Content Generation in the Age of Mechanical Reproduction

Sarah Bull

Book History, Volume 26, Issue 2, Fall 2023, pp. 324-361 (Article)

Published by Johns Hopkins University Press

DOI: https://doi.org/10.1353/bh.2023.a910951

For additional information about this article
https://muse.jhu.edu/article/910951
Content Generation in the Age of Mechanical Reproduction

Sarah Bull

In July 2020, the artificial intelligence research and development company OpenAI released GPT-3, at the time the world’s most powerful natural language generating algorithm, for testing by select developers. GPT-3’s facility for writing in a seemingly endless variety of styles and genres with minimal direction from human operators set off a media frenzy. It wrote an op-ed about itself for the Guardian. It wrote stories about modern love for the New York Times. It was hailed as an incredible breakthrough. But critics worried that GPT-3 was too sophisticated, and not sophisticated enough. Its ability to write in a human-like fashion was very impressive. Its work also made appallingly racist, sexist, and homophobic assumptions, and sometimes displayed a “total lack of common sense” about the extra-digital world. Drawing on one expert’s speculations, Will Douglas Heaven, a senior editor at MIT Technology Review, opined: “Exactly what’s going on inside GPT-3 isn’t clear. But what it seems to be good at is synthesizing text it has found elsewhere on the internet, making it a kind of vast, eclectic scrapbook created from millions and millions of snippets of text that it then glues together in weird and wonderful ways on demand.” In the wrong users’ hands, critics argued, large language models like GPT-3 could rapidly “fill Twitter, email, and the web up with reasonable-sounding, context-appropriate prose” that reproduced factual errors and amplified harmful biases derived from the billions of pages of internet content used to train them to write on command. That prose could prove “impossible to filter” from the internet precisely because it was such a good imitation of human writing.

Concern about these kinds of outcomes intensified from late 2022, as GPT-3 was rapidly integrated into platforms available to the general public, and a series of newer, more powerful language models was released.

The debates I describe reprise themes in nineteenth-century writing about “book-making.” Book-making was once a common label for uninspired
writing that enfolded a theory of mechanical authorship grounded in the idea that the text of books was routinely created by assembling existing materials, just as a codex is made by binding parts together. A fixture of the Anglophone world between the late eighteenth and the early twentieth centuries, it was associated with a range of authorial practices that relied on some form of reuse. None of these practices involved making a straightforward copy of a text. Arguably, they resembled the “regular, law-like” processes of reproduction less than they did looser, more mysterious procreative actions allied with the term’s predecessor, “generation.” Yet, critics framed them as mechanistic and mindless, and expressed concern that their proliferating, uneven output was “impossible to filter” from the print marketplace. With its focus on mediocre literature, writing about bookmaking has never attracted much scholarly interest. This article takes what readers were saying about their media moment seriously. In what follows, I draw on their words and a bibliographical study to explore how something like our concept of content generation came to operate as a framework for understanding textual production, and how the practice that became most associated with it—unattributed excerption—produced epistemic anxieties and informational effects that anticipated issues that we are grappling with today.

The first section of the article argues that the birth of the Author also marked the birth of the content generator. With the elevation of originality, the early modern conception of authorship as a process of skilled reconfiguration was refracted through two fictions of capitalist efficiency: (1) that there were two discrete categories of labor, the intellectual/creative and the bodily/mechanical; and (2) that the materials of production were extractable, fungible resources. While this adaptation upheld a bold new model of literary production—that of inspired authorship—it also produced an opposite model: literary manufacture. The rise of the book-maker as a figure for a mechanistic, market-oriented mode of literary production, whose quintessential product was the compilation, had important ramifications for the subsequent histories of authorship, information and technology. The most crucial for my purposes is that it popularized the idea that all text and anything that could be rendered into text was content—alienable filler for a medium, in this case, the book—and that this material could be generated pretty much automatically. Language-generating algorithms like GPT-3 were as yet unimaginable. Yet, the basic premise that underpins them was manifest in Victorian visions of book-making machines that mixed and mastered old prose to make new on demand.
Nineteenth-century anxieties about literary manufacturing were similar to our own. Readers worried that it spread factual errors, obstructed attempts to verify knowledge, homogenized reading material, and clogged up the print sphere. The second section argues that these anxieties, and the theory of book-making itself, addressed everyday conditions of writing and reading that have largely gone unnoticed. The practice of compiling unattributed excerpts from old texts to write new books—once routine, acknowledged, and accepted—did not decline with the elevation of originality nearly as quickly as scholars tend to assume. It appears to have been common across a wide range of genres in Britain and the United States into the early twentieth century, making text reuse, as Petra McGillen has suggested, even more central to literary production than scholarship on scrapbooking, anthologization, abridgement, reprinting, and “scissors-and-paste” journalism has already demonstrated. As any compiler knows, independent acts of copying add up. Rampant excerption consequently amplified and extended the shelf life of strings of text, creating disorienting echoes in print across oceans, genres, and generations. In the past, these phenomena had been anticipated. Now that the (admittedly tenuous) expectation of originality was set, however, glimmers of recognition undermined readers’ ability to trust that the text of any book was what it appeared to be.

As my outline suggests, this article is less a critical history of natural language generation than it is an attempt to think through how some habits of thinking and feeling now associated with it came into being by reconsidering compiling and its role in the modern history of authorship. The past few decades have witnessed increasing scholarly interest in the role that text reuse has played in literary culture, and scholars have shown that compiling has served a number of important functions: as a means of information management, as a basis for creative labor, as a method of political resistance, as infrastructure for shared culture, and more. At the same time, legal historians have demonstrated that rules for writing now taken for granted in many parts of the world, such as that parts of texts are not common property, were naturalized slowly and unevenly even in Britain, the birthplace of authorial copyright. Yet, scholars often assume that writing by excerpt played a marginal role in literary history after 1800, confined to ephemeral formats and idiosyncratic genres. In delving deeper into the history of “uninspired” writing, this article complicates previous understandings of how information circulated during the nineteenth century and identifies effects of routine copying in this context that deserve closer scrutiny. At the same time, its findings emphasize how much large language models—and our varied reactions to them—owe to major historical shifts.
in thinking about the nature of labor and materials. In the crucible of emergent industrial capitalism, text became a resource and compiling became a technology that made fully automated literary production feel inevitable, though not necessarily desirable.

I: From Craftsman to Machine, From Materials to Content

Washington Irving’s 1819 sketch “The Art of Book-Making” takes us into the reading room of the British Museum, where its narrator, Geoffrey Crayon, encounters rows of authors “in the very act of manufacturing books.” Intently, he watches a man who has all “the appearance of an author on good terms with his bookseller” dip into the room’s “pools of obsolete literature […] taking a morsel out of one [book], a morsel out of another.” “The contents of his book,” Crayon remarks, “seemed to be as heterogeneous as those of the witches’ cauldron in Macbeth. It was here a finger and there a thumb, toe of frog and blind worm’s sting, with his own gossip poured in like ‘baboon’s blood,’ to make the medley ‘slab and good.’”

At first, Crayon situates this author’s pilfering in the tradition of *imitatio*. Like many Romantic poets, he considers authorship something of an ecological affair, relying on cycles of dispersal, ingestion and transformation akin to mysterious natural processes:

We see that Nature has wisely, though whimsically provided for the conveyance of seeds from clime to clime, in the maws of certain birds; so that […] the lawless plunderers of the orchard and the cornfield, are, in fact, Nature’s carriers to disperse and perpetuate her blessings. In like manner, the beauties and fine thoughts of ancient and obsolete authors are caught up by these flights of predatory writers, and cast forth, again to flourish and bear fruit in a remote and distant tract of time. Many of their works, also, undergo a kind of metempsychosis, and spring up under new forms. What was formerly a ponderous history, revives in the shape of a romance—an old legend changes into a modern play—and a sober philosophical treatise furnishes the body for a whole series of bouncing and sparkling essays. Thus it is in the clearing of our American woodlands; where we burn down a forest of stately pines, a progeny of dwarf oaks start up in their place; and we never see the prostrate trunk of a tree mouldering into soil, but it gives birth to a whole tribe of fungi.
But when he falls asleep, lulled by the hush of the reading room, Crayon’s perspective shifts. The “sage magi” before him become a “threadbare throng” who seize parts of books as impoverished men seize cast-off clothes. One book-maker “trimmed himself magnificently from an illuminated manuscript […] and having put Sir Phillip Sidney’s hat on one side of his head, strutted off with an exquisite air of vulgar elegance.” A less convincing book-maker “bolstered himself out bravely with the spoils from several obscure tracts of philosophy, so that he had a very imposing front, but he was lamentably tattered in rear.” Book-making is exposed in Crayon’s dream. He has not witnessed genius at work. He has observed a “literary masquerade.”

