

Web archiving in Tunisia post-2011: The National Library of Tunisia's experience

Raja Ben Slama

Abstract: The National Library of Tunisia had undertaken the digitization of written heritage and the "heritagization" of digital documents. To meet the second requirement, a web archiving unit was created outside the legal deposit service, as the organic law regulating this procedure only provided for the voluntary deposit of digital documents produced by publishers and authors, while social networks were becoming a virtual agora and a stage hosting several forms of literary and artistic creativity.

Keywords: BnT, heritagization, legal deposit, web archiving.

Tunisia's national library (Bibliothèque nationale de Tunisie – BnT) began to digitize its written heritage for preservation back in the 2000s and, from 2016 onwards, to offer free of charge, open access to Tunisian works already in the public domain. Digital technologies had proven their worth as a document preservation and dissemination tool. However, preserving native digital documents and *moving from the digitization of heritage documents to treating digital documents as heritage*, was one step further for which the BnT was not really prepared. Firstly because its thirty or so librarians, most of them trained at the Higher Institute of Documentation (Institut Supérieur de Documentation), gained their degrees before a web archiving module was introduced in 2018 and, secondly, because the Organic Act n°2015-37 of 22 September 2015 on registration and legal deposit makes no reference to this particular preservation method. Clause 4 of this act states that "Works of the types listed below that are made available to the public, in any form and on any medium, must undergo the registration and legal deposit procedures:

- all written matter and printed, engraved, illustrated, audio, audiovisual or multimodal documents, drawings, maps, photos, digital artworks, abstract words, and other content aimed at the general public
- software packages, databases, and the associated web pages and news sites."

However, the act does not specify the authorized body with whom websites are to be 'deposited' and makes no mention whatsoever of web archiving in the sections relating to the BnT. Despite its quest to be exhaustive, its openness to the new reality of document and artistic production, and its implicit abolition of the prior screening of publications,

Raja Ben Slama, Manouba University, Tunisia, raja.benslama@flah.uma.tn, 0000-0003-3059-7154

Referee List (DOI 10.36253/fup_referee_list)

FUP Best Practice in Scholarly Publishing (DOI 10.36253/fup_best_practice)

Raja Ben Slama, *Web archiving in Tunisia post-2011: The National Library of Tunisia's experience*, © Author(s), CC BY 4.0, DOI 10.36253/979-12-215-0413-2.05, in Sophie Gebeil, Jean-Christophe Peyssard (edited by), *Exploring the Archived Web during a Highly Transformative Age. Proceedings of the 5th international RESAW conference, Marseille, June 2024*, pp. 37-42, 2024, published by Firenze University Press, ISBN 979-12-215-0413-2, DOI 10.36253/979-12-215-0413-2

it simply places digital in the same media category as books. It falls short of the new socio-political and cultural reality brought about by the internet revolution and the closely associated Tunisian revolution.

Admittedly, by inventing the notion of *digital heritage* and encouraging the preservation of digital documents, the 2003 UNESCO charter on The Preservation of Digital Heritage goes some way to fill this legal loophole. However, where it falls short is that native digital documents are not elevated to the same status as classical heritage documents in every country. This involves a lengthy, almost involuntary process and there is generally a gap between manuscripts and books, on the one hand, and what archivists place in the 'non-book documents' category. Digital documents are short-lived, hence the need to archive them, but because of this ephemerality they are unfavorably compared to the great intellectual works which, when mentioned in Arab literature, are often accompanied by the adjective 'immortal' or considered to be the 'mothers of books'. It would take a shift in politics and document procedures for web archiving to become necessary and possible without the use of force. And, for Tunisia, the turning point was the revolution that broke out in the country on December 17, 2010. Triggered by a traveling salesman named Bouazizi setting himself on fire, it brought down the Ben Ali regime on January 14, 2011. How did this political event become a document and archiving issue without which web archiving could not have been contemplated?

