

ISSN 2516-8568



Digital Archive Review: The Internet Archive

Source: Midlands Historical Review, Book Review

Published: 27/05/2021

URL:

<http://www.midlandshistoricalreview.com/digital-archive-review-the-internet-archive/>

Digital Archive Review: The Internet Archive

Robert Frost

Abstract

This review is the first of a new series, intended as a learning resource, and aimed primarily at undergraduates about to embark on individual research projects and dissertations, but will also be relevant to anyone interested in the rich potential of digital archives for accessing primary sources. Here, Robert Frost discusses using the Internet Archive to access out-of-copyright books from the early nineteenth century. The Internet Archive is an indispensable resource for all those interested in modern British history, cultural history, and beyond, and one now more important than ever due to lockdown restrictions this past year

Biography: Robert Frost is an AHRC-funded doctoral student with joint Geography and History department supervision. He is interested in Georgian and early Victorian travel, exploration and field studies in the Eastern Mediterranean.

Throughout my PhD on the work of the Egyptologist and antiquary Sir Gardner Wilkinson (1797–1875), I have found the Internet Archive (IA) invaluable.¹ In this piece, I give a short introduction to the IA, recount how I have used it in my research, and cover a few of the problems that using such a digital resource inadvertently brings.

The IA started off in the mid-1990s, and is now an organisation with a number of branches: most famously, it runs the ‘Wayback machine’—an online archive of billions of website pages, as well as an ‘Open library’ which ‘loans’ new books for a limited time period. The IA also holds film and audio media. My focus here though is on a more specific part: its massive collection of out-of-copyright books, from the eighteenth, nineteenth, and twentieth centuries (and a few from even further back), sourced in large part from major public and university libraries in the United States.

Being open-access, the IA is simple to access: no passwords are required, although there is an option to create an account to access additional features, including the lending library of more recent books (Fig. 1). The interface is also easy to use: if you know what book you want, then you only need to type the title into the search bar at the top right-hand corner of the page. If the IA holds it, then it will come up (Fig. 2). Also worthy of note is the ‘Advanced Search’ functionality: it is possible to narrow the results down to individual years, or keywords. Unlike some other online archives (including at least one subscription one which I know of), the IA allows users to export entire out-of-copyright books as PDFs, rather than simply view them, or just download a limited page range. The option is available in a panel when scrolling down the page. I have used this feature to download copies of all of Gardner Wilkinson’s books—including multi-volume works such as his *Manners and*

¹ Internet Archive. < <https://archive.org/> >, accessed 21.3.2021.

Customs of the Ancient Egyptians (1837), *Modern Egypt and Thebes* (1843), *Dalmatia and Montenegro* (1848)—amongst much else (Fig. 3).²

Had this resource not existed—excluding other online archives from consideration for a moment—then I would have had to buy reprinted copies of Wilkinson’s books from Cambridge University Press (at a price of £30 per volume) where available, and take photographs of books in archives for his more obscure printed works. As Wilkinson’s published books run to over 6,000 pages—by no means all of which are available as reprints—that that would be a sizeable number of photographs, which I would then have had to spend even more time organising. While noting that a lot of extremely good research was conducted in the pre-digital age, it is interesting to note what else becomes easier when you have digital copies of books immediately to hand (which you can annotate, and which are not at risk of being recalled by other library users). More cross-referencing is one possibility. So is spending more time on looking at the corresponding Wilkinson manuscript collection.

Not consulting the original copies comes at a cost, however. The most significant drawbacks to using the IA to read nineteenth and early twentieth-century books are those associated with materiality. At one level, reading a book published in the Victorian era on a screen is quite different to how it was originally read—an argument forwarded by John Berger in relation to paintings.³ Whether this makes a significant difference or not is debatable. The closest that I would come to making any sort of complaint about a book being abstracted by the IA is that you can all too easily lose any conception of its size: was a book ‘Octavo’ size and therefore read by a large audience? Or was it an ‘elephant folio’ which would hardly ever have escaped a scholarly library? These questions need to be asked (and answered) if any attempt to put a book in its context is attempted—yet it is only one I started to think about seriously after seeing several enormous volumes of antiquarian

² J. G. Wilkinson, *Manners and Customs of the Ancient Egyptians* (London, 1837); J. G. Wilkinson, *Modern Egypt and Thebes*. (London, 1843); J. G. Wilkinson, *Dalmatia and Montenegro* (London, 1848).

