

To the editors at *Information Technology and Libraries*:

The Richard Brzustowicz article entitled “From ChatGPT to CatGPT” in your Sept. 2023 issue sparked much discussion in several online cataloging communities, much of it consisting of amazement that such a poorly designed experiment could have made it through the peer review process at *ITAL*.

The structure of the article demonstrates a clear misunderstanding of what generative artificial intelligence [AI] even *is*—from asking the program itself questions regarding its training data to saying that the program follows cataloging rules. Starting from a flawed premise leads to flawed results.

So what *is* generative AI? Generative AI programs such as ChatGPT are language-learning models, a subclass of neural network. Neural networks are machine-learning algorithms whose structures are modeled on the structure of the human brain. They solve problems through trial and error and, with the increasing affordability of cloud computing and processing power, can process vast amounts of training data and draw conclusions from it. This is an extremely useful tool. As Janelle Shane says, “They’re great at matching patterns and finding subtle trends in highly multivariate data. Crucially, they make progress towards their goal even if the programmer doesn’t know how to solve the problem ahead of time.”¹

How does this apply to ChatGPT? Generative text AI are essentially extremely advanced predictive text generators. “ChatGPT is always fundamentally trying ... to produce a ‘reasonable continuation’ of whatever text it’s got so far, where by ‘reasonable’ we mean ‘what one might expect someone to write after seeing what people have written on billions of webpages, etc.’”²

Because it is trained on a lot of natural language materials, it can produce very convincing sentences that seem to carry meaning.

OpenAI’s own FAQ explains it in this way: ChatGPT is called “a language model trained to produce text.” In other words, it “uses human demonstrations and preference comparisons to guide the model toward desired behavior.” “These models were trained on vast amounts of data from the Internet written by humans, including conversations, so the responses it provides may sound human-like.” They warn, however, that, “ChatGPT is not connected to the internet, and it can occasionally produce incorrect answers. It has limited knowledge of world and events after 2021 and may also occasionally produce harmful instructions or biased content,” and “ChatGPT will occasionally make up facts or “hallucinate” outputs. If you find an answer is unrelated, please provide that feedback by using the “Thumbs Down” button.”³

ChatGPT does not think for itself. It is not self-aware and cannot meaningfully answer questions about itself. It also cannot be trained like a human because it doesn’t “understand” anything. It repeats back information and data based on statistical probabilities. Even the people who created ChatGPT, in their own FAQ, warn users that it can give incorrect or made-up answers that are unrelated to the questions and input that a person gives it.

About this Section

Letters to the Editor reflect the opinions of their authors and are not necessarily those of the *ITAL* Editorial Board or ALA’s Core Division. Each letter’s copyright is held by its authors and is published under a Creative Commons CC-BY-NC-4.0 license.

Furthermore, it is baffling that Brzustowicz claimed that ChatGPT was able to generate accurate records when even a cursory glance at the author's own appended data shows multiple mismatches between generated records and existing records.

However, we have taken more than a cursory glance, and we believe that it is important to really dig into all the ways in which ChatGPT fails at even basic cataloging tasks.

Let's take this item by item.

The first example the author puts forth consists of a ChatGPT record and a professional cataloger's record for "the 1996 reprint of *Interview with the Vampire* by Anne Rice using RDA."⁴ This first example contains several critical informational differences between the two records.

Starting at the top, the 020 fields, representing the ISBN of the work, contain different numbers. Searching for the ISBN listed in the ChatGPT record in OCLC Connexion brings up many records from various years, published by Ballantine Books, but no record from 1996. A search for the ISBN from the professional record also brings up many records, but published by Knopf, not Ballantine Books, and there is a 1996 edition featured. Additionally, the ISBN for the ChatGPT record is labeled as a paperback edition, and the ISBN for the professional record is labeled as a hardback edition.

The 040 field, which records the source of the cataloging, in the ChatGPT record features the code DLC, which is the code for the Library of Congress. This is incorrect: this is not a Library of Congress-created record, but one generated by ChatGPT. (This false attribution is common to all of the ChatGPT records save one.) Additionally, the 040 shows that ChatGPT did *not* generate an RDA record, as it is missing the subfield e which would indicate the use of RDA.

Continuing down the record, the 250 field in the professional record holds an edition statement, while the ChatGPT record has no 250 field at all. The author did not describe which edition they were basing their search off of, so it is difficult to proclaim one as correct for the item in hand and the other incorrect, but either way, an edition statement is a core element of a descriptive record, and the difference between these two records is not encouraging.

As can be surmised from the ISBN differences, the publisher featured in the 260 \$b subfield differs between the ChatGPT record, which attributes it to Ballantine Books, and the professional record, which attributes it to Knopf. This alone is deeply worrying; the publisher is such a vital piece of information for identifying which record to apply that ChatGPT's failure to provide an accurate value invalidates the record. Again, we do not know which company published the item the author had in hand, but as there is no 1996 Ballantine edition represented in OCLC at the time of this writing, it is not hard to see the inaccuracy.

Even more glaringly, the 300 field, which contains the extent, contains both a significantly different page count value (372 pages from ChatGPT and 340 pages from the professional record), but also physical size of the book (18 cm versus 22 cm). The 300 field in the ChatGPT record also has a period at the end of the field, which is incorrect (the 300 only ends in a mark of punctuation when there is a series statement, which there is not in either record).