In employing the term “book-making” to deflate defenses of excerption, Irving was recycling his own reading material. Outside the literary sphere, book-making referred to the practice of taking bets on sporting events, such as horse racing, and within it, to the material processes of book manufacture. However, it also referred to making a book’s text. For most of the early modern period, book-making was used as a synonym for “writing a book.” Over the course of the eighteenth century, it became associated with a range of more specific authorial practices. By the nineteenth century, as Irving’s story suggests, book-making was primarily associated with the practice of assembling excerpts from old texts to make, or bulk up, new works. It was so tightly linked with this style of composition that dictionaries often defined “book-maker” as a noun for “compiler” or “one who makes books from the writings of others,” and “book-making” as “the act of compiling books” or “the art or practice of compiling books from the writings of others.” It did not matter whether these writings were attributed. Overusing quotations and filling pages with stolen paragraphs both constituted book-making. However, book-making did not necessarily involve such literal forms of reuse. Many people maintained that travelling with the intent of writing about the experience was a form of book-making: it rendered the writing process a simple act of reporting. Others claimed that merely imitative or formulaic writing constituted book-making, as did padding chapters with superfluous descriptions, anecdotes, or sentimental reflections.

In other words, book-making was not a euphemism for plagiarism. Authors were occasionally accused of book-making and plagiarism, but only the latter indicated a moral offense. Book-making was a labor-oriented term: it addressed the practice of making derivative or redundant texts, and implied an uninspired, market-driven sensibility in those who took it up.
This rendered the term something akin to our own “content farming,” a label for cheap processes of media production whose products are—as Tess McNulty has argued of internet “content” more broadly—intended to do “little more than facilitate some profit-driven process.” Critics agreed that book-making did not necessarily create bad reading material. “There is an abundance of sound fact and substantial matter in this workmanlike piece of book-making,” was the Sporting Gazette’s backhanded praise of one 1900 travel narrative. However, they framed it as a lesser mode of writing, one oriented around filling pages in the interest of creating a book to sell, rather than around intellectual or aesthetic goals, such as crafting beautiful stories or sharing new insights. As one writer put it in 1859, book-makers were not “those who have something to write about, and therefore write a book,” but “those who have a book to write, and therefore write about something.” As suggested by the publisher John Blackwood’s epithet for the novelist Elizabeth Gaskell’s biography Life of Charlotte Brontë (1857), “book-making out of the remains of the dead,” book-makers did not write literature. They generated content for the book market by mining what already existed, “pouring [old materials] out of one phial into another for sale.”

The way the term collapsed the literary and the material aspects of book production not only underscored the market orientation of book-making, but aligned it with the repetitive and increasingly mechanized labor involved in creating consumer items. Satires and jokes about book-making habitually emphasized these parallels. An 1837 article in the Satirist called book-making “a trade, like cabinet-making,” for example, while in 1861 the Saturday Review cast book-makers as bakers, churning out a steady flow of books as if from a giant oven. During a period in which the laboring body, the workshop, and the machine were often analogized as “mechanical” entities—each relying not on the powers of the mind, but on habit, instinct, or coordinated action—this view of book-making as an essentially mechanical process inevitably inspired visions of machines that would render it fully automatic (fig. 1). In 1849, a writer for the Athenæum offered up Archibald John Stephens’s history The Book of Common Prayer (1849) as proof that some genius, inspired by Charles Babbage’s designs for calculating machines, had “invented a machine for the composition of books,” a “Book-Making Machine” comprising an “alphabetical compartment” that issued transition phrases, mechanisms for cutting excerpts from a selection of old books, and a “great cauldron” that mixed them all together. Like Babbage’s demonstration pieces, the machine left something to be desired:
its alphabetical compartment was wont to belch out nonsense lines, and the excerpts it combined in its cauldron were often repetitious. Forty years later, in his novel *New Grub Street* (1891), George Gissing anticipated a similar but far more successful device for manufacturing book text through the fantasy of Marian Yule. “Surely before long some Edison would make a true automaton; the problem must be comparatively such a simple one,” Marian reflects. “Only to throw in a given number of old books, and have them reduced, blended, modernised into a single one for to-day’s consumption.” In the absence of such a machine, Marian must resign herself to the role of “machine for reading and writing.”

**Figure 1.** “This novel application of mechanism, to the purposes of periodical publications, is the invention of an ingenious littérateur. The hoppers above being fed with subject of all sorts, from “Criminal Trials” to “Joe Millers,” the handle is turned, and the fountain-pens immediately begin to write articles upon everything.”

It is not a coincidence that the meaning of book-making shifted over the course of the eighteenth century, in tandem with the rise of the notion of authorial genius. As McGillen, who has examined a similar German representation of literary production, points out, authorship was previously framed as a form of craftsmanship. Students in Renaissance classrooms were “encouraged to view all literature as a system of interchangeable fragments” and composition as a process of reconfiguration, centered around “gathering and framing” traditional materials with the help of commonplace books and other paper tools. The techniques of the author were not considered radically different from those of the student. As Martha Woodmansee has demonstrated, he was viewed as a master craftsman, skilled in manipulating the same kinds of materials to achieve effects dictated by his patrons. In cases when an author seemed to achieve something that could not be explained by craft alone, he was said to have been inspired by God, or a muse. Gradually, however, authors sought to establish ownership over their writing, and gain protection of their property through copyright. In doing so, Woodmansee has influentially argued, they “minimized the element of craftsmanship [...] in favor of the element of inspiration, and they internalized the source of that inspiration” in themselves, inaugurating the emergence of the notion of the author as an autonomous individual who produced novel works “inspired by ‘originality,’ ‘sincerity,’ or ‘genius.’”

As Christine Haynes has commented, the “genius” model of authorship became “such a fixed component of the modern consciousness that even the most self-conscious and critical commentators” still find it “difficult to view authorship in any other way.” On its introduction, however, it was dangerously at odds with the material realities of the print marketplace. By definition, genius is limited; literary production is not. In offering, as Jessica DeSpain has suggested, “a counterpoint to authorial genius,” the figure of the book-maker emerged as an explanation for the increasing fecundity of the press, one that bolstered the notion of inspired authorship by offering a foil for it: literary manufacture. The idea that the figures of the genius author and the hack writer came into the world as a pair is not novel. What writing about book-making highlights is that in the process, old ideas about writing were imported and refracted through two increasingly dominant ideas: (1) that there were two discrete categories of labor, intellectual/creative and bodily/mechanical; and (2) that materials of production were extractable, fungible resources. Once respected, the craftsman’s identity was split in two. One part became the Author. The other was collapsed with that of his old antithesis: the unskilled compiler, who would clumsily “lard [his] lean books with the fat of others [sic] works.” A combined identity
emerged: a laborer-machine that spewed texts into the world day in and day out. The craftsman’s materials transformed in tandem. Once hailed as precious fragments of knowledge, painstakingly gleaned through study, they were recast as fungible resources that could be endlessly broken down and recombined to fill new pages bound for the market.

