

Here come the robot authors

Charles Day

“**L**ithium-ion batteries (Li-ion batteries) have been commonly used as power sources in consumer electronics including laptops, cellular phones, and full and hybrid electric vehicles because of their long cycling life, high energy capacity, and eco-friendliness. Considerable efforts have been devised to examine useful electrode materials for Li-ion batteries with long cycle life and high capacity.”

Thus begins *Lithium-Ion Batteries: A Machine-Generated Summary of Current Research* (2019). The 278-page book is the first from Springer Nature that has a machine, not human, author. Named Beta Writer, the robot author is an algorithm developed by computer scientists and editorial subject-matter experts at Springer Nature and Goethe University Frankfurt in Germany.

In the book's human-written preface, Goethe University's Christian Chiacros and Niko Schenk reveal that Beta Writer combines two subfields of artificial intelligence: natural language processing and machine learning. The raw material for *Lithium-Ion Batteries* consists of 1086 scientific papers. After processing them to index and standardize their content, Beta Writer set to work identifying common themes. The result was an ordered set of chapters, sections, and subsections. Producing the text for the book involved analyzing, simplifying, and synthesizing text from the papers. For example, each chapter's introduction and conclusion were generated from the introductions and conclusions of the papers included in the chapter.

How successful was the enterprise? Whereas each chapter starts with an introduction, the book itself lacks a machine-written preface. My hunch is that providing an overview of an entire field is beyond Beta Writer's current ability, possibly because writing a good overview entails making strong, idiosyncratic choices about what to include and exclude. Also likely beyond Beta Writer's ability: explaining why it wrote the book in the first place. Not surprisingly, given how the chapters were generated, their titles tend to be lists of technical terms, some of which are puzzling or obscure. The fourth chapter bears the title “Models, SOC, Maximum, Time, Cell, Data, Parameters.”

On the other hand, the writing itself, though dull, serves for the most part to achieve one of the book's objectives: To organize those 1086 papers into a coherent, structured, abstracted whole. Indeed, the algorithm's most likely use-case (to use a term from the software industry) is to trawl through literature on a given topic—topological insulators, say—and distill and arrange its main findings. Beta Writer should prove lucrative for Springer Nature.

What should we make of the fact that Beta Writer more or less successfully reviewed a substantial scientific area? When I examined the author guidelines of a handful of journals in

physics and its closely related sciences, I discovered that none of them explicitly require papers to have an introduction. Yet all the papers I encountered had one. You can imagine that papers about cathodes made from lithium manganese oxide (one of the subsections in Beta Writer's book) might all begin with similar introductions that cite the same prior papers. And you might also imagine that each paper reports an improvement of some kind in the cathode's performance. Insofar as the paper is a thick wrapper of words around a set of quantitative results, is it any wonder that Beta Writer could successfully extract results and put them in context?

Beta Writer isn't the only robot science writer in town. Two years ago, a company called sciNote launched an AI tool within its electronic notebook software, ELN. Once a scientist has all of his or her data and lab notes in ELN, the Manuscript Writer tool intervenes to write a report. A visit to sciNote's website revealed this quote from a contented user: “Not only does the new feature generate manuscripts quickly, it also provides several versions that can be used to assemble that perfect publication for your data.”

The formulaic style of modern scientific papers makes the work of Beta Writer and Manuscript Writer possible. The style is efficient because it's predictable. But if 21st-century scientists want to engage readers in addition to informing them, they might consider adopting the personal, conversational style of their 19th-century predecessors. Lord Rayleigh opened his derivation of what became known as Rayleigh scattering with the following lines:

It is now, I believe, generally admitted that the light which we receive from the clear sky is due in one way or another to small suspended particles which divert the light from its regular course. On this point the experiments of Tyndall with precipitated clouds seem quite decisive.¹

How long will it be before robots can write like that?

Reference

1. J. W. Strutt, *London, Edinburgh, Dublin Philos. Mag. J. Sci.* **41**, 107 (1871). 

