



Citation: K. Ozment (2025)
Digital Bibliography in the Age
of Linked Data. *Jems*. 14:
pp. 33-45. doi: [http://dx.doi.
org/10.36253/JEMS-2279-7149-
16517](http://dx.doi.org/10.36253/JEMS-2279-7149-16517)

Copyright: © 2025 K. Ozment.
This is an open access, peer-
reviewed article published by
Firenze University Press ([https://
oajournals.fupress.net/index.php/
bsfm-jems](https://oajournals.fupress.net/index.php/bsfm-jems)) and distributed under
the terms of the Creative Com-
mons Attribution License, which
permits unrestricted use, distri-
bution, and reproduction in any
medium, provided the original
author and source are credited.

Data Availability Statement:
All relevant data are within the
paper and its Supporting Infor-
mation files.

Competing Interests: The
Author(s) declare(s) no conflict
of interest.

Editors: D. Pallotti, P. Pugliatti
(University of Florence)

Digital Bibliography in the Age of Linked Data

Kate Ozment

Case Western Reserve University (<kate.ozment@case.edu>)

Abstract

The article explores how the interplay of ideological values and technological capacities have shaped the digital bibliography of British print history. Using a misgendering in the English Short Title Catalogue (ESTC) as a case study, the article explores how information flows through resources like Eighteenth Century Collections Online (ECCO), and Early English Books Online (EEBO), library catalogues, WorldCat, and retail outlets like Amazon. The article argues that as data from the ESTC is reproduced through linked data structures, information is 'authorized' far beyond what a single resource would do alone or what its original authors imagined or designed. While feminist, queer, and critical race scholarship has discursively created and revised new histories of textual production, in contrast foundational resources like the ESTC perpetuate old assumptions with unfixed errors and editorial practices that render the *who* and the *why* of their metadata choices opaque. The article concludes that radical revision is necessary if we are to disrupt centuries of a white and male norm in British print history.

Keywords: Early English Books Online, Eighteenth Century Collections Online, English Short Title Catalogue, Feminist Bibliography, Print History

In early 2023, I noticed an editorial addition in the English Short Title Catalogue (ESTC) that incorrectly attributed the work of a woman to that of a man. With the gumption of a feminist book trade scholar, former copy editor, and digital humanist who has worked with bibliographic resources for a decade, I went about trying to fix this error. My inability to do so revealed that beneath the seeming logical ease of a digital database was a tangled web, metaphorically and literally, that is the current digital bibliographic landscape of print in the Early Modern period. This article uses my failure as a launching point to examine when and how information is introduced into the ESTC and the circulation of data beyond the intentions of the project's architects or the database's virtual address. By detailing a media history of the ESTC from multiple perspectives, I show how chronic misunderstandings of what the database is and is not

have been magnified as the data is imperfectly duplicated in commercial enterprises, scholarly projects and other bibliographic formats. The result is that we have never known more nor misunderstood more about print history in the early modern period, especially as it relates to the history of women's textual labour.

The editorial addition at the centre of this narrative was on record P2961 for the March 1715/16 edition of *Chit-Chat*, a periodical linked to Richard Steele. The imprint read: 'Printed; and Sold by R[ichard]. Burleigh in Amen-Corner', with the bracketed additions expanding the 'R' initial to 'Richard.' A check in the British Book Trade Index (BBTI) and other sources, including the Stationer's Company archives and Henry R. Plomer's *Dictionary* (1922), indicated that there was no Richard Burleigh working in the trades at this time, especially not at that address. Biographical research did not yield any Richard Burleigh in this period or location.

Eventually, I determined that the business must belong to a woman named Rebecca, not Richard. Rebecca Burleigh was born Rebecca Falls and married Ferdinando Burleigh in 1712; they had a son named Ferdinando. Ferdinando Burleigh Sr. was a bookseller with a shop in Amen Corner, which was north of Ludgate Street and to the west of St. Paul's Churchyard in London. He died in 1715, and Rebecca continued on in the business until 1719. In 1717, she married William Graves, also in the trades through his family, but she continued to publish as 'R. Burleigh' rather than changing her business name after her second marriage. I cannot find traces of when Rebecca left the trades or why, but William's father John Graves indicated in his will in 1726 that his son was 'in foreign parts' (quoted in Treadwell 1999).¹ Perhaps Rebecca went with him.

Rebecca Burleigh is well documented for a woman in the trades in the sense that she is in the BBTI, which is an imperfect but useful resource that aggregates references like Plomer (1922) and Foxon (1975). Despite the fact that 'R. Burleigh' accounts for 193 records in the ESTC, it is her husband Ferdinando with his 57 imprints who makes it into Plomer's *Dictionary*. Because the name 'Richard' is on ESTC record P2961 and there is no 'Rebecca Burleigh' or 'Mrs. Burleigh' in any other imprints listed in the ESTC that I could find with text searches, I went through an extensive process of elimination to arrive at the most logical conclusion – that Richard Burleigh is the invention of an editor who assumed an initial would be a man instead of a woman in the absence of other information. Why 'Richard' rather than 'Robert' is harder to guess, but perhaps with 'Richard Steele' floating in the cataloguer's head, that seemed the right name for the publisher as well.