That these fictions informed a new bipartite conception of literary production is hardly surprising: they underwrote the mechanics of Western political economy so completely by the early nineteenth century that it is difficult to identify any form of industry they did not touch. Overtaking early modern theories of “oeconomy” that prized the prudent cultivation, reuse and repair of materials, the myth of the raw resource, subject to ever more efficient procedures of extraction and refinement, drove the expansion of European empires and transformed the very soil of their capitals.44 Labor was not just reorganized to optimize efficiency, but fundamentally reconceived in ways that justified rigid and often oppressive hierarchies of work. The figuration of the body-machine that mindlessly produces helped maintain the economically expedient myth of “the negro [...] [the] dumb beast of burden fit only for slavery” and undergirded the creation of the factory system, with its strictly defined organizations of labor.45 Other forms of mental labor were aligned with the body to accommodate new economic requirements. Once celebrated as the augury of dazzling intelligence, difficult calculations of the kind needed for statistical research were reframed as menial work. Like literary genius, mathematical virtuosity was now proven through “spontaneous, unanalyzable eruptions of the imagination.”46 Calculations were “manufactured” by armies of (increasingly female) computers, for which Babbage sought to create more reliable, more docile metal substitutes.47 As Ranjodh Singh Dhaliwal has recently highlighted, the divisions of labor that offered Babbage metaphors for mechanical computation remain encoded in computational system design: every day, legions of UNIX “slaves” and internet “daemons” perform immeasurable operations so we, the users, can call ourselves productive.48

Unlike the computer, the factory worker, or the slave, the book-maker was not quite (not yet) a class of labor. As some authors hinted when they spoke of the “communistic principles” of authorship, or grumbled, as Ralph Waldo Emerson did, that they did not weave “their web from their own bowels,” the exact dividing line between inspired writing and manufacturing literature was difficult to ascertain.49 What author did not draw from previous works to create new ones? What author did not write, at least partly, for profit? Any writer—every writer—could be a book-maker. Who counted was a matter of interpretation. Alongside the fact that the difficulty
of detecting copied text could render book-making in its most basic sense invisible, as the next section emphasises, the ambiguous boundaries of literary manufacture help explain why the book-maker ultimately served more as a figure for a kind of practice than as a label for a type of worker. Male and female writers were equally likely to be charged with book-making. Wealthy authors and struggling authors, amateur authors and celebrated authors were accused of it. But this should not lead us to underestimate how influential the idea that a great deal of writing was merely mechanical was, nor, the history of copyright suggests, its troubling implications for African-Americans and other racialized groups identified with mechanical labor. Thomas Jefferson’s infamous claim that the poet Phyllis Wheatley did not have the capacity to write work worthy of copyright because, as a Black woman, she was only capable of an imitative, body-centred form of literary production offers an unvarnished example of the racist logic that would influence copyright disputes involving African-Americans over the next two hundred years. White authors could aspire to more than book-making. Black authors could do nothing else.

While the locus of literary manufacture was fuzzy, the view of text enfolded in the concept was not. Text was now content: media matter that “can move across contexts and that comes in discrete chunks like bits or packets,” facilitating economic activity wherever it goes by dint of filling space. Attempts to historicize this way we have of thinking about (and conflating) information, text, images, sound, and data are still in their infancy. Kate Eichhorn considers it a recent development, born in the rise of digital media. Lisa Gitelman and Thomas S. Mullaney have suggested a longer history connected with the introduction of media technologies like the telegraph, which circulated information by remediating it from text to electricity to sound. In my view, the way book-making was represented as a process of manufacturing filler for codices—filler that could be efficiently mined from and moved between various bibliographical contexts—signals the concept’s deeper roots in the modern reconceptualization of materials as resources that historians like Simon Werrett have highlighted. To be sure, earlier writers sometimes represented text as a kind of codex-filler. By the early nineteenth century, though, this had become a habit. In writing about book-making, all text and anything that could possibly be rendered into text is chronically cast as a “resource” for the hungry machine of the publishing industry. Even works of “genius,” which seem to stand above the market as inspired wholes, are ripe to be carved up, reconfigured, reused or imitated to fill up new pages.
I want to propose, then, that the emergence of the Author also marked the emergence of the content generator. Literary production was never truly mechanized during the nineteenth century, of course. Even John Clark’s *Eureka* (1843), a permutation device that eventually succeeded at grinding out Latin hexameters at a rate of one per minute, was seen more as a curiosity than as a viable replacement for human labor. “We fear we cannot recommend it to despairing school-boys,” one journalist sighed. But as a theory of literary production, it existed as much as inspired authorship did. The book-maker’s emergence signalled the crystallization of a premise that would eventually undergird much of the digital economy and make large language models possible: that text is a resource that can be mechanically mined and remade to serve various market purposes.

Nineteenth-century readers took this premise for granted but rarely with much enthusiasm. The book-maker was an object of disappointment and often of derision, a burlesque of the publishing industry’s insatiable capacity for product. With the rise of the white-collar office at the century’s twilight, however, as Marian Yule’s reverie suggests, content generation started to look more like a fantasy. For men, it would almost be manifested in a new kind of worker, the “typewriter,” who filled in and tidied up prose for her employer behind the scenes. While they have a range of applications, the development of large language models—and certainly of platforms to exploit them—has been substantially motivated by this gendered fantasy of a labor-saving prosthesis to assist, and even replace, workers tasked with churning out ever-greater supplies of content. Platforms like Copy.ai and Jasper work up marketing copy, mission statements, and other formulaic business documents to fill corporate websites, while others digest financial reports and sports scores into news summaries. Still others facilitate a managerial (or perhaps more accurately, seminal) approach to creative success, optimizing the production of what Eichhorn calls “entertainment content” for platforms like Amazon Kindle Direct Publishing. “Start with a story seed and let the AI take the wheel,” advises one devotee of SudoWrite, a GPT-3-driven fiction-writing tool.

II: Compiling Compilations: Book-Making In Aggregate

Why did compiling, a practice typically associated with older literary traditions, dominate conceptions of literary manufacture even at the end of the
nineteenth century? And why were readers so obsessed with it? References to book-making pop up in every place one might possibly mention a book, from personal correspondence, book reviews, and reader’s reports to novels, jokes, and songs. The earlier tradition of thinking of authorial materials as fragments and fact that books are compiled sheet by sheet offer plausible answers to the first question. Nineteenth-century readers’ well-documented ambivalence about the expansion of the print marketplace offers a very reasonable answer to the second. For every social improver who hailed the advent of cheap print as a boon to public education, there was a reader who looked at “the overgrown jungle of ‘information’” it had wrought with undisguised dismay.\(^6\) But there was an additional, and I would argue especially important, reason why readers imagined book-making in the way that they did, and why they wrote so much about it: they believed that their print world was flooded with compilations.

Scholars now recognize that the print market relied heavily on reprinted and reworked material. Editors filled newspapers with articles sourced from other periodicals; gathered and stitched serialized stories together to make volume editions; cherry-picked the most beautiful parts of old works to create new anthologies and gift books; abridged them; and reprinted them wholesale.\(^6\) But readers insisted that this reliance on old text went deeper. Throughout the century, they claimed that the marketplace was deluged with new books composed of “disguised excerpts”: biographies, travel narratives, histories, sermons, schoolbooks, medical and scientific works, and even novels were routinely created with scissors and paste.\(^6\) An 1838 article in *Bentley’s Miscellany*, posing as a guide to would-be book-makers, pummeled its readers with example after example:

At sunset, these industrious creatures (like homeward-bound bees) return, laden with the sweets of centuries, to their garrets, to toil through the night at the work of reproduction. [...] Your “petit Literateur” is [...] essentially, a beast of prey—he is, moreover, a gross feeder, and decidedly omnivorous. A grave metaphysician of this class will occasionally plunder a thought from Byron’s “Cain,” and a hungry small-beer poet will sometimes snatch a mouthful from the “Philosopher of Manbury.” I have known a whig-doctor to quietly appropriate a leading article from the “Standard,” and a tory parson lay violent hands upon a whole chapter from Jeremy Bentham [...] The miniature book of science [...] is principally “worked up” out of old materials with the assistance of the scissors [...]\(^6\)
Although they sometimes framed text reuse as text theft (“plunder,” “snatch”), the authors of such articles were not “plagiarism hunters.” The focus of their concern about this “system” of literary production, as they conceived it, was not its singular effect on authors, but its cumulative effect on readers. In the estimation of many of these critics, unattributed excerption was a public nuisance: it clogged up the circuits of communication by flooding the market with redundant and sometimes nonsensical books, making an already bursting print sphere even more difficult to navigate.

Others worried that it destabilized the integrity of knowledge by introducing or reproducing errors (side effects of hasty and indiscriminate copying) and obstructing attempts to set knowledge within genealogies. As John Aikin and William Johnston diplomatically put it in their *General Biography* (1801),

> It is far from our intention to speak slightly or degradingly of this species of literary industry, by which the most interesting parts of human knowledge have been held forth and communicated to thousands who might else have remained without instruction. [...] [But] the editors dishonestly pretend to originality, which does not belong to them; and to support that pretence, they conceal the sources from which they have copied. This practice renders the history of facts, of inventions, and the moral dependence upon the narrators, confused and inextricable.