Before the revolution, the executive government was using the legal deposit procedure for the purposes of prior censorship, thus preventing the recognition and circulation of the books published by opponents. The Head of the Legal Deposit Office at the BnT was a member of a Reading Committee that sat within the Ministry for Internal Affairs. The censored books were kept in a locked cupboard and could not be consulted. It was as if they had been imprisoned. Because of the revolution, they were once again made available to readers. The Reading Committee was abolished, as was prior censorship. It was on its own initiative that the national library extricated itself from the control of the political authorities. The web archiving initiative would not have been possible had this not happened.

There are several factors that contributed to the web becoming a key political issue in Tunisia and to the Tunisian revolution taking on archiving and digital significance. Before the uprising was sparked in December 2010, there was a revolt against internet censorship that some of its instigators referred to as a "virtual revolution". In May 2010, seven months before the people's revolt, young digital natives organized a campaign against internet censorship after the government closed down several dozen websites and blogs. The campaign was known as Ammar 404, referencing the error code 404 (file not found) and linking it to the first name of a public figure

(Ammar Bouzwir) who had been contacted to obtain the release of Salah (another public figure), the imprisoned madman depicted on an old, early-20th-century postcard.

After this pre-revolution for the liberation of the internet, the cyber activists, who could get round the censorship by using proxies (there's a Bendirman song about this), used blogs and social media to garner support, primarily circulating videos of the victims of repression and capitalizing on the fact that the written press and TV channels were under full governmental control. As such, they were almost the only source of media coverage at the local and international level.

In 2016, the threat was no longer the issue of government censorship but the extreme fragility of digital heritage. It was established that the lifespan of a domain name was around 3.8 years. This brings us to the notion of loss of heritage and to the findings of a 10th-century poet and critic from Kairouan, Ibn Rashiq (d. 1064 A.D.), who wrote in the book *Al-Umdah*: "It is reported that there is more good prose spoken by Arabs than good poetry. Of this prose less than a tenth survives; of poetry less than a tenth has been lost." Ibn Rashiq's aim was to commend poetry and demonstrate that rhyme and rhythm were factors of permanence and preservation. Moreover, I demonstrated that this was over-optimistic, that much of his own work had been lost and that only a tiny portion of Medieval Arab-Islamic works survive, particularly those related to Tunisia which was a juncture of struggles, influences, and migratory movements (Ben Slama 2022). We wanted to counter this fear of loss and had noticed that social media was becoming a virtual forum and a stage for several types of literary and artistic creation and that fewer and fewer periodicals were being published and websites were taking over. So we decided to create a selective archiving unit and a web manual without using automated extraction software. Several months later, we heard about the digital heritage of the revolution and the need to archive it. "Stored away in mobile phones, on desktops, on the web, and in foreign TV archives, these digital sources are perishable. Casualties of piracy and falsification, dispersed among amateur video-makers and photographers, Facebook accounts, dissident websites such as *Takriz*, image-sharing platforms such as *YouTube* and *DailyMotion*, websites of bloggers or news sites such as *Nawaat*, these digital revolution documents were doomed to disappear. The situation is alarming and, because it threatens the preservation of the nation's memory, it has not failed to attract the attention of so many archivists and human rights defenders" (Ben Achour et. al. 2021, 16).

Two courses of action were required: the retrospective archiving of revolution-related materials and day-to-day 'current' archiving. So the BnT joined forces with the multi-disciplinary group formed in 2016 of several public institutions (the National Archives of Tunisia (les Archives

nationales de Tunisie), the Higher Institute of Documentation (Institut supérieur de documentation – ISD), the Higher Institute for the History of Modern Day Tunisia (Institut supérieur d’histoire de la Tunisie contemporaine – ISHTC)), associations (FEMDH, the Euro-Mediterranean Foundation of Support to Human Rights Defenders) and voluntary researchers to collect and archive these documents that were dispersed over the internet and on the phones of ordinary people.

Not only had the BnT extricated itself from the prior censorship that had been secretly carried out in collusion with the executive government, but it was also able to fully fill its preservation function again and found itself inheriting revolutionary content it had to collect. It so happens that the word for revolution in Arabic (*thawra*) is an anagram of heritage (*wartha*).