After posting on social media about my consternation at the spectral Richard Burleigh, a cataloguer at the British Library edited the record to reflect the name 'Rebecca' rather than 'Richard'.² This was a helpful and satisfying thing to do, and I briefly experienced all the gratification that a feminist bibliographer feels when correct information about a woman in the trades is now circulating out in the world. Except it did not circulate. While Rebecca now sits on her throne on ESTC record P2961, that is the extent of her bibliographic kingdom. The usurping Richard remains in the duplicates of the ESTC's metadata that circulate beyond the catalogue's digital residence.³ He is in the Eighteenth-Century Collections Online database (ECCO) hosted by

¹ I was unable to get a copy of this will through a request with the probate records office of the National Archives.

² This change overwrites the MARC (Machine Readable Cataloguing) field where I would have been able to see the last edit on the record, something I did not realize until later. However, it likely would not have yielded much information, as similar edits have all been twenty or more years ago.

³ See <<https://estc.bl.uk>>. The British Library experienced a cyber-attack in 2023 that rendered this URL and many others, inoperable at the time of publication. See the *Coda* for more.

Gale, identifiable by searching with the ESTC record locator. As a consequence, he is also in all the corners of the Internet where Gale's bibliographic metadata has travelled including WorldCat, where the title in the ESTC reference has listed holdings and an assigned OCLC (Online Computer Library Center) number. Gale sells ECCO's metadata to subscribing libraries so that ECCO holdings come up in library catalogue searches by users. Richard appears in dozens of library catalogues to be pulled up in full-text searches by undergraduates and their professors. Gale also sells books on demand on Amazon that are reprints of its ECCO holdings – allowing me to hold in my hand a codex reprint of a digitization of another microfilm digitization of a codex that was printed in the 1700s and now lives in the Bodleian. Matryoshka doll technologies aside, in the description for Amazon's listing of *Chit-Chat*, it reads 'The below data was compiled from various identification fields in the bibliographic record of this title. This data is provided as an additional tool in helping to insure edition identification'.⁴ The subsequent information is exactly what is on the ECCO record, including Richard and the surrogate's original copy location.

As months passed after Rebecca was listed on ESTC and no changes occurred anywhere else, I began to unravel my own assumptions about how our digital bibliographic ecosystem communicates. I had erroneously assumed that ESTC information appearing in full transcription in ECCO meant that there was an ongoing dialogue between the ESTC and Gale, which was my experience working in relational database platforms. This is not the case. ECCO's description page, accessible only with a log-in, emphasizes that the ESTC information within is 'static' and that the ESTC should be consulted for the most up-to-date information ('About This Archive' n.d.). ECCO's old interface had an error report button that sent information back to the ESTC when a user located an error, but the button has been removed on the new Gale Primary Sources interface. The metadata on ECCO is a copy of the ESTC made at a certain point in time, done for the benefit of the scholarly community so that ESTC numbers can be searched, and digital surrogates connected to (hopefully) relevant bibliographical information. But as the case of Rebecca Burleigh emphasizes, this veneer of easy identification covers cracks in the accuracy and comprehensiveness of information that is available to ECCO users who may see the ESTC's data and not feel the need to go to an external website and check it.

As I will detail in this article, many scholars are aware of the significant drawbacks of our key databases. However, I do not think that my experience is so singular that a *mea culpa* is sufficient for explaining why I was confused about the information flow between resources – and that is because the resources do not gesture to their own cracks both by design and by happenstance. From its origins, what the ESTC is and could be has been debated by its architects and users, with articles published simultaneously in different journals to haphazardly overlapping audiences. While the history of the ESTC has been told previously, it is necessary to give a short account of the database's creation focused on its technological negotiations and from differing points of view to emphasize that the ESTC has always been more of a contested territory than a neutral reference source – and that much of these debates have been muted for the generation of scholars that have 'grown up' with an open-access ESTC database, Google Books, and scholarly databases sharing bibliographic information.

What is now known as the ESTC began as distinct bibliographical projects and library catalogues that eventually coalesced into the database that we now use. The structure of the ESTC was built as the Eighteenth-Century Short-Title Catalogue (18th STC), which began in earnest in 1977 as a collaboration between the British Library and a collective of researchers in

⁴ <<https://www.amazon.it/principles-methodically-exhibited-explaiend-Nicholas/dp/1140996940>>, accessed 1 December 2024.

the United States. From the beginning, Henry L. Snyder explains, the 18th STC project was intended to use MACHine-Readable Cataloguing files, or MARC records, because ‘The enormous expansion of printing in the eighteenth century seemed to obviate the method employed for the manual compilations that preceded it’ (2003a, 105). These ‘manual compilations’ include the Short Title Catalogue (STC), which covers 1475-1640 and is a compilation by many hands including Alfred W. Pollard, Gilbert R. Redgrave, William A. Jackson, Frederic S. Ferguson, and Katharine F. Pantzer. The other is the Wing Short Title Catalogue, covering 1641-1700 and short-handed as ‘Wing’ for its author, Donald Wing. The STC and Wing set the standard for bibliographical work in England, but their origins as manually compiled print volumes meant that they could not be direct templates for the work of the 18th STC.

The 18th STC embraced new technologies, but this required significant financial investment from the U.S. and U.K. governments – in the neighbourhood of \$30 million, according to Snyder (2003a). It is unlikely such an investment would be made now, which speaks to both the accomplishments of the project and how difficult it would be to make large-scale changes to nearly 500,000 records. Eventually the 18th STC became the English Short Title Catalogue as it incorporated the STC, Wing, and data from other major catalogues like that of the American Antiquarian Society. The ESTC was available first through the Research Libraries Group RLIN system and the British Library’s BLAISE system. Users were able to purchase CD-ROMs in the 1990s before it migrated online in 2006, where it has continued as an open-access resource maintained largely through the British Library and the University of California at Riverside’s small editorial team at the Center for Bibliographical Studies and Research (Gadd 2009, 684). Thousands of libraries contribute to the catalogue to expand its listed holdings. Now that the database can be hyperlinked to other resources, the ESTC directs users to ProQuest and Gale databases, Hathi Trust, and Google Books where digital surrogates are available.

