume and efficiency. We informally run tested these tools to obtain references to AI-driven emotion recognition and their results included sources only from computer science but nothing from psychology or anthropology, which should be relevant for *any* researcher working on emotions. However, this reductionism does not matter in the intensive science model, where monoculture is the rule and for which interdisciplinarity may even present a threat to the meaning of their task.

Conversely, if we are driven by extensive science and seek to explore the true complexity of emotions, starting with the multidimensionality of the concept rather than settling for simplistic interpretations, tools like Elicit and Complexity Copilot are likely to be of limited use.

This leads us to a different response to the question posed earlier. In the context of extensive science, the answer is that automating bibliographical research is not an easy task, as it is not a mere instrument. Aristotle said in his *Nicomachean Ethics* that the practice of medicine did not merely consist of cutting or not cutting, or of prescribing medicine, but rather of doing so *in a certain way*. We can draw a parallel with bibliographical research—it is not just about citing literature and referencing sources, but rather about doing so in a certain way.

In short, we have two preliminary answers: for intensive science, bibliographical research can indeed be automated, while for extensive science it can only be done partially. This point warrants examination in greater depth, which is what we will do next.

2 The intrinsic and instrumental values of bibliographic research

Philosophers distinguish between two types of values (or ends): intrinsic and instrumental. Intrinsic here refers to everything that has value in itself, such as friendship, health, fun and justice. Instrumental refers to things the value of which depends on their relation to something valuable, either by obtaining or preserving that thing. In this sense, instrumental values are also important.

For example, a drill has no intrinsic value, but it is useful for making holes in a wall so that we can hang pictures and enjoy them. On the other hand, the aesthetic pleasure derived from contemplating said pictures does have intrinsic value. We would certainly find little meaning in the question of

why we want to obtain aesthetic pleasure, given that it is something desirable in itself.

In some situations, a single thing can integrate both types of values: that is, it can be useful for something and simultaneously have value in itself. Take, for example, friendship. A friend can help us to get a better job or move house. At the same time, friendship has value in itself, and the value of having friends does not depend on a friend being useful for something. If we compare the value of friendship with that of a drill, the difference is obvious: a common drill has little to no value beyond helping us to make the holes we need. After this detour, let's return to bibliographical research.

Conducting bibliographical research exceeds instrumental as it enables us to develop our own research ideas, thus increasing our capacity for critical thinking and creativity, which have intrinsic value. Good bibliographical research situates our work and positions us within an academic tradition, which can and does affect our identity and the values and beliefs we adopt as researchers and as individuals. Doing bibliographical research is part of the process of constructing a conceptual and theoretical scaffold for our own thinking about the topics we are researching. These scaffolds, in turn, are part of dialogical agreements and disputes over vocabularies, traditions, and methods with those that preceded us. A literature review is a very sophisticated way of deliberation.

Bibliographical research is thus connected with the two types of values: it is useful for something and it also has value in itself. Based on this, we can ask a series of questions: should we automate these tasks connected to the most profound parts of our profession and which even lend it meaning and constitute it?

In our experience, Elicit's results aren't currently better than those of Google Scholar, but answering the questions we just posed doesn't depend on whether they are better—one can grant that the quality of these results may improve in the future. Here, the strategic question lies in the academic-scientific practices themselves. What happens to their intrinsic values and ends when bibliographical research is automated? What happens to the meaning and sense of these human practices when instrumental values displace and erode mindsets, norms, and activities that have intrinsic value? And what happens when the intrinsic value obtained from bibliographical research is replaced by an answer from a kind of oracle rather than one coming from those who make up the practice?

Let us imagine the best-case scenario: we obtain a 'hallucination'-free shortlist of the most relevant articles. But for whom are they relevant? Does it make sense to talk about relevance as if it were a neutral, universal notion? If we separate the researcher from this search process, if we separate them from this dialogue with the various

traditions that take place during bibliographical research, what conceptual tools do we have for developing a notion of relevance? What's more, how can we talk about relevance without a subject that imprints meaning and commitment on this stage of research?