The next significant difference between the two records comes in the subject headings. The ChatGPT record only has two 650 fields (each, interestingly and incorrectly, repeated twice), and they are very basic, only "Vampires \$v Fiction" and "Horror tales." The professional record, on the

other hand, has three distinct subject headings, one including the established 600 for the main character of the work, and a Library of Congress Genre/Form Terms (LCGFT) term, bringing the total subject descriptive fields up to four (double the amount of ChatGPT, and more useful ones).

The records for the album *Low* by David Bowie contain similar discrepancies. The ChatGPT record has no place of publication or publisher, the title contains a deprecated subfield h, and the 336, which represents the content type, is incorrect (“notated music” indicates that the music is written down and “intended to be perceived visually,”⁵ whereas the professional record correctly has “performed music” in the 336, which is correct for an album). Additionally, the professional record contains a track listing, which the ChatGPT record lacks, and more accurate subject headings.

The ChatGPT record for the German translation of Paulo Freire’s *Pedagogy of the Oppressed* in the original article has been compared to a Dutch-language record, not an English-language record, and so we will not attempt to analyze the differences. We will say, however, that the 240 field, meant for the uniform title, in the ChatGPT record reads “Pedagogy of the oppressed. \$l German,” when it should use the original title of the work, which was in Portuguese, and should instead read “Pedagogia do oprimido. \$l German.”

Furthermore, the call number in the 050 field has a second indicator 0, which indicates that it was generated by the Library of Congress, which it was not, and is also incorrect—Paulo Freire’s works are generally classed into LB880.F73, and the classification number used by ChatGPT, LB875, is used for American educators only. There is also an 042 field claiming that the record is a PCC-generated record, which is false.

The record attributed to ChatGPT for Cixin Liu’s *The Three Body Problem* is character-for-character identical to the record attributed to a professional cataloger. One wonders whether there was a copy-paste error, especially given that the text of the article claims that the ChatGPT record *did* have differences from the professional record.⁶

As we do not have a copy of Mood Rings’ “Pathos y lagrimas” in hand, we cannot assess the accuracy of the ChatGPT record. That being said, this record has an OCLC number in the 035 field, which, when searched, has already been assigned to an open-access electronic resource record created in 2018. (The only other example record in the original article with an 035 was also a ChatGPT record, for the German translation of *Pedagogy of the Oppressed*, and the OCLC number was also already assigned elsewhere.)

We have gone into some detail regarding the cataloging fields and specifications in the above analysis. This was a deliberate choice. The author of the original article gives their job title as Instruction and Outreach Librarian. We cannot speak to their cataloging experience, or lack thereof, but anyone trained in cataloging practices would have caught at least some of the above errors in ChatGPT’s output on an initial read.

There is more to cataloging than the look of the record, or the existence of certain fields. Just because a record has a title field, a publication information field, and a handful of subject headings does not make it an accurate record, or useful to the researcher in any meaningful way. Cataloging is a precise, detail-oriented practice, as it must be in order to distill the contents of a work into a single, searchable record.

The original article is careful to mention that review of ChatGPT records is needed before they can be loaded into a library catalog. Rather than starting with a blank slate which can be filled in from

the start by a trained cataloger with the item in hand, the author would have us start with an error-riddled tangle of probability-predicted words and spend our time instead unpicking the mistakes of a generative language model. ChatGPT cannot understand the rules of MARC or RDA, as a human can. All it can do is generate predicted text strings.

Yes, cataloging takes time, care, and attention to do correctly. This does not mean that automation would make the process easier or faster; instead, as we have seen here, attempting to automate the process only slows things down. Rather than proving that ChatGPT is a useful, accurate tool for generating records, the article has instead proven that ChatGPT cannot be a successful replacement for a trained, professional cataloger. We recognize that artificial intelligence is the hot new thing in a wide range of industries, including in librarianship, but it would behoove a respected scholarly journal to do its due diligence, rather than jump on the bandwagon.

Sincerely,

Tess Amram

Special Materials and Continuing Resources Cataloging Librarian

University of Colorado Boulder

tess.amram@colorado.edu

Robin Goodfellow Malamud

Cataloger and Classifier I

Boston Public Library

rmalamud@bpl.org

Cheryl Hollingsworth

Cataloguing Librarian

University of Dallas

chollingsworth@udallas.edu

NOTES

¹ Janelle Shane, "Neural Networks, Explained," *Physics World*, last modified July 9, 2018, <https://physicsworld.com/a/neural-networks-explained/>.

² Stephen Wolfram, "What is ChatGPT Doing ... and Why Does it Work?," Stephen Wolfram Writings, last modified February 14, 2023, <https://writings.stephenwolfram.com/2023/02/what-is-chatgpt-doing-and-why-does-it-work/>.

³ Natalie, "What is ChatGPT?," OpenAI Help Center, accessed October 27, 2023, <https://help.openai.com/en/articles/6783457-what-is-chatgpt>.

⁴ Richard Brzustowicz, "From ChatGPT to CatGPT," *Information Technology and Libraries* 42, no. 3 (September 2023): 2, <https://doi.org/10.5860/ital.v42i3.16295>.

⁵ "RDA Registry | Vocabulary," American Library Association, accessed October 27, 2023, <https://www.rdaregistry.info/termList/RDAContentType/>.

⁶ Brzustowicz, "CatGPT," 3.