The ESTC’s data structures have been a constant negotiation between different national standards, existing research networks, and legacy projects that were retrofitted into its information architecture. At the heart of the database are MARC records, which required standardization between the U.K. and U.S. to allow libraries in both countries to contribute. MARC is the pioneering work of Henriette D. Avram at the Library of Congress, and it has three separate components: ‘the *structure*, or “empty container,” the *content designators* (tags, indicators, and subfield codes) used to explicitly identify or to additionally characterize the data elements, and the *content*, the data itself (author’s names, titles, etc.)’ (1975, 10). The goal of MARC, according to Stephen H. Gregg, was ‘not only about designing a record to be parsed by a computer; it also set the standard for bibliographical records that libraries across the world would follow’ (2020, 8). Avram’s dream, which has become a reality, was that MARC records would decrease the cost of cataloguing by allowing for copy cataloguing rather than original cataloguing where possible. That is, if the Library of Congress created a MARC record for an object, a subscribing library could adapt that record for their similar object without cataloguing it from scratch. Copy cataloguing has become the norm even as libraries transition to BIBFRAME, which is marketed as more flexible than MARC, although may not be (see Edmunds 2023). By using MARC in the ESTC, the database has become the original cataloguing record against which others are copy catalogued, marking it as the authority in the field. However, it is worth noting that many records in the ESTC were copy catalogued from subscribing libraries like the Bodleian to begin with. Other data was pulled from diverse sources including card catalogues that had to be hand-keyed, requiring that British Library cataloguers learn how to type (Snyder 2003a, 122). The standards set by the MARC structure decided on in the 1970s and 1980s have been maintained: when new records were incorporated into the ESTC with data fields that did not

fit what was created, they were 'shorn'; record structures were imperfectly adapted in the 1990s to accommodate serials, which were initially left out (133). Texts deemed not applicable for the database were left out, which included genres like 'domestic papers' that tended to reflect women's writing, as Margaret J.M. Ezell (2007) has argued.

Reflecting back on their work at a 2003 conference, several of the ESTC's major stakeholders and advisers articulated their vision for what the project does and could be. G. Thomas Tanselle commented that 'the database that has emerged is by now so comprehensive that the work can be regarded as substantially complete' (2003, xi). Snyder confirms that, despite the fact the ESTC is a 'living, constantly growing and changing organism' (2003b, 21), the team was 'close to our goal of creating a union catalogue that includes all the known imprints of the English-speaking world to 1801' (2003a, 149). The use of 'union catalogue' is intentional; the design of the ESTC is to be a catalogue of catalogues, a record of other records. It was supposed to be the 'one-stop shop' for all bibliographical information, library holdings and digitizations of titles in English prior to 1801 (Crump 2003; Snyder 2003a). While the 2003 conference coincided with the last CD-ROM edition of the catalogue before it was moved online by the British Library, the editors were already imagining how HTML and hyperlinking could make their 'one-stop shop' a reality – the original, accurate, and comprehensive record against which all pre-1801 books would be copy catalogued.

The editors of and contributors to the ESTC – both named people, who are all men, and the many unnamed people who did the manual labour of typing and cataloguing and were mostly women – are right in celebrating its monumental achievements. Yet the idea that the ESTC is either 'comprehensive' or 'complete' is far from how it was perceived by users in the early 2000s or now. Stephen Karian comments that 'many errors and inconsistencies pervade the file' (2011, 286), and James E. May adds that 'Many, including the ESTC's editorial staff, seem not to realize the incompleteness of ESTC's records for even eighteenth-century imprints, the project's original focus' (2001, 288). Errors proliferate from the fallible hands that keyed data and omissions multiply from the ESTC's origins as three separate editorial projects, formats, and staffs (Gadd 2009, 684). Part of the reason for the divide between Karian and May's perceptions of the ESTC and Tanselle and Snyder's assessments of their work is a misunderstanding, on both sides, about what the ESTC is and does. As noted, it was designed as a 'comprehensive union catalogue', which Ian Gadd defines as 'a single catalogue that merges together the existing catalogue records of other libraries' (2009, 684). In practice, this means that the ESTC represents what is in contributing *libraries*, not the complete output of the period. The ESTC also notably excludes many handbills, broadsides, and other ephemera (Hume 2007; Suarez 2009). It may be that 'comprehensive' describes everything that is in a contributing library and correctly catalogued as from this period rather than everything that existed in the period itself, but in either case the ESTC is not actually comprehensive.

There is also debate about whether or not the name is descriptive of what the ESTC is. Hugh Amory quips that 'In many ways the ESTC resembles the Holy Roman Empire: it is neither English, Short-Title, nor a Catalogue, since the "cataloguing" is only a response shaped by the system at the user's request' (2001, 8). The catalogue has focused on titles printed in the U.S., U.K. and Ireland, which is far beyond England. The titles are not all in English, nor does the ESTC encompass the entirety of the Anglophone world. The 'short title' aspect was happily abandoned as digital space became cheaper and now cataloguing emphasizes full titles, although many truncations remain in the database.