To consider which source to include or discard is a crucial part of bibliographical research. It's key to the construction of a hypothesis or theory and we posit that it cannot and should not be fully automated using a chatbot primarily operating on the basis of statistical regularities. Delegating this task to a machine leads to an enormous loss for research itself by voiding of meaning the fraught task of determining the relevance of a source and whether to include it or not.

Should we, then, not automate anything at all? Of course, we can automate and optimize. We automate the generation of a reference list with bibliography managers, and we delegate grammatical checking to the word processor. However, we can defend these automations because they optimize the process without fundamentally eroding its intrinsic values. This critique is not a rant against automation, but rather a call to reflect on what can be automated and what shouldn't be delegated to machines. Does it make sense to automate the generation of hypotheses, bibliographical reviews, the design of experiments, or the discussion of results? If it does, which parts and to what extent? Are we sacrificing intrinsically valuable things in exchange for the instrumental value of efficiency?

The answers to these questions cannot be left in the hands of engineers with reductionist and 'solutionist' views. Nor can they be left in the hands of academic managers interested only in 'quality indicators.' The answers must be given principally by the people committed to the pursuit of the intrinsic values that academic–scientific practices offer. We are well aware that all of this also entails a radical questioning of how we communicate and evaluate scientific production.

To conclude, let's consider agriculture again. Fertilizers and industrial machinery are not mere tools but rather technological devices that shape a type of practice (Berry 2009). Technologies are not neutral instruments but rather ways of putting certain visions into practice while excluding others. Raising calves in cages is at odds with extensive grazing. We believe something similar occurs with chatbots such as Elicit and Perplexity Copilot, which

seem to fit better with the instrumental values of intensive science than with the intrinsic ends of extensive science. Thinking about the eventual adoption of systems like these is a good opportunity to think also about the direction we want our academic–scientific practices to take. Do we want a science of texts (partially) automated by machines that other machines then process and summarize and in which no one reads what we write? Do we want an intensive, instrumental science optimized for quantity and efficiency or, in contrast, do we prefer to imagine and pursue an extensive science that seeks value, meaning, and depth, and which is not just an instrument for something but rather an end in itself?

Acknowledgements A different version of this text (originally written in Spanish) was previously published in a non-indexed, low-circulation magazine published by the Open University of Catalunya, Spain:

Guersenzvaig, Ariel; Sánchez-Monedero, Javier. Nadie escribiendo y nadie leyendo: los generadores de texto con inteligencia artificial y la ciencia que queremos. *Mosaic* [on line], June 2023. ISSN: 1696-3296. DOI: <https://doi.org/https://doi.org/10.7238/m.n199.2309>

Curmudgeon Corner Curmudgeon Corner is a short opinionated column on trends in technology, arts, science and society, commenting on issues of concern to the research community and wider society. Whilst the drive for super-human intelligence promotes potential benefits to wider society, it also raises deep concerns of existential risk, thereby highlighting the need for an ongoing conversation between technology and society. At the core of Curmudgeon concern is the question: What is it to be human in the age of the AI machine? -Editor.

References

- Bender EM, Gebru T, McMillan-Major A, Shmitchell S (2021) On the dangers of stochastic parrots: can language models be too big? In: Proceedings of the 2021 ACM conference on fairness, accountability, and transparency. <https://doi.org/10.1145/3442188.3445922>
- Berry W (2009) The gift of good land: further essays cultural and agricultural. Counterpoint
- Else H (2023) Abstracts written by ChatGPT fool scientists. *Nature* 613(7944):423–423. <https://doi.org/10.1038/d41586-023-00056-7> <https://blog.perplexity.ai/faq/what-is-copilot>. Accessed 23 Dec 2023 <https://elicit.org/faq>. Accessed 1 Aug 2023

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.