

Conversation on Digital Archiving Practices

With Geert Lovink

By Davide Giorgetta & Valerio Nicoletti

DG/VN: The scanning of texts — for instance of out-of-print books — and their subsequent storage as digital files has developed a new figure: the amateur librarian. Which are the features, responsibilities and limits of this figure?

GL: I do not think in terms of amateurs vs. professionals. I'd prefer terms such as '[geniale dilletanten](#)'. At our applied science school here we train students to become professionals. In our emerging internet context there is a battle about the name and description of these professions, which in turn start to appear in job adds (think of SEO professional, app designer etc.). These days, it is more and more uncertain what the role of education should be. Will you obtain a set of (technical) skills such as programming or is higher education there to offer eternal universal knowledge (which can then be applied to a variety of fields)? In short, liberal arts or professional education? This is relevant if we want to find an answer to your question. One could say that all people who work with information are amateur librarians: books, records, films, we all download them, day and night. Every computer is a library. We collect music, text files, e-books, films, photos. In the case of personal portable offline libraries (as described by Henry Warwick in his *Radical Tactics of the Offline Library* text) we only say that this is done in a systematic and somewhat larger scale... 50.000 e-books, 150.000 songs, 8000 films etc. Yet, these numbers no longer impress us. We want specific knowledge, such as the librarian and the archivist had in the past. Is this work these days done by amateurs or professionals? We don't know and in a way it doesn't matter. An archive is a materialised condensation of memory. What's interesting in this context is the fact that archives attract other archives. An archive will, in most cases, at some point become part of another, bigger archive. So what starts off as a personal initiative may very likely become part of an official reality. This is no different in the context of digital books. And with such a centralization comes expertise... and new professions.

DG/VN: Marcell Mars has written [a vademecum for possible contributors of this amateur librarianship](#). Are there some features that define the ideal digital book, from the operative perspective of a librarian, eg. file formats and tools to be preferred, how to organize files and metadata, how to

manage distribution and access to content?

GL: Right now, the digital book universe is exploding, not so much from the usage perspective (Nicolas Carr is right here), but in term of content. The simple reason for that is that we're reaching the end of the digitization phase. Over the past decades lots of investments were made in terms of scanning. Scholars in my field have come to know this because of the 'digital humanities', a top-down meta-discipline that was invented to deal with the question what to do with all the digitized documents (recently renamed as 'big data'). Now that there is plenty of data and tools to analyse and visualize them, we start to see the rise of boutique data collections that are neither open nor closed. That's where someone like Marcell Mars comes into play. The problem in this field that we have is the proliferation and constant change of e-book formats of a growing number of non-compatible devices of competing players from Google, Apple, Amazon and Microsoft to Samsung, Ubuntu and Chinese hardware manufacturers. E-pub is supposed to resolve all this but many fear that HTML-5 is yet another possibility, not the 'standard to end all standards'. In that sense it is still 1995 in e-publishing land—and maybe we get stuck there.

DG/VN: In which way, if any, are digital archiving practices (such as Monoskop Log, Aaaarg.org etcetera) stimulating new publishing phenomenons? Are there any innovative outcomes, apart the obvious relation to p.o.d. tools?

GL: I hesitate to use business terms such as innovation in a context like this. Rather, I think it is '[data dandyism](#)', a phenomenon our Adilkno wrote about twenty years ago and that only recently started to manifest itself with the phenomenal rise of i-pods, smart phones and tablets. Exclusive data collections do not work according to the social media logic. The way to show off with this information is much more indirect, and hidden, in the seclusion of the sub-culture. Why should you share that on such banal platforms such as Facebook or Twitter? We all know that such one-off collections have not been built with the interest of the individual author in mind. I compare them to 1980s record collections that are now out there, in the open. We must emphasize—and celebrate—the idiosyncratic nature and the civil courage of its founders and maintainers to put them out in the public domain. You cannot compare them to the Wunderkammer. It is more useful to compare it to the specialized libraries of the mid 19th-20th century, founded by influential scholars, art lovers and rich industrialists. It is tempting to

compare them to the large universal collections such as the Library of Congress or the British Library but I do not find that productive. In comparison they are tiny parasites that lack the ambition of, let's say, Google Books. It is important to stress freedom and mobility here. For Henry Warwick and me they must be personal and portable, with a possibility to still function in a strictly offline environment. The offline mobility is only truly innovative aspect.

DG/VN: Google Books launched a mass digitization program. Keeping it simple, it on the one hand provided access to public domain of out-of-print publications, which are lately searchable with a full text functionality. On the other hand it may eventually result in a corporate monopoly over human knowledge. What's your perspective on Google Books?

GL: I can recommend everyone to look at the 2013 BBC documentary [Google and the World Brain](#). The film is not yet outdated, with the exception of the last part, which is all about the court case and the failed settlement with the American Authors Guild. Despite all the legal battles, Google Books is still out there, the venture recently announcement to large digitization projects, here in the Netherlands. Here at INC we're extensively studying the 'evil' intentions of do-go companies like Google. It's impossible for me to repeat or summarize that work here. If you're interested, have a look at Ippolita' Dark Side of Google. In the Society of the Query program we're mainly looking at the search engine but I would love have separate research networks that look specifically into the politics of Google Books (or Google Maps, for that matter). Of course it is all about the creation of 'natural' monopolities which can then start to monetise their content. Think of extensive print-on-demand services, e-book delivery for tablets, combinations with high content zones such as YouTube (think of embedded videos in school text books). In the end, it is all about searchability. This goes way beyond the old print-web dichotomy. Google Books is not about digital versus analogue.

DG/VN: Do you notice any unresolved or raising questions in the contemporary context of digital archiving practices and their relation to the publishing realm?

GL: Whatever happens, try to stay away from the paper vs. digital debate. Digital archiving aesthetics should establish itself as a vital field of study as soon as possible. After digitization and the DRM disaster, the real challenge is findability. This is only in part covered by issues such as metatagging. In the end, metatags is human information, it is a process which cannot—and

maybe should not—be automated. This turns the metatag into an all-too-human digital object: interesting but marginal. Full-text or full-image search, on the other hand, is too messy. It is causing info smog of the worst kind. On the flip side, we also know about the hidden politics of search results. I am not interested in what others like. Why should that be the defining criterium for knowledge production? The herd mentality of Google cries for a devastating Nietzschean critique. What we need is a next generation of singular search engines, much like the libraries that we discussed above.

—

Davide Giorgetta & Valerio Nicoletti are researcher at DaISIA Urbino in Italy (<http://www.isiaurbino.net/>) and involved in a thesis about digital archiving practices in relation to the publishing domain. They are interested in the way both amateur-diy tools (e.g. diy scanning) and curated online libraries (e.g. Monoskop Log, Ubuweb etc.) relate — consciously or not — to the current state of the art of publishing.