What is particularly interesting for my quandary is whether or not the ESTC is a 'catalogue', given that it is used as the authority record for copy cataloguing in other libraries and

resources. Amory argues that a union catalogue is ‘in contrast to a bibliography, which lays bare an unseen, inexplicit reality from documents as well as books’ (*ibid.*). He sees the primary benefit of the ESTC is its role as an index ‘whose accuracy and exhaustiveness depend on the illogical whims of language’ (*ibid.*). Additionally, David Vander Meulen posits that the ESTC might be better described as a bibliography and its users as bibliographers ‘for every search produces a list that represents a selection from infinite possibilities, some new imposition of order on the plenitude that is the ESTC database’ (2003, 67). Stephen Tabor concurs with this assessment, not because of the querying aspect of the resource, but because ‘its records describe groups of copies rather than specific copies’ making it more akin to a bibliography (2007, 369). Expanding on this idea, Gregg examines Patrick Browne’s *The Civil and Natural History of Jamaica* (1789) and notes the ESTC record (T89758) is a ‘palimpsest of all the book copies consulted’ rather than a description of a single artifact that exists. He concludes that ‘no catalogue could possibly account for all the variations of all book copies held in the world’ (2020, 30). All ESTC records are palimpsests to some degree, representations of omnibus descriptions that may or may not exist as copies.

In sum: when ESTC data is replicated on ECCO, it may or may not apply to the surrogate that viewers are encountering, as the digitized copy may have different bibliographical features than the description of many copies could account for. Gadd has a similar observation for Early English Books Online (EEBO), the ProQuest database that accounts for print output in England up to 1700 and similarly incorporates ESTC data to help users search for information:

By bringing together the bibliographical record for an edition and (usually but not always) only a single witness of that edition, *EEBO* is obviously aiming to provide a useful scholarly mechanism in terms of searching but by doing so are implying – albeit not deliberately – that the record and the copy *are one and the same thing*. It would be better, perhaps, if *EEBO* represented itself as a library of copies, rather than a catalogue of “titles”. (2009, 687)

EEBO has ‘no formal mechanism for synchronising the data’ between it and the ESTC, which has led to ‘an increasing likelihood of significant discrepancy between the two resources’ (686).

Within EEBO and ECCO, then, are several different versions of ESTC data, pulled at different points from the core database from palimpsestic concepts of books that may or may not exist and with limited if any mechanisms for synchronizing what is on each with the latest bibliographical information. This network does not function like a database system, as I erroneously assumed, which would allow one change in information to automatically populate across the entire ecosystem. Competing commercial companies, different systems and different timelines for creation has likely rendered this a financial impossibility if not a technological one.

It is debatable how much the average user of the ESTC, EEBO, or ECCO keeps the layered epistemologies of what they are viewing in mind as they interact with these platforms. The debate over what the ESTC ‘is’ is not simple semantics, but involves differing ideas and competing viewpoints about what data in the ESTC represents and how it should be used – often divided between the cataloguers and information architects who populate and maintain the data and scholars deeply engaged with its information structures from the perspective of a user. What users interact with online has the feel of a unified resource, and now that the ESTC has been online for 17 years and ECCO for 20 years, there is an entire generation of scholars for whom the above history is learned rather than a lived experience as it was for scholars who experienced the growing pains and powers of the digital turn in Early Modern book cultures firsthand.

Although the ESTC has grown beyond a small number of cataloguers and editors to thousands of contributing libraries, its perception has remained that of a highly controlled and

verified union catalogue. Whether or not it is an authority record, it is used as one similarly to the Virtual International Authority File (VIAF) for biographical subjects. ESTC and VIAF records are the default, but it is much harder to create or change a VIAF record than an ESTC record. VIAF records must be changed by Name Authority Cooperative Program-certified librarians, and the information must meet rigorous standards to be included. In contrast, an ESTC record can be altered by a quick error report, without the editor completing independent research. Yet ESTC records function as bibliographical authorities, and accordingly they legitimize information beyond what a single catalogue, union or otherwise, could have ever hoped to do.

If we return to Richard and Rebecca Burleigh, we can see this distinction in concrete terms. A cataloguer introduced Richard Burleigh into the ESTC without any of the requirements that would have been asked of someone creating a VIAF record, and accordingly Richard has populated all the places where ESTC data has been replicated. However, when the database was updated to reflect Rebecca rather than Richard, other entities that use the 'authority' of an ESTC record through transcription rather than linking back to the record were not also updated. Worse, there is seemingly no mechanism for doing so. At what point, then, is the fictional Richard more 'real' than the empirical reality that Rebecca Burleigh existed and by all accounts sold this book? Tabor argues that the tension between the 'fluidity of content' (2007, 367) in the ESTC and its 'status as an authoritative resource' is precisely what is at stake when we consider the repetition of errors. He expands, 'Transcription errors acquire the weight of authority when the records in which they occur accumulate multiple verified matches ... The same human qualities that give rise to transcription errors by cataloguers allow their persistence through numerous proofreadings by matchers' (372).

The 'weight of authority' is both bibliographical and psychological. 'Truth' is a subjective human perception, and repetition is a powerful mechanism for its creation. Psychologists have established a link between repetition of information and the likelihood of the hearer believing that information is true, which has been called the illusory truth effect. Put another way, 'if people think they have heard something before, they will be more likely to believe it' because we correlate familiarity with validity (Arkes, Hackett and Boehm 1989, 82). Dozens of studies have expanded on and nuanced how this phenomenon occurs in terms of facts and beliefs, ranging from politics to perceptions of trust in medical practitioners. Jessica Udry and Sarah J. Barber summarize that 'this repetition-based increase in belief occurs even when people know the repeated information is factually incorrect ... or comes from unreliable sources' (2023, 1) and that the effect works even through 'paraphrases, or non-verbatim repetitions' (11).