³ J. Berger, *Ways of Seeing* (London, 1972).

books—which I had previously only been familiar with through the IA—in ‘hard’ copy, in a research library.

What is of far more practical significance—at any rate to my own research—is that the paratextual material—maps and images—frequently fairs less well in the copying process than the text. This issue has affected my own research on several occasions, in relation to one of Gardner Wilkinson’s books, *Dalmatia and Montenegro* (1848)—a travel and regional history book on the southernmost regions of the Habsburg Empire, with forays into the contested borderland with the Ottoman Empire. It was only after studying this text for a year that I realised that the original edition included a fold-out map of the eastern Adriatic and Balkan coast. This feature had not been reproduced in the copy that I had downloaded from the IA—I had to rely on another website to see this. Finding out that the original book had a map was not a surprise—the publisher, John Murray often included them in regional and travel books—but at the same time my thoughts at various points during the past year were more along the lines of ‘It’s so frustrating having to use google earth—why doesn’t this book have a map?’. The question that I should have been asking was, of course, ‘I wonder if the original had a map?’.

The story with images is better, but has some of the same problems. One problem is that some images are landscape and it takes more effort to rotate them (they need to be downloaded first) than it would do for a book—a case where John Berger’s critique really does matter: it is all too easy to simply skip over them, or be lazy or simply glance at them and move on. This is a problem that can be easily solved, given a few moments. But there are also more serious issues: some images in scans are blurred or otherwise distorted (and on a few occasions, I have even found that whole pages are missing). Usually this has not been too much of a problem, but there have been times when it has been an issue: I still recall one supervision in which myself and my supervisors disagreed as to whether a particular image of a track on the hills above the plain of Tzetinie and Lake Scutari (in Montenegro) showed a man on a donkey or not. It was hardly a major issue, but was one which would not have occurred had I been using a hard copy rather than an inadequately-copied image.

Whilst I have dwelt on some of the drawbacks involving the IA—in part because these are not often discussed—they definitely should not suggest that the IA is a resource with major problems. Even when it comes to research on materiality, there are advantages. Unlike the digitisation produced by some scanning technology—such as ECCO—the images of book pages are usually of a high standard: annotations and marginalia—now an area of major interest amongst scholars—come out well, allowing signs of reader engagement and contemporary responses to be studied.⁴ On several occasions I have even come across a reader noting the full name of an author, where only initials were printed—a potentially invaluable piece of help—although one of course which depends on the copy being scanned: if copiers choose the books in best condition, with clean pages, then historians interested in this angle may not even be aware of a lost opportunity. Especially during the last year, the choice has been, far more often than not, either to use a digital copy of a book (especially a nineteenth-century one), or make do without it altogether. Other options exist: the *Hathi Trust*, covers a similar time period. *Early English Books Online* features books from the late medieval and early modern periods. For me though, the IA has quite literally been the most important website on the internet.

⁴ H. J. Jackson, *Marginalia: Readers Writing in Books* (New Haven CT, 2001); H. J. Jackson, *Romantic Readers: The Evidence of Marginalia* (New Haven CT, 2005).

Figure 1: The Internet Archive homepage.



Figure 2: Results for *Manners and Customs of the Ancient Egyptians* (several volumes and editions).