That book-making could be done well thus presented as much of a problem as the fact that it could be done badly. Anthologists and other more respectable renovators of old texts aimed to bring them into conversation with the present and sometimes radically edited them for this purpose. Yet, as Tom Mole has emphasized, they never completely unmoored them from their original contexts. For Victorian readers, the works of Shelley and Byron felt modern when they were tamed into digestible, apolitical lyrics or presented alongside illustrations of Don Juan dressed in the latest fashions. Even so, they remained tightly associated with their authors, with literature, with Britain, and with the recent past. Disguised excerption deliberately occluded textual history, papering over fissures between excerpts from multiple sources to present them as a single entity. Readers could never be certain of knowing where information was coming from when they read a book, or how old it was. In a world in which print production was rapidly expanding, even learned readers could neither recognize nor trace the origins of reused passages reliably.
Of course, the fact that readers like Aikin and Johnston thought that they knew they couldn’t know raises the question of what they actually knew. Victorian book reviewers, who, as Elizabeth Carolyn Miller has noted, often represented texts as “arrangements of materials,” frequently made brief, evaluative references to book-making that give the impression that they read for signs of writing by excerpt, always scanning for fissures in the text:

It is seldom, indeed, that we meet with a work so thoroughly honest in purpose and so interesting in material as these two volumes, not the least grateful feature of which is the absence from them of everything in the shape of book-making.

We are not inclined to rank it so high as the more spontaneous of his efforts, and there is a considerable flavour of book-making about it.

It is thoroughly substantial — quot verba tot pondera — and has not the faintest trace of book-making throughout.

Notably, however, phrases such as “the flavour of book-making,” “the shape of book-making,” and “the trace of book-making” do not communicate facts. They communicate intuition. Did compiling’s fall from grace produce generations of paranoid readers, always on the lookout for counterfeit authorship? Or were nineteenth-century readers trying to negotiate an issue that we are confronting today, as AI-generated texts and images, “deepfake” video, and other synthetic media become ever more sophisticated: a media environment in which what they consumed was not always, or even often, what it appeared to be?

The need to develop a more robust understanding of unattributed excerpt’s role in modern culture has become clear over the past decade, as the digitization of historical material has facilitated increasing discoveries of text reuse in nineteenth-century books, often through serendipity. These discoveries have begun to alter scholarly understandings of books and the people who made them in uneven and sometimes troubled ways. Responses to Geoffrey Sanborn’s findings in the course of his scholarship on the African-American abolitionist William Wells Brown, who by the end of his career had, in Sanborn’s words, “plagiarized at least 87,000 words from at least 272 texts” in speeches, histories, an autobiography, and a novel, Clotel (1853), offer a trenchant example.

Brown, who winkingly lifted a chunk from Irving’s “The Art of Book-Making” in his own memoir, The
American Fugitive in Europe (1855), seems to have embraced the unusual and potentially defiant identity of virtuoso book-maker. However, he has long been considered a “plagiarist” in scholarly circles. Sometime in the 2010s, Sanborn set out to identify all of Brown’s sources. In the process, he discovered that three other writers, William Hunt, Wilson Armistead, and Pauline Hopkins, had copied phrases from Brown or a common source text without attribution. Further research revealed that Hopkins, like Brown, relied heavily on excerption; more than 20 percent of her novel Of One Blood (1902–3), for instance, comprises “passages imported from other people’s publications,” as does at least 25 percent of her novel Winona, A Tale of Negro Life in the South and Southwest (1902).

While scholars had known that Hopkins, a pathbreaking figure in African-American literary history, appropriated some text in her novels, revelations of the extent to which she did so were explosive. Against the backdrop of a history in which Black creators have often been denigrated as imitators or thieves, literary scholars wrestled to reconcile their desire to defend Hopkins with their bafflement as to why she would deliberately engage in practices that they believed “would have been understood as an illegitimate appropriation.” However, just as writing about book-making strengthens doubt already raised in scholarship on copyright that Hopkins’s technique would have universally been seen as plagiarism, situating these discoveries alongside other recent work on text reuse offers compelling evidence that for much of Hopkins’s lifetime “unacknowledged borrowing” on this scale was simply not unusual.

Drawing on a series of author’s notebooks, McGillen has shown how, with the help of amanuenses, the celebrated late-nineteenth century German novelist Theodor Fontane amassed a vast archive of snippets of newspaper reports, journal articles, letters, historical documents, and monographs. When he sat down to write, he would “survey these textual building blocks, outline rearrangements with the help of lists, and combine them into a new text.” By digitally comparing the text of novels in the Wright American Fiction Collection with the text of newspapers in the Library of Congress’s Chronicling America database, Ryan Cordell and Avery Blankenship have further demonstrated that white American authors “regularly borrowed snippets of prose from the newspaper for use in their fiction.”

Other scholars have detected similar patterns of text reuse in nonfiction on both sides of the Atlantic. Maria Damkjær has shown that Isabella Beeton’s infamous plundering in her best-selling Book of Household Management (1861) was commonplace. Authors of domestic management manuals
and cookbooks routinely created them by compiling excerpts from other sources. Many used the same excerpts as Beeton.\textsuperscript{81} Other scholars have drawn attention to extensive text reuse in travel narratives, biographies, and textbooks, and Jeff Loveland has demonstrated that up until the mid-nineteenth century, most of the content in encyclopaedias was recycled from other sources (mainly other encyclopaedias).\textsuperscript{82} Taken together, this growing body of literature suggests that nineteenth-century readers were right: unattributed excerption was far more central to literary production than we tend to assume. Relative to the size of the print market, however, it does not offer a lot of information about how this style of composition was practiced, and what information it does offer is uneven. In general, just \textit{how} common was making books out of excerpts from others? How, and how quickly, did it become uncommon? Was it performed differently in different genres? Moreover, if it was routine within or across genres, did it have aggregate aesthetic effects or affect the circulation of knowledge in potentially deleterious ways, as some readers suggested?

These are big questions that cannot be answered satisfactorily in a single journal article, nor, I suspect, by a single scholar. Below I propose some tentative answers, drawing on a computer-assisted study of text reuse in relation to four works first published in Britain: the 1840 and 1852 editions of J.L. Curtis & Co’s sexual health manual \textit{Manhood: The Causes of its Premature Decline with Directions for its Perfect Restoration}; the obstetrician Michael Ryan’s sweeping \textit{The Philosophy of Marriage, in its Social, Moral, and Physical Relations} (1837); and Lady Augusta Hamilton’s equally ambitious \textit{Marriage Rites, Customs, and Ceremonies of all the Nations of the Universe} (1822).\textsuperscript{83} The method that I used is labor-intensive and involves a lot of careful reading. However, it requires neither specialized technical skills nor access to paywalled collections of digitized material. I copied and pasted text from each book into the Google Books search bar phrase by phrase twice, first with and then without quotation marks around each phrase. After I keyed in each phrase, I scanned Google Books for hits and manually compared the text of the work I was studying with the text of works that contained potential matches. I then recorded verbatim or near-verbatim text matches in a spreadsheet alongside a copy of the original text. I did not record attributed copying (e.g. quotations or excerpts from books published in periodicals or miscellanies) unless it indicated the appearance of the text in a work that is not wholly reproduced in Google Books, or was the first known appearance of the text.
The results of computer-assisted studies of text reuse are always speculative. The total publishing output of the world will never be digitized. Google Books contains only a portion of the published works that have been digitized. Comparing the text of a single work with the text of every work in Google Books can therefore only tell us so much about text reuse. It cannot tell us how much of the work was copied or adapted from other sources, nor—this will become very clear—can it tell us what works its author derived their material from. It is also important to recognize that although search algorithms can identify overlaps between texts more efficiently than human readers, they often fail, partly as a result of errors introduced into the text during the digitization process.\textsuperscript{84} Even with these drawbacks, these kinds of studies are helpful because they enable the creation what Cordell calls “speculative bibliographies,” which “model a particular idea of textual relation or interaction” by surfacing textual relationships between works.\textsuperscript{85} These bibliographies can be used to develop theories about textual relationships that can be tested and expanded on through further research. Although I cannot claim that it is representative of the literary field, the speculative bibliography that I created of textual relationships surrounding these four works provides further evidence that writing books by excerpt was common across a range of genres during the nineteenth century. It also offers evidence that in the aggregate, text reuse did have effects that deserve further study.