Traveling ESTC data has created discrete repetitions that are not perceived by all users as interactions with the *same* information, but interactions with *repeated* information. The average user of the ESTC or ECCO, including students and scholars, are likely not aware of the full extent of the ecosystem of Gale's products and its relationship with ProQuest, nor of how library information structures work and what happens behind the scenes to enable an ECCO holding to appear on WorldCat with an OCLC number. Literary and historical students are not taught about copy cataloguing nor the financial constraints on original cataloguing for even well-funded libraries like the Houghton at Harvard. Therefore, when users encounter Richard Burleigh on the ESTC, ECCO, WorldCat, their library catalogue and Amazon, there is a good chance they might understand these as discrete interactions with a repeated fact, feeding into the illusory truth effect. This will be especially true for students who are taught in information literacy and research courses to look for reliable academic sources, most of which these would be. This is amplified further when we consider the high likelihood that such information would be referenced in citations for publications, creating more opportunities for the information

to be repeated and therefore verified as ‘true’. The editorial expansion, then, of ‘R. Burleigh’ to ‘R[ichard]’ rather than ‘R[ebbecca]’ has impact far beyond users looking up the item in the ESTC, and it verifies Richard as ‘real’ with far more strength than a correction that stays in its discrete space on the ESTC, a lone note of feminized dissonance.

It is not a coincidence that my example involves the elision of a woman’s labour, and the impacts of bibliographic sexism on the creation of the ESTC, ECCO, and EEBO need far greater scrutiny than they have been given by the scholars who have discussed the limits of these resources in other terms. As I have worked with ESTC data in my research on women publishers, I have encountered at least a dozen of these errors and assumptions, and the team members that I work with at the Women’s Print History Project (WPHP)⁵ each has their own examples of similar moments. The WPHP references the ESTC, but it requires original cataloguing rather than copy cataloguing. This is a time-intensive process, but re-cataloguing digital surrogates and printed objects with ties to women’s labour has yielded far more usable and reliable information than working with the ESTC alone, because doing so disrupts repetition as authorization. We can see an example of this disruption in ESTC record T19848, where an unknown editor put ‘[sic]’ next to the ‘K’ in ‘K. Sanger’ to indicate it was a typo.⁶ This may be because the previous edition of this book, *Calypso and Telemachus* by John Hughes, was printed by Egbert Sanger. However, the initial actually refers to Katherine Sanger, Egbert’s wife who is clearly named in his will. Given that Egbert was not living when *Calypso and Telemachus* was published in 1717, the ‘[sic]’ seems like an assumption rather than a researched observation.

It is not only cataloguers but users who make these assumptions. John Horden’s work on the Nutt family assumes that Elizabeth Nutt’s work can be ascribed to her son, Edward. Horden argues that Edward was at the Royal Exchange as ‘E. Nutt’ from 1722-1745, adding: ‘Foxon ... apparently assumes each “E. Nutt” to be Elizabeth Nutt, but in a number of instances ... it seems much more probable that Edward Nutt, whom Foxon does not notice, is indicated’ (1988, 23). Horden gives no justification for this argument, which is in a footnote to the main article largely uninterested in Elizabeth or her daughters. These arguments are despite significant documentation of Elizabeth’s work with her daughters in the Royal Exchange, as summarized by John D. Gordan: ‘[Elizabeth], in her own right a “mercury” and retailer of books and newspapers near the Royal Exchange, carried on the law-printing business under the patent until her retirement about 1740’ (2014, 258). Elizabeth was deposed a number of times due to the nature of the pamphlets she was selling, adding further evidence to imprint information that places her, not Edward, at the Royal Exchange (Hunt 2008). While family businesses were common, it seems clear from the examples of ‘Mrs. Nutt’ and ‘Eliz. Nutt’ in imprints, easily accessible by searching the ESTC, that while the Nutt children likely worked with their mother, she was the business’ proprietor as well as the ‘E. Nutt’ who continued to print under her late husband John’s law patent with her son, Richard. While Horden did not have the benefit of a full-text-searchable ESTC for his article in the late 1980s, he did have Wing and the 18th STC to reference.

As these examples suggest, publishers’ initials on imprints are assumed to be male even for the most well documented of women like Elizabeth Nutt. As with Rebecca Burleigh, men are even invented to explain away a woman’s clearly documented labour. Casual sexism in book trade scholarship or literary history more broadly is not a new revelation for those of us who have encountered the transformative recuperative work of Margaret J.M. Ezell (1987), Paula

⁵ <<https://womensprinhistoryproject.com/>>, accessed 1 December 2024.

⁶ As with the error on Rebecca Burleigh, I have since gotten this mistake fixed in the ESTC.

McDowell (1998), Lisa Maruca (2007), Helen Smith (2012), Maureen Bell (2014), Cait Coker (2018), Alan Farmer (2020), and Georgina Wilson (2022), nor those of us who have received antagonistic responses to our feminist scholarship in reader reports and at conferences. However, the impact of this sexism, casual and deliberate, is magnified by the structures that digital bibliography has built. It has been much of my work in the last few years to painstakingly look under every metaphorical stone to prove that fictitious men do not exist, spending hours correcting an assumption that likely took seconds to make and now carries the authority of the ESTC, ECCO, and WorldCat with it. Some of these cases are so glaringly frustrating that one wants to require every book history scholar to read Joanna Russ' *How to Suppress Women's Writing* (1983).