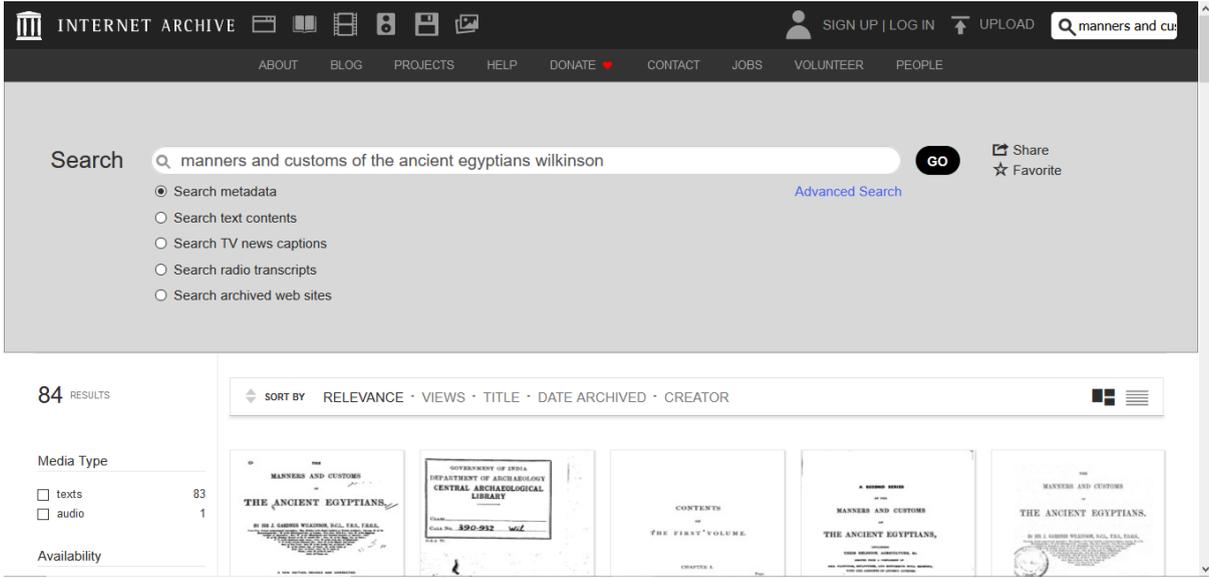
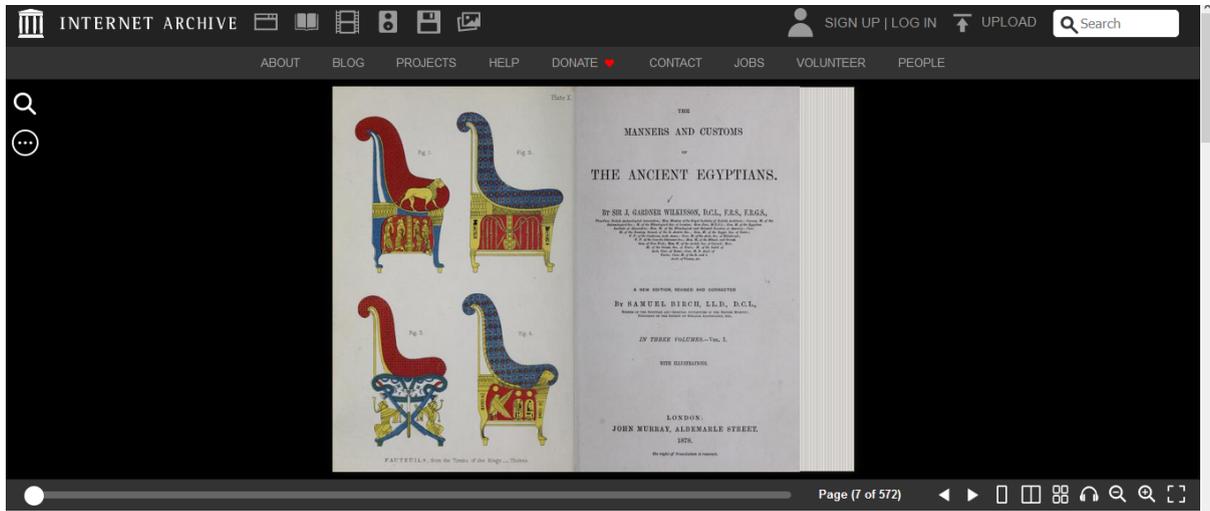


Figure 3: Wilkinson's *Manners and Customs of the Ancient Egyptians*—a navigable book, which is also available to download.



Bibliography

Berger, J., *Ways of Seeing* (London, 1972)

Jackson, H. J., *Marginalia: Readers Writing in Books* (New Haven CT, 2001)

Jackson, H. J., *Romantic Readers: The Evidence of Marginalia* (New Haven CT, 2005)

Internet Archive. < <https://archive.org/> >, accessed 21.3.2021.

Wilkinson, J. G., *Manners and Customs of the Ancient Egyptians* (London, 1837)

Wilkinson, J. G., *Modern Egypt and Thebes* (London, 1843)

Wilkinson, J. G., *Dalmatia and Montenegro* (London, 1848)