All of the works that I examined were substantially derived from previously published material. \textit{Manhood} presents itself as a “popular treatise” written and published by J.L. Curtis, the pseudonym of a medical practitioner based at 15 Albemarle Street in London’s West End. In fact, the bulk of the first English edition was derived from the French physician Leopold Deslandes’s 1835 treatise, \textit{De l’onanisme et des autres abus veneriens} (Onanism and other venereal abuses).\textsuperscript{86} \textit{Manhood}‘s author seems to have customized material translated from Deslandes’s work by incorporating verbatim or near-verbatim excerpts from at least seven English-language works into the text. For example, commentary that appears in \textit{Onania; or the Heinous Sin of Self-Pollution} (c. 1712), a screed on the dangers of masturbation, appears in \textit{Manhood}‘s introduction and contributes to the work’s moral frame.\textsuperscript{87} A review of physicians’ opinions of the perils of uncontrolled sexual indulgence is a verbatim match for the surgeon F.B. Courtenay’s in his \textit{Practical Essay on the Debilities of the Generative System} (1839).\textsuperscript{88} Most of \textit{Manhood}‘s sixteen-page chapter on the causes, symptoms, and treatment of gonorrhoea matches text in a chapter in the surgeon William Lawrence’s
Lectures on Surgery (1830). Here, Manhood’s author seems to have lightly adapted the source text, which was aimed at medical students, for a new audience. Where Lawrence instructs surgeons how to relieve painful symptoms of gonorrhoea:

In the first place, you may get a basin of cold water, and let the patient, with a sponge or piece of lint, bathe the part so as to cool it as much as possible. Then you press gently upon the swollen glans with the thumb, or thumb and finger of one hand, while you gradually draw over it the contracted orifice of the prepuce with the thumb and finger of the other hand. If you proceed slowly, squeezing out the blood from the glans as well as you can, so as to reduce its size, then managing to push it gradually into the opening of the orifice, at the same time that you draw the prepuce gently forward, you will usually succeed in replacing the parts, and thus relieve the patient from a state which to him is one of considerable alarm and apprehension, besides being very painful.

Manhood explains the procedure to prospective patients:

What we would advise in this case is, to get a basin of cold water, and let the patient, with a sponge or piece of lint, bathe the part so as to cool it. Then the surgeon is to press gently with the thumb the swollen glans, or with the thumb and finger of one hand, whilst he gradually draws over it the contracted orifice of the prepuce with the thumb and finger of the other. He is to proceed slowly, squeezing out the blood from the glans, as well as he can, so as to reduce its size, then managing to push it gradually into the opening of the orifice, at the same time that he draws the prepuce gently forward, and he will then succeed in replacing the parts, and in relieving the patient from a very distressing and painful situation.

The expanded 1852 edition of Manhood incorporated more unattributed material. Text in this edition matches verbatim or near-verbatim with text in twenty-one separate previously published sources, including Ryan’s Marriage, and two sources published in the same year. Superficially, Ryan’s Marriage is a very different kind of book. It is longer and was three times as expensive as Manhood, written by a well-known obstetrician and medical journal editor, and issued by the eminent medical and scientific publisher John Churchill. Much of the text of Marriage also announces itself as compiled material through devices such as quotation
marks and paragraph beginnings such as “Pliny says —.” I did not count this material as copied for my purposes nor search for it in Google Books. However, many of the portions of Marriage that are presented as original are not: the text matches verbatim or near-verbatim chunks of text, ranging from paragraphs to pages, in nine books and articles issued prior to Marriage’s publication, from Timothy Dwight’s Theology: Explained and Defended, in a Series of Sermons (1825) to Hamilton’s Marriage Rites. Hamilton’s Marriage Rites, a proto-anthropological world tour first published in 1822, is similar to Ryan’s Marriage in that it contains portions of quoted or otherwise attributed text. However, text presented as original in the introduction and first chapter alone matches verbatim or near-verbatim with text in twenty-three previously published works, including Elias Habesci’s The Present State of the Ottoman Empire (1784), François, Baron de Tott’s 1786 Memoirs, and a plethora of encyclopedia entries. Some of these text fragments are just a line or two long. These may be survivors from source texts that were subjected to intensive editing. In some cases, however, whole paragraphs and or pages of Marriage Rites match verbatim or near-verbatim with a previously published source.

What is more significant for the purposes of this article is that these authors were not the only ones taking scissors to old works. As the visual representation of textual relationships in figure 2 suggests, their books can be thought of as nodes in a network of books that shared verbatim or near-verbatim chunks of text. Before the first edition of Manhood was published in 1840, parts of its sources had already been reused in other publications. The authors of some of these sources, including Lawrence, silently reused their own writing, recycling sections from published lectures and journal articles in their books, and often again in other books and articles. However, most of these figures, and a number of others, also recycled the work of other authors without attribution. The physician Wooster Beach’s American Practice of Medicine (1833) shares pages of text with Astley Cooper’s Surgical Lectures, Robert Thomas’s The Modern Practice of Physic (6th ed, 1819), and Samuel Cooper’s Dictionary of Practical Surgery (4th ed, 1822), including some of the same passages as the 1840 Manhood. The authors of two competing manuals, Brodie’s Medical Work on Virility (1844) and The Silent Friend (1847) lifted text that appeared in Ryan’s Marriage before the 1852 Manhood, which contains the same text, was issued. These knots of textual relationships make it impossible to determine exactly what the sources of either edition of Manhood actually were. They may not be any of the works I have suggested, but others that have not been digitized and uploaded to Google Books.
The silent copying did not stop after the publication of either version of the manual. At least twenty-four authors lifted text from *Manhood*‘s sources, from *Manhood* itself, from works that had lifted from one or the other, from both, or the sources of *Manhood*‘s sources for decades after *Manhood* was first published in 1840. In some cases, the same chunks of text appeared in new publications repeatedly. As the appendix to this article shows, for example, *Dr. Bates’ True Marriage Guide*, published in Boston in 1889, includes a paragraph on impotence that seems to have originated in the surgeon John Roberton’s *Diseases of the Generative System*, which was first published London in 1811. It was reused without attribution, with minor editorial changes, in at least six other books, on both sides of the Atlantic, over a span of at least seventy-two years. *Manhood* was merely one staging post in the excerpt’s circulation.

These patterns repeated when I examined Ryan’s *Marriage* and then Hamilton’s *Marriage Rites*. Most of the sources that they drew on had been copied before, and either their books, the sources they used, the sources of their sources, or sources that copied from them were copied by other authors for years, and sometimes for many decades, after they published them. A description of a Turkish harem in *Marriage Rites*, for instance, had appeared in at least four other works before *Marriage Rites* was published, in productions ranging from Elizabeth Craven’s *A Journey Through the Crimea to Constantinople* (1789) to an anonymous pamphlet called *The

---

**Figure 2.** This graph, created using the open-source software package Gephi, visualizes data discussed in the article. The dots, or nodes, indicate books, and the lines, or edges, between the nodes indicate shared text. The size of the nodes refers to the number of edges they share with other nodes in the network.

**Credit:** Figure by author.
Surprising Savage Girl, who was Caught Wild in the Woods of Champagne (1821). And after Marriage Rites was published, its discussion of Tahitian polyandry appeared again in John Armstrong’s The Young Woman’s Guide to Virtue, Economy, and Happiness (1826), Theophilus Moore’s Marriage Customs and Modes of Courtship of the Various Nations of the Universe (1830), and Michael Ryan’s 1837 Marriage. As with Manhood, Marriage Rites rarely appears to have been the originator of an excerpt or the stopping point in its circulation. A description of a seraglio in Constantinople, for instance, was published without attribution in at least fifteen separate works between 1738 and 1830. Lady Hamilton was just one of many authors who put it to use.