This is not simple personal irritation, however. It is another instance of the illusory truth effect that has rendered Richard Burleigh more 'real' than Rebecca Burleigh – many resources on the British book trades in particular and book history more broadly only offer examples of male tradespeople and define book trade production according to men's experiences, as I have enumerated elsewhere (Ozment 2020). Assumptions of masculinity are enmeshed in our key resources and narratives. It can hardly be the fault of cataloguers who are working with dominant narratives in their research that they repeat what book trade scholars continually repeat themselves. Nor is it the fault of subscribing libraries whose cut budgets have meant we must rely on copy cataloguing despite the fact that 'touch-it-once cataloguing can be problematic, since many of the characteristics of rare materials are unique to the holding repository, and the awareness of certain characteristics may not even exist when the initial cataloguing is done' (Ascher 2009, 99). At some point, the 'truth' that 'the little world of the book' has been a male domain' (Howsam 1998, 1) became such through repetition, rather than through documentary evidence, and despite the discursive histories feminist scholars have written to the contrary.

Justified and nuanced critiques have been levelled at the ESTC, ECCO, and EEBO and the ways that they have shaped our perceptions of print production in the hand-press period, even as they have enabled significant scholarship without the burden of travel.⁷ Gadd cautions that 'The canny user of *EEBO* needs to be conscious of the resource's history, the limitations of its coverage, the origins of its bibliographical data, and the nature of the relationship between its bibliographical catalogue and the individual copies available' (2009, 687), and Cassidy Holahan's analysis 'raises questions about the degree to which archives shape the research that comes out of them' (2021, 804). I will add that it is also necessary that we consider the political implications of how the ESTC, ECCO, and EEBO have shaped the research we complete on the Early Modern period's printed objects, especially as it relates to the ways the archives capture the history of women and their textual labour. The ESTC is not alone in defaulting to masculinity. The original selection criteria for ECCO includes more comprehensive digitization for 'widely recognized authors', all of whom are white men ('About This Archive' n.d.). Perhaps unsurprisingly, the list of the top fifty most printed authors in ECCO are overwhelmingly male, with Hannah More and the omnibus 'A Lady' as the lone feminized names. The former feeds into the latter – if only men were deemed 'recognized' and therefore were digitized at a greater rate than women, of course the most digitized pieces in ECCO will be men. If we consider Gadd and Holahan's cautions in light of gendered disparities in the ESTC and ECCO, it is clear that multiple authorizing functions are in play to convince a user, canny or not, that books and their makers are the province of men and male readers.

⁷ The high cost of access to EEBO and ECCO is a different sort of burden, of course.

What, then, is the feminist bibliographer to do? If they write discursive histories about women in the trades, as McDowell, Smith, Farmer, and Maruca have done, the work is deemed niche and does not make it into mainstream scholarship. This information is unlikely, then, to be read by generalist cataloguers who might work with ESTC data but are not experts on the nuances of different books or publishers. If the bibliographer is able to get the error fixed in the ESTC through reporting, the digital bibliographic landscape does not allow for the free flow of information that would be ideal with linked data. The error remains. The ESTC's latent replication is a surprisingly less reliable system than the one that was used in the 1980s when British and American cataloguers sent physical tapes of their work across the Atlantic to synchronize the systems (Snyder 2003a). If the bibliographer builds a separate project as we have done with the WPHP, they are able to work with data in new ways without navigating the slow-moving behemoth of international library cataloguing systems, but the project will always be a separate node compared to the 'one-stop shop' that is the ESTC and, more realistically given the changing landscape since 2003, EEBO and ECCO.

When the ESTC began in 1978, its architects and editors imagined a world in which they could make cataloguing easier for libraries by sharing information, and consequently empower scholars to do more powerful work than they ever could before. Now that the ESTC is nearing its fiftieth birthday, perhaps it is time for another reconsideration of not just what the database is or is not, but how it can be used as a tool for the visibility rather than invisibility of a multiplicity of perspectives, rather than the veneer of objectivity in a union catalogue. Karian advocates that a relational database is one way of accomplishing this vision, and this is what the WPHP team is exploring on a much smaller scale (2011, 294). To do this on a large scale would require Gale, ProQuest, and the ESTC to have open dialogue of data in complementary if not similar coding languages, not something that seems possible to imagine with the high cost of paywalls on one side and decreasing funding for cultural heritage projects on the other.

Ultimately, however, we may need to work beyond these systems as well as within them. Digital databases and other forms of data that do not inherently interrogate sexist epistemologies – including how sexism works in tandem with other interlocking systems of power such as racism, cisheteronormativity, and ableism – will not address the core issue. To address it, digital British print history can learn from other fields in digital humanities that have analysed how power structures work within data models. In particular, the work of Safiya Umoja Noble (2018), Roopika Risam (2019), Catherine D'Ignazio and Lauren F. Klein (2020), and Derrick R. Spires (2022), are key. As more feminist, postcolonial, queer, and anti-racist projects complete original cataloguing and supplemental metadata, we may be able to create a network of linked resources that are not bound by geographical locations or nationalist traditions and instead are linked by their commitment to doing better with the tools we have alongside sound bibliographical approaches. If these motivations were an essential aspect of rather than an addition to bibliographical training, over time we could see slow changes that disrupt illusory truths and instead move us closer to the empirical reality that women worked in and around the British book trades and beyond.