The man who took on the role of representing the library within the group, Faycel Hamdi, was a young opponent of the Ben Ali regime and one-time member of the Progressive Democratic Party that was founded in 1983. As he himself had played a part in the revolution, he knew the Tunisian web well and also the ins and outs of publishing on dissident blogs and social media. The National Archives of Tunisia had been tasked with pulling together the videos and photos so Faycel set about collecting caricatures, sketches, graffiti, songs, slogans, poems, tweets, and Facebook posts. The National Archives collected, dated, authenticated, cataloged, and saved 1,100 videos and as many photos. The parties got together on 11 March 2017 to take stock of the situation and came up with the idea of staging an exhibition of the “digital heritage, but also the documentary, artistic and satirical heritage of the revolution” entitled *Before the 14th, instant tunisien* (Archives of the Tunisian Revolution) and producing a trilingual catalog. The exhibition opened at the Bardo National Museum, under the patronage of the Head of State, and ran from January 14 until March 31, 2019. It opened at the Mucem (Museum of European and Mediterranean Civilizations) in Fort Saint-Jean in Marseille on March 19, 2019.

The internet holds documents with a limited lifespan that need to be archived and saved, but it also conceals a matrix of things that no longer exist in real life. For instance, the walls of the sumptuous villas of the family and stepfamily of President Ben Ali were defaced and covered in humorous revolutionary graffiti. Unfortunately, to the great dismay of the ‘archivists’, the authorities confiscated and repainted these walls. However, there was still some evidence of this graffiti on the web and we were able to reproduce at least part of it.

Once the revolution heritage had been archived, the TnT undertook other targeted initiatives, among them the archiving of the 2019 *hirak* movement in Algeria, given its African and Maghreb remit.

During lockdown, 143 webinars (a total of 194 hours) were recorded.

Manual, selective daily web archiving began on November 1, 2016. The decision was made to archive the opinion pieces published on the news sites and the Facebook and Twitter pages and accounts of intellectuals, activists and opinion leaders and also caricatures and photos with informative and/or artistic value that were in the public domain. To reconcile copyright, privacy protection, and preservation, the following restrictions were agreed:

- Only public (and not private) posts would be kept. For Facebook, for example, the security setting must be public and not friends only.
- Until there was clear legislation in place, the information would not be made open access. It would be available for consultation at the library.

On February 15, 2017, a further two archivists joined the group and the number of websites, blogs, and Facebook and Twitter pages and accounts scrutinized each day increased. According to the group's 2022 report: around 120,000 documents have been saved, equating to 88.6 GB (gigabytes) of files. These documents have been archived as non-editable files and saved in two locations:

- Storage arrays for long-term retention,
- NAS servers for ease of access and use.

I remember the debate surrounding the technology to be used to catalog the documents. A technical solution was to be provided by the supplier of the new BnT platform that went online in May 2022. However, the archivists had already created a tree structure for the web archiving catalog with a menu linking the different categories (videos, caricatures, graffiti, posts, tweets, etc.) to the authors' names and a calendar, which provides the browser with several entries and consultation options.

Conclusion

Many documents archived by the BnT since 2016 had already disappeared from the net by 2023. Others had been changed, for instance the videos of comic singers had had the sound removed. The new challenges facing national, and more generally, archive libraries are the 'digital deluge'—and the volatility of its fallout—and the risk of partisan manipulation that has grown with the progress of artificial intelligence. These will change their role and *raison d'être*. In the past, librarians fought a constant battle against book-destroying insects (cockroaches, termites, booklice, etc.). The threat today comes from faceless, unspecified protagonists in a rapidly growing and increasingly sophisticated document-related environment.

References

- Ben Slama, Raja. 2022. “Parmi les vestiges du patrimoine écrit”, Inaugural talk of the Rencontres humanités numériques, *Insaniyyat Tunis*, Cité de la Culture, Thursday 22 September 2022. Translated by Kmar Bendana. <https://hctc.hypotheses.org/3732>
- Ben Achour, Rabâa, Bel Haj Yahia, Fathi, Ben Mhenni, Sadok, and Mkada, Amina. 2021. *Before the fourteenth, Instant tunisien*, directed by Rabaa Ben Achour, Tunis: Ministry of Cultural Affairs.