Ryan seems to have made use of newer material than Hamilton. His discussion of the duties and obligations of marriage is typical. It had already been published, but the earliest sources I could find were issued only a decade before Marriage: the Reverend H.C. O’Donoghue’s Marriage: the Source, Stability, and Perfection of Social Happiness and Duty (1828) and a prize essay in an 1827 issue of the Ladies’ Monthly Museum. However, material associated with Ryan’s Marriage was reused for much longer than material associated with Hamilton’s Marriage Rites. While few authors seem to have lifted material associated with Hamilton’s introduction and first chapter after the 1830s, when the popularity of tales about life in Turkey faded in the Anglophone print sphere, text from Ryan’s Marriage or its sources was copied regularly for decades after its publication, appearing without attribution in at least thirty works issued on both sides of the Atlantic into the twentieth century. Readers of B.G. Jefferis and J.L. Nichols’ Searchlights on Health: The Science of Eugenics, published in Illinois in 1921, were presented with a passage about German fertility that had previously appeared in works on homeopathy, physiology, reproduction, marriage, and masturbation in Britain. This material was at least ninety years old at the time Searchlights was issued.

In all three cases, unattributed excerption brought chunks of text across time, space, and genre, quietly transporting information and perspectives developed in a particular cultural context for a particular kind of readership into others. Clear misattributions in other works affirm my impression that few of these excerpts were commonplaces—that is, text fragments that most readers would have easily recognized as the work of another, older author. Many excerpts appear to have been chosen precisely because they were so generic that they required little to no adapting. Most of them are descriptions—of the workings of the body, the symptoms of ailments, the
features of plants, the beauty of faraway places, the appearance and behavior of foreign peoples, and so forth—that could be put to work in a variety of literary contexts. Like Hamilton, Ryan, and “Curtis,” the authors of a number of the works I examined appear to have shaped them by gathering and arranging less generic material, often quoted or translated from foreign-language sources. They then fleshed the text out by arranging unattributed excerpts around that material to add color, depth, and detail; and then they often added to and bridged those excerpts with original writing or more extensively adapted material from the same or other authors. Such authors were not passing off old writing as their own so much as they were using old writing as a repository from which to build a new text.

But while each compiler selected excerpts independently, regular excerpt compounding information: the more times an excerpt was reused, the more likely it was to be reused again. Judging from this study, the maximum shelf life of an excerpt was around ninety years. After that, excerpts were either edited substantially before they were reused in new works, no longer appeared without attribution, or disappeared entirely. However, this finding may be an effect of the limitations of using Google Books to track text reuse: copies of nineteenth-century books are overrepresented in Google Books compared to those published prior to 1800, and few works published after 1924 are displayed in full text due to copyright concerns. In any case, the fact that a number of different works often shared the same text suggests that writing by excerpt had, at the very least, a weak cumulative effect. As well as encouraging misattributions, potentially replicating factual errors, and facilitating the use of information in contexts very different than those originally intended (as in Searchlights’ eugenicist cast on old writing about German marriages), regular excerption amplified representations, creating uncanny redundancies and echoes in print that readers sometimes noticed—not just in works on the same subjects published at roughly the same times in roughly the same places, but also across borders, genres and generations. The same old descriptions appeared over and over again in new books. Authors (let alone readers) were almost certainly unaware of how old such descriptions were or the original context in which they appeared.

In one case, writing by excerpt had a stronger, more localized aggregate effect. Incestuous excerption spread the same text around and around a single genre over about fifteen years. Ultimately, this homogenized the genre. Both editions of Manhood and most of its competitors, including R.L. Perry and Co’s The Silent Friend, C.J. Lucas and Co’s Manly Vigour, R.J. Brodie and Co’s Brodie’s Medical Work on Virility, Walter de Roos’s Silent
Hints on Certain Subjects, and Samuel La’Mert’s Self Preservation share large amounts of text. The 1852 edition of Manhood and the 1844 edition of Brodie’s Medical Work alone share over four thousand words of verbatim text, around 10 percent of each manual. It appears that the authors of these cheap sexual health manuals copied from their competitors as well as from other sources when they created them in the first place, and when they updated them for frequent new editions. For example, the 1840 edition of Manhood contains a paragraph that seems to have originated in Onania; or, the Heinous Sin of Self-Pollution. The 1847 edition of The Silent Friend contains the exact same paragraph, and follows it with a paragraph that matches one in Roberton’s 1811 Generative System, which does not appear in the 1840 Manhood. The 1852 edition of Manhood reproduces the exact same 950-word, two-paragraph block of text containing material from both Onania and Generative System. As in this example, exchanges of text between these manuals often did not involve more than a few hundred words. However, they accumulated so much over time that they seem to have become obvious to readers. By the 1850s, readers were complaining that these manuals were all the same.

This was a very small study. I examined text associated with only four works, two of which are different editions of the same work. Hamilton’s Marriage Rites is also so lengthy and shares text with so many other works that I was only able to systematically examine the text of its introduction and first chapter. However, the fact that this tiny body of literature shares unattributed text with at least 125 other works, and that many of those works share unattributed text with each other, is telling in my view. In combination with the work of other scholars, and the words of readers themselves, it suggests that writing by excerpt was common into the early twentieth century across a range of genres, including histories, sermons, geographies, medical works, advice manuals, travel narratives, biographies, cookbooks, and, potentially, novels. It also suggests a few narrower trends. For most nonfiction, writing by excerpt does seem to have declined over time, especially from around 1870. It seems to have been increasingly, though not exclusively, associated with cheap and middlebrow works. And the ways writing by excerpt was performed seem to have been influenced by generic demands. When making nonfiction, authors used roughly similar techniques, often lifting large chunks of text from their sources almost verbatim. When making fiction, authors usually reworked material more extensively, often interleaving borrowed and original text, or subjecting excerpts to multiple rounds of rearranging, condensing, and expanding to arrive “at the perfect blend.”
Observed in concert, these practices make the notion that new books were being “manufactured” out of old text seem perfectly natural in this cultural context. They also help explain certain idiosyncrasies in writing about bookmaking, such as the fact that exposés that described how and where writing by excerpt was practiced circulated alongside satires that treated the general idea as common knowledge. Readers had a sense that many books were pieced together, but they could not necessarily tell which ones or how many. Both kinds of writing addressed this condition in different ways. The degree to which the practices I have described affected genres or knowledge is less certain. Rampant excerption homogenized the text of a particular genre of sexual health manual. To what extent could this have been true of other genres? Did the regular copying of descriptions of people, places, and things fix ideas about race, gender, culture, or nation, or cause them to change more slowly than they might have if authors had written the descriptions themselves? What, if any, were these practices’ effects on “fact”? Fiction writers mined nonfiction for excerpts. Did nonfiction writers frame fictional passages as nonfiction? Cordell and Blankenship’s scholarship suggests that this could easily have happened by accident. Fictional excerpts were frequently reprinted in newspapers without attribution and were often unidentifiable as fiction in this context.

To answer these and other questions I have raised, we will have to study text reuse involving books in more systematic ways. Scholarship by Katherine Bode, Cordell, Blankenship, and David Smith, Robert Morrissey and Min Chen, and others has laid the groundwork for enquiry along these lines using more sophisticated computational methods, while the work of scholars like McGillen has demonstrated the value of careful qualitative studies. As Cordell and Blankenship’s cross-media study emphasizes, both kinds of scholarship would need to account for the fact that books were part of a larger media ecology in which chunks of text moved in and out of different media contexts, sometimes attributed, sometimes not. Although I have focused on textual relationships between books here, many of the authors I studied may not have copied from other books but from excerpts in periodicals. Such work would also need to acknowledge the fact that the practices I have described were not unique to Anglophone print culture, though they may have been understood in different ways elsewhere. If this work is done, it will take time. In the meantime, I think that one thing is clear: those of us who study nineteenth-century books for a living cannot assume that textual originality was the norm.
Conclusion

From certain standpoints, aligning the work of nineteenth-century compilers with the work that large language models perform is misleading. It is worth emphasizing that although they were often cast in the role, compilers were not machines. They were individuals who made new works out of old texts for many reasons, some creative, some political, some personal, and yes, some mercenary. It is also worth emphasizing that compilers worked with text very differently than language models do. While compilers copied, arranged, and edited texts to varying degrees, language models generate texts through predictive modelling, relying on complex statistical models developed through exposure to many texts. As a result, their products are distinct, too. A compilation will share verbatim or near-verbatim text with its sources, while the texts that language models produce are—for the most part—“original.”