Coda

The digital ecosystem that is the subject of this piece experienced further stress and fragmentation when the British Library was the victim of a cyberattack in October 2023 that, among other things, took down the ESTC from its home. As of October 2024, there is no timeline for its rehabilitation. An incomplete but useful stopover has been established at the Print &

Probability website (Vogler 2024). The original cost of building the ESTC (\$30 million) would never be invested today, much less at the higher cost of labour in the 2020s, by U.S. and U.K. governments who chronically underfund systems of higher education and cultural heritage centres. If this data were lost, would it be recoverable?

The ESTC has been replicated enough times that its disappearing entirely is unlikely. But the stability of this data is now in question beyond my previous concerns, and instability is an issue for a resource used as an authority file. The ESTC is accessible only in a copied format with no ‘copy-text’ to compare it to. Without the authority of the ESTC to compare to other information, even more emphasis is placed on EEBO and ECCO and their matching of digital surrogates with bibliographic information. We have shifted the maintenance of bibliographic authority from cultural heritage institutions to commercial enterprises with high buy-in prices and little accountability from academic institutions or scholars except with their pocketbooks. While the work of a good-hearted cataloguer could resuscitate Rebecca Burleigh, no access point exists for the public with Gale and ProQuest.

The long-term impacts of the attack are still unknown, but the immediate effect is that scholar-produced data has been removed from scholars’ control until a new access point is built. Neither Gale or ProQuest has announced financial or material support for the British Library, even though they use the Library’s resources in their commercial products. This one-way relationship is increasingly extractive, and for my concerns, not conducive to the necessary large-scale changes that need to happen if we are to get a more complete picture of pre-1800 print in England.

Works Cited

- ‘About This Archive’ (n.d.), *Gale Primary Sources: Eighteenth Century Collections Online*, <https://web.archive.org/web/20250225144605/https://go.gale.com/ps/helpCenter?userGroupName=cwru_main&prodId=ECCO&docId=EFZIPA587871271&sid=bookmark-ECCO&asid=a4c37dbc>, accessed 1 December 2024.
- Amory Hugh (2001), ‘Pseudodoxia Bibliographica, or When is a Book not a Book? When it’s a Record’, in L. Hellinga, ed., *The Scholar & the Database: Papers Presented on 4 November 1999 at the CERL Conference Hosted by the Royal Library, Brussels*, London, Consortium of European Research Libraries, 1-14.
- Arkes H.R., C. Hackett and L. Boehm (1989), ‘The Generality of the Relation Between Familiarity and Judged Validity’, *Journal of Behavioral Decision Making* 2, 2, 81-94.
- Ascher J.P. (2009), ‘Progressing Toward Bibliography; or: Organic Growth in the Bibliographic Record’, *RBM: A Journal of Rare Books, Manuscripts, and Cultural Heritage* 10, 2, 95-110, doi: 10.5860/rbm.10.2.321.
- Avram H.D. (1975), *MARC: Its History and Implications*, Washington, Library of Congress, <<https://eric.ed.gov/?id=ED127954>>, accessed 1 December 2024.
- Bell Maureen (2014), ‘Women and the Production of Texts: The Impact of the History of the Book’, in J. Hinks and V. Gardner, eds, *The Book Trade in Early Modern England: Practices, Perceptions, Connections*, New Castle and London, Oak Knoll Press and British Library, 107-131.
- Chit-chat. In a letter to a lady in the country. ... By Humphry Philroye. Volume 1 of 3* (n.d.), <<https://www.amazon.com/Chit-chat-letter-country-Humphry-Philroye/dp/1170927556/>>, accessed 1 December 2024.
- Coker Cait (2018), ‘Gendered Spheres: Theorizing Space in the English Printing House’, *The Seventeenth Century* 33, 3, 323-336.
- Crumph Michael (2003), ‘The Origins of the ESTC: The Case for Vision’, in H.L. Snyder and M.S. Smith, eds, *The English Short-Title Catalogue: Past, Present, Future*, New York, AMS Press, 45-63.
- D’Ignazio Catherine and L.F. Klein (2020), *Data Feminism*, Cambridge, The Massachusetts Institute of Technology Press.