And yet, conversations about natural-language generation today and conversations about book-making in the nineteenth century look similar, at least right now. I have suggested that one reason for this is that in these different historical moments, unattributed excerption and natural-language generation have (already, despite the latter’s relative novelty) had some similar effects. While neither process exactly replicates what has been written before, both recirculate old ideas encoded in old texts. In both cases, it has been difficult for people to identify which old texts inform works that have been produced using these processes and whether these processes have been used to create a work at all. Both have consequently fostered the sense that one cannot quite trust the text of any work to be as it appears—a reprisal of a sentiment that Adrian Johns has argued was fostered by the introduction of the printing press in early modern Europe, given its capacity to reproduce piracies and forgeries. Fixing knowledge in a medium and making it trustworthy has, apparently, never been easy. The history of book-making, such as it, is, emphasizes that these problems have only ever been temporarily resolved. It also emphasizes that they are not solely connected with the introduction of new technologies. What counts as authoritative changes over time. A trusted way of circulating knowledge in one moment can become untrustworthy in the next.

Conversations about book-making also feel familiar because our current predicament has some deep historical roots. In the eighteenth and nineteenth centuries, broad shifts in thinking about labor and materials facilitated the rise of a view of authorship that soldered it with originality and thereby jus-
tified authorial copyright. At the same time, I have argued, these shifts made it possible to imagine automating literary production by reifying text. Even as authorship was recast as an almost god-like enterprise, text was increasingly thought of as a resource that could be endlessly mined, remade, and recirculated to sustain flows of capital. Unattributed excerption does seem to have eventually declined amid the expansion of copyright and publishers’ increasing efforts to define, and condemn, the sin of plagiarism. However, these models of literary production have only become more fixed. As this article goes to press, several authors have launched lawsuits against technology companies for using their books to train language models, and more are likely on the way. Though they are not unprecedented, the ruptures to order represented by language models are different from those we have experienced in the past. The history of book-making helps us see that they are a logical outcome of the capitalist transformations of the eighteenth and nineteenth centuries and their internal contradictions.

Appendix

Many gentlemen, turned of forty years of age, and sometimes even earlier, who have lived rather freely, are not unfrequently, about that period, greatly altered in their powers of sexual intercourse. They may, indeed, in general health, be stout, and, for several years, not very sensible of their powers degenerating, but the frequency of their inclination, for such duties, gradually becomes much less, and that symptom is, at all times portentous of approaching impotence; for the inclination gradually and entirely ceasing, the power soon follows. In others, about the same time of life, the physical power ceases first (which occasions a most farcical catastrophe), and the inclination continuing, often for many years after, they are obliged to gratify themselves in amusements which are the mere pantomime of amorous indulgences!—such individuals, otherwise in tolerable health, are recoverable.


Many individuals scarcely turned of forty years of age, and sometimes even earlier, who have lived rather freely, are not unfrequently, about that period, greatly altered in their powers of sexual intercourse. They may, indeed, in general health, be stout, and, for several years, not very sensible of their powers degenerating, but the frequency of their inclination for such duties, gradually becomes much less, and that symptom is at all times portentous of approaching impotence; for the inclination gradually and entirely ceasing, the power soon follows. In others, about the same time of life, the physical power ceases first (which occasions a most farcical catas-
trophe,) and the inclination continuing, often for many years after, they are obliged to gratify themselves in amusements which are the mere pantomime of amorous indulgences!—such individuals, otherwise in tolerable health, are recoverable.


Many individuals scarcely turned of forty years of age, and sometimes even earlier, who have lived rather freely, are not unfrequently about that period of life, greatly changed in their powers of sexual intercourse. They may, indeed, in general health and personal appearance, be stout, and, for several years, not very sensible of the degeneration of their powers, but the frequency of their inclination for such duties gradually becomes much diminished, and that is a symptom which is at all times indicative of approaching impotence; for the inclination gradually and entirely ceasing, the power speedily follows, or rather both are lost together. In others, about the same time of life, the physical power ceases first, and the inclination continuing, often for many years after, they are obliged to gratify themselves in amusements which are the mere pantomime of amorous indulgences! such individuals, being otherwise in tolerable health, are recoverable.


Many individuals scarcely turned of forty years of age, and sometimes even earlier, who have lived rather freely, are not unfrequently about that period of life, greatly changed in their powers of sexual intercourse. They may, indeed, in general health and personal appearance, be stout, and, for several years, not very sensible of the degeneration of their powers, but the frequency of their inclination for such duties gradually becomes much diminished, and that is a symptom which is at all times indicative of approaching impotence; for the inclination gradually and entirely ceasing, the power speedily follows, or rather both are lost together. In others, about the same time of life, the physical power ceases first, and the inclination continuing often for many years after, they are obliged to gratify themselves in amusements which are the mere pantomime of amorous indulgences!—such individuals, being otherwise in tolerable health, are recoverable.


Many individuals scarcely turned forty years of age, and sometimes even earlier, who have lived rather freely, are not unfrequently about that period of life greatly changed in their powers of sexual intercourse. They may, indeed, in general health and personal appearance, be stout, and for several years not very sensible of the degeneration of their powers, but the frequency of their inclination for such duties gradually becomes much diminished, and that is a symptom which is at all times indicative of approaching impotence; for the inclination gradually and entirely ceas-
ing, the power speedily follows, or rather both are lost together. In others, about the same time of life, the physical power ceases first, and the inclination continuing, often for many years after, they are obliged to gratify themselves in amusements which are the mere pantomime of amorous indulgences! Such individuals, being otherwise in tolerable health, are recoverable.


Many men, about the age of forty, who have lived freely, are not unfrequently greatly altered in their power of sexual intercourse. They may, indeed, in general health, be stout and hearty, and for several years not very sensible of diminution in this respect but the frequency of their inclination for such duties, gradually becomes less, a symptom at all times portentous of approaching impotence; for the inclination ceasing, the power soon follows. In others about the same time of life, the physical power ceases first, and the inclination continuing often for years, their amusements become but the merest pantomimes of amorous indulgence! Such individuals otherwise in tolerable health, are recoverable.


Many individuals scarcely turned forty years of age, and sometimes even earlier, who have lived rather freely, are not unfrequently about that period of life greatly changed in their powers of sexual intercourse. They may, indeed, in general health and personal appearance, be stout, and for several years not very sensible of the degeneration of their powers; but the frequency of their inclination for such duties gradually becomes much diminished, and that is a symptom which is at all times indicative of approaching impotence; for the inclination, gradually and entirely ceasing, the power speedily follows, or rather both are lost together.

—J. W. Bate, Dr. Bate’s True Marriage Guide: A Treatise for the Married and Marriageable, both Male and Female. ([Chicago]: [s.l.], [1889]), 98.

Notes

1. A grant from the Social Sciences and Humanities Research Council of Canada supported research for this article. I am grateful to Alexandrea Fiorante, who helped me organize several hundred pieces on book-making. Thanks to Jim Secord, Leslie Howsam, Maria Damkjær, the anonymous readers, Yuri Cowan, and Ryan Cordell, who gave feedback on drafts at different stages in the article’s development, and to the many colleagues who suggested readings or listened to scraps of this work at the Research Society for Victorian Periodicals/Victorian Studies Association of Western Canada joint conference (2018), the University of Toronto’s Work in Nineteenth Century Studies seminar (2020), and the Migrations of the Book conference at Texas A&M University (2021).


10. This article’s characterization of how book-making was represented draws on surveys of the following: (1) Pieces in digitized historical newspapers originally published in the UK (Gale’s *British Library Newspapers, Parts I-IV, 19th Century UK Periodicals, Parts I and II, and the British Library and FindMyPast’s British Newspaper Archive*), the US (Gale’s *Nineteenth Century U.S. Newspapers*) and Australia (the National Library of Australia’s *Trove*); (2) Digitized historical books and periodicals in Google Books and the Internet Archive, Gale’s *Early English Books Online* and *Eighteenth Century Collections Online*, and JISC’s UK *Medical Heritage Library*; (3) Reader’s reports issued between 1866 and 1925 in Add MS 55931-55996, Macmillan Archive, British Library, UK; (4) Previous scholarship that mentions the use of “book-making” as a label for textual production, or quotes from a source that uses the term in this context. The examples that I present in this article are primarily from Britain. However, there were no obvious differences in the term’s usage in American and Australian sources.


14. One of the more interesting aspects of this history—one that I wish I had room to explore here—is how much format mattered. As Ryan Cordell explains in “Reprinting, Circulation, and the Network Author in Antebellum Newspapers,” American Literary History 27, no. 3 (2015): 417–445, nineteenth-century readers not only expected material in periodicals to be derived from other sources but considered pieces that achieved “viral” status valuable because they had achieved it. The same was clearly not true of material in books.


16. For example, see Will Slauter, Who Owns the News? A History of Copyright (Stanford: Stanford University Press, 2020); Catherine Seville, Literary Copyright Reform in Early Victorian England: The Framing of the 1842 Copyright Act (Cambridge: Cambridge University Press, 1999).


21. For example, see Simon Latham, Lathams Falconry Or the Faulcons Lure, and Care in Two Bookes... (London: John Beale, 1614), 1, Early English Books Online.

22. The earliest example I have seen of “book-making” being used as described in this and the following paragraph is in “Art. 11. Divers Parts of the Holy Scriptures done into English, Chiefly from Dr. J. Mills’s Printed Greek Copy with Notes and Maps,” The Critical Review, Or, Annals of Literature 15 (1763): 484–485. Early English Books Online, but this was clearly not the first instance. It seems to have been common to use the term in this way from the mid-1780s.


25. For example, see Rev. of George Eliot’s *The Makers of Florence, The Saturday Review*, February 24, 1877, 237–8, Google Books.


45. Cedric J. Robinson, *Black Marxism: The Making of the Black Radical Tradition* (Chapel Hill, NC: University of North Carolina Press, 2000), 115. See also Ranjodh Singh Dhaliwal, “The Cyber-Homunculus: On Race and Labor in Plans for Computation,” *Configurations* 30, no. 4 (2022): 377–409, 405–6; Trevor Burnard and John Garrigus, *The Plantation Machine: Atlantic Capitalism in French Saint-Domingue and British Jamaica* (Philadelphia: University of Pennsylvania Press, 2016); Schaffer, “Babbage’s Intelligence.” As Hopwood notes in “Keywords” 289, some historians have argued that “thinking in terms of capitalist (re)production de-humanized women as baby-producing machines” along these lines. Hopwood argues that this theory doesn’t hold up for white women, who were denigrated in other ways. However, Black women were often defined by their reproductive capacity.


48. Dhaliwal, “The Cyber-Homunculus,” 394–5, 402. Dhaliwal’s analysis of transhistorical figurations substantially informs my thinking in this section of the article, and led me to several of the sources cited in it.


50. For instance, George Eliot, Charles Babbage, and (retrospectively) Johann Wolfgang von Goethe were all accused of book-making. See E.S. Shaffer, “George Eliot and Goethe: ‘Hearing the Grass Grow’,” *Publications of the English Goethe Society* 66, no. 1 (1996): 3–22, 17; “Makers of Florence,” *The Saturday Review*, 237; Rev. of *The Ninth Bridgewater Treatise. The British Critic, Quarterly Theological Review, and Ecclesiastical Record*. Vol. XXII (London: Printed for J. G. and F. Rivington, 1837), 88. Google Books. The gendered associations of book-making would benefit from more careful exploration. My sense is while they were positioned as opposites, both inspired authorship and literary manufacture were initially associated with masculinity—the latter because it was associated with the marketplace—while practices like scrapbooking, which look a lot like practices associated with book-making, were associated with femininity because they were practiced in the domestic sphere.


53. Eichhorn, *Content*, 1–29; McNulty, “Content-Era Ethics,” also argues for the recent emergence of “content” as a medial category. However, McNulty advocates a much narrower, genre-based definition of content than I and other authors cited here do.


56. For instance, see Burton, The Anatomy of Melancholy, 6–7.


59. Eichhorn, Content, 25.


65. For more on “plagiarism hunters,” see Macfarlane, Original Copy, 40–45.

66. For example, see “Book-Making,” Saturday Review, 215. These concerns—clearly connected with the concerns about information management that Lee examines in Overwhelmed—are somewhat ironic when we consider that premodern readers dealt with similar concerns by creating compilations. See Blair, Too Much to Know.


73. See Henrickson, Computer Generated Texts, 1–2 for a brief discussion of some of the ways we already consume AI generated text without knowing it.
74. Sanborn, Plagiaram!, 14.
75. See discussion in Sanborn, Plagiaram!, 27–8.
83. The study took shape organically during research for a different project. Initially, I suspected that Manhood was derived from a variety of sources and set out to trace them in two versions of the manual. When I began to recognize the patterns of text reuse described below, I decided to examine a more prestigious work in the medical genre (Ryan’s Marriage), and then a nonmedical work (Hamilton’s Marriage Rites). I also used the method I describe to examine parts of histories and travel guides and detected similar patterns. However, due to concerns about the length of the article, I focus only on these four works here.
84. Ryan Cordell, “‘Q i-jtb the Raven’: Taking Dirty OCR Seriously,” Book History 20 (2017): 188–224. Other issues include the fact that Google’s search algorithms attune to user behaviour, and that phrase length affects results. Much to my chagrin, Google Books does not handle searches for long passages very well. Searches for six- to eight-word phrases tended to return the most relevant results. Frankly, this is not a method that I would recommend for anything more than exploratory work.
Content Generation in the Age of Mechanical Reproduction

Orme, Brown and Green, 1830, 814; Manhood (1832), 86–7 and John Baco, “A Treatise on Syphilis” in W.M. Atkinson, Popery Unmasked (Leeds: T. Inchbold, 1829), 123–5; Manhood (1832), 103 and “Original Communications: Mr. Cane’s Cases of Moon Blindness,” Dublin Journal of Medical Science. Vol. 17–18 (1Dublin: Hodges and Smith, 840), 176. Note: I do not count the Cooper and Lawrence reprints in periodicals as separate sources in the main text of the article, but I include them here as Google was not able to locate matching text in the books themselves.


95. See citations in n. 70 and Michael Ryan’s Writings on Medical Ethics, ed. Howard Brody, Zahra Meghani and Kimberly Greenwald (London and New York: Springer, 2009), 17–29 for further information about Ryan’s reuse of his own writing.


97. Compare Curtis, Manhood (1852), 6 with The Silent Friend (1847), 109–110; Brodie’s Medical Work on Virility (1844), 115; Ryan, Marriage, 64.

98. Compare Hamilton, Marriage Rites, 15 with Craven, Journey, 233; The Surprising Savage Girl, 12; Clarke, Picturesque and Moral Geography, 205; Damberger, Travels Through the Continent of Africa, 356.


100. Compare O’Donnoghue, Marriage, 4 and “Prize Essay: Marriage,” 248.

102. Like Manhood, these works were issued in many editions. I compared text from the 1840 and 1852 editions of Manhood with the following editions in Google Books: The Silent Friend (1841; 1847), Manly Vigour (1841; 1842), Brodie’s Medical Work on Virility (1844), Silent Hints on Certain Subjects (1852), and Self Preservation (1841; 1852).

103. See Manhood (1840), 5–8; Onania (1752), 186–7; John Roberton, On the Generative System, (1817), 77; Silent Friend, (1847), 16–18; Manhood (1852), 9–11.


105. McGillen, Fontane Workshop, 20. See also Cordell and Blankenship, “Reprinting Wright.” Although recent studies of Brown’s and Hopkins’s work emphasize how much text they borrowed from other works, the examples they provide underscore that these authors usually edited borrowed material extensively in their fiction.

106. Lisa O’Connell, “‘Matrimonial Ceremonies Displayed’: Popular Ethnography and Enlightened Imperialism,” Eighteenth-Century Life 26, no. 3 (2002): 99–116 suggests that works the proto-anthropological genre to which Hamilton’s Marriage Rites belongs were very similar because they copied so much from each other, which does point to routine copying affecting other genres in similar ways.

107. Cordell and Blankenship, “Reprinting Wright.”


109. McGillen’s work suggests that book-making had a counterpart in the German lands and probably other European regions, but here I am thinking mainly of countries like China, which developed different norms around copying. See William P. Alford, To Steal a Book is an Elegant Offense: Intellectual Property Law in Chinese Civilization (Stanford: Stanford University Press, 1995).

110. For a longer discussion of this issue with language models and a description of techniques that might be used to learn more about their training data, which is typically a closely guarded corporate secret, see Ryan Cordell, “Towards a Bibliography for AI Systems,” Programmable Type: Thoughts on Book History, Bibliography, Information Arts, & Digital Humanities, April 21, 2023, https://ryancordell.org/research/aibibliography/.