- Edmunds Jeff (2023), 'BIBFRAME Must Die', *ScholarSphere*, n.p., doi: 10.26207/v18m-0g05.
- Ezell M.J.M. (1987), *The Patriarch's Wife: Literary Evidence and the History of the Family*, Chapel Hill and London, The University of North Carolina Press.
- Ezell M.J.M. (2007), 'Domestic Papers: Manuscript Culture and Early Modern Women's Life Writing', in M.M. Dowd and J.A. Eckerle, eds, *Genre and Women's Writing in Early Modern England*, London, Routledge, 33-48.
- Farmer A.B. (2020), 'Widow Publishers in London, 1540-1640', in V. Wayne, ed., *Women's Labour and the History of the Book in Early Modern England*, London and New York, Arden Shakespeare, 47-74.
- Foxon D.F. (1975), *English Verse, 1701-1750: A Catalogue of Separately Printed Poems with Notes on Contemporary Collected Editions*, London and New York, Cambridge University Press, 2 vols.
- Gadd Ian (2009), 'The Use and Misuse of Early English Books Online', *Literature Compass* 6, 3, 680-692.
- Gordan J.D. (2014), 'John Nutt: Trade Publisher and Printer "In the Savoy"', *The Library* 15, 3, 243-260.
- Gregg S.H. (2020), *Old Books and Digital Publishing: Eighteenth-Century Collections Online*, Cambridge, Cambridge University Press.
- Holahan Cassidy (2021), 'Rummaging in the Dark: ECCO as Opaque Digital Archive', *Eighteenth-Century Studies* 54, 4, 803-826.
- Horden John (1988), '“In the Savoy”: John Nutt and His Family', *Publishing History* 24, 5-26.
- Howsam Leslie (1998), 'In My View: Women and Book History', *SHARP News* 7, 4, 1-2.
- Hume R.D. (2007), 'The ECCO Revolution', *Eighteenth-Century Intelligencer* 21, 1, 9-16.
- Hunt M.R. (2008), 'Nutt [formerly Carr], Elizabeth', *Oxford Dictionary of National Biography*, doi: 10.1093/ref:odnb/66882.
- Isaac P.C.G. (2015), *British Book Trade Index*, <<http://bbti.bodleian.ox.ac.uk/>>, accessed 1 December 2024.
- Karian Stephen (2011), 'The Limitations and Possibilities of the ESTC', *The Age of Johnson. A Scholarly Annual* 21, 283-297.
- Literary Print Culture: The Stationers' Company Archive, 1554-2007*, <<https://www.literaryprintculture.amdigital.co.uk/>>, accessed 1 December 2024.
- Maruca Lisa (2007), *The Work of Print: Authorship and the English Text Trades, 1660-1760*, Seattle, University of Washington Press.
- May J.E. (2001), 'Who Will Edit the ESTC? (And Have You Checked OCLC Lately?)', *Analytical and Enumerative Bibliography* 12, 3/4, 288-303.
- McDowell Paula (1998), *The Women of Grub Street: Press, Politics, and Gender in the London Literary Marketplace 1678-1730*, Oxford, Clarendon Press.
- Noble S.U. (2018), *Algorithms of Oppression: How Search Engines Reinforce Racism*, New York, New York University Press.
- Ozment Kate (2020), 'Rationale for Feminist Bibliography', *Textual Cultures* 13, 1, 149-178, doi: 10.14434/textual.v13i1.30076.
- Ozment Kate (forthcoming), 'Catherine Sanger: Publisher in Bartholomew Close', in A. Smith, R. Stenner and K. Kramer, eds, *The People of Print, 1700-1800*, Cambridge, Cambridge University Press.
- Plomer H.R. (1922), *A dictionary of the printers and booksellers who were at work in England, Scotland and Ireland from 1668 to 1725*, Oxford, printed for the Bibliographical Society at the Oxford University Press.
- Pollard A.W., G.R. Redgrave, W.A. Jackson, F.S. Ferguson and K.F. Pantzer (1976-1991), *A Short-Title Catalogue of Books Printed in England, Scotland & Ireland and of English Books Printed Abroad, 1475-1640*, 2nd ed, London, Bibliographical Society, 3 vols.
- Risam Roopika (2019), *New Digital Worlds: Postcolonial Digital Humanities in Theory, Praxis, and Pedagogy*, Evanston, Northwestern University Press.
- Russ Joanna (1983), *How to Suppress Women's Writing*, Austin, University of Texas Press.
- Smith Helen (2012), *Grossly material things: Women and Book Production in Early Modern England*, Oxford, Oxford University Press.
- Snyder H.L. (2003a), 'A History of the ESTC in North America', in H.L. Snyder and M.S. Smith, eds, *The English Short-Title Catalogue: Past, Present, Future*, New York, AMS Press, 105-154.

- Snyder H.L. (2003b), 'The Future of the ESTC: A Vision', in H.L. Snyder and M.S. Smith, eds, *The English Short-Title Catalogue: Past, Present, Future*, New York, AMS Press, 21-32.
- Spires D.R. (2022), 'On Liberation Bibliography: The 2021 BSA Annual Meeting Keynote', *Papers of the Bibliographical Society of America* 116, 1, 1-20.
- Suarez M.F. S.J. (2009), 'Towards a Bibliometric Analysis of the Surviving Record, 1701-1800', in M.F. Suarez S.J. and M.L. Turner, eds, *The Cambridge History of the Book in Britain*, vol. 5, 1695-1830, Cambridge, Cambridge University Press, 37-65.
- Tabor Stephen (2007), 'ESTC and the Bibliographical Community', *The Library* 8, 4, 367-386.
- Tanselle G.T. (2003), 'Introduction', in H.L. Snyder and M.S. Smith, eds, *The English Short-Title Catalogue: Past, Present, Future*, New York, AMS Press, xi-xv.
- Treadwell, J.M. (1999), 'John Graves', <https://www.trentu.ca/english/treadwell/Booksellers_g-n/GRAVES_John_bs_1r.html>, accessed 1 December 2024.
- Udry Jessica and S.J. Barber (2023), 'The Illusory Truth Effect Requires Semantic Coherence Across Repetitions', *Cognition* 241, 1-13, doi: 10.1016/j.cognition.2023.105607.
- Vander Meulen D.L. (2003), 'The ESTC in Practice: A Bibliographer's Perspective', in H.L. Snyder and M.S. Smith, eds, *The English Short-Title Catalogue: Past, Present, Future*, New York, AMS Press, 67-73.
- Vogler Nikolai (2024), 'English Short Title Catalogue (ESTC) Search', *Print & Probability*, <<https://estc.printprobability.org/>>, accessed 1 December 2024.
- Wilson Georgina (2022), 'Surface Reading Paper as Feminist Bibliography', *Criticism* 64, 3-4, 369-383.
- Wing D.G. (1945-1951), *Short-Title Catalogue of Books Printed in England, Scotland, Ireland, Wales, and British America, and of English Books Printed in Other Countries, 1641-1700*, New York, Index Society, 4 vols.