

Claude D. Conter:

Thinking the data of the future.

Webarchiving and Ephemera as challenges for
National Libraries

Annual General Meeting 2022

19 - 21 June 2022

National Library of Türkiye, Ankara

Thinking the data of the future.

Webarchiving and Ephemera as challenges for National Libraries

Claude D. Conter (National Library of Luxembourg, 20.06.2022)

The Luxembourg journalist and author Georges Hausemer wrote the humoristic guide *What You Always Wanted to Know About Luxembourg but were afraid to Ask*, whose title naturally recalls Woody Allen's movie *Everything You Always Wanted to Know About Sex (But Were Afraid to Ask)*. Hausemer, however, does not devote his not entirely serious book to the sex life of Luxembourgers, but paints a picture of the Grand Duchy, from bizarre peculiarities to necessary information on politics, history, culture, etc. He wrote one chapter to the epicurean delights of the Luxembourgers, and I quote: "Luxembourg is considered the land of gourmets and gourmands - with the well-known consequences. At least the Luxembourgers are not the only ones struggling with obesity."

It is not surprising that food has become a subject, even a topos, and sometimes even a cliché in the literature and culture of the Grand Duchy. I have chosen this simple example as a starting point in order to ask to what extent national libraries as traditional data providers can give answers to the demands of research. Imagine that a cultural historian would like to write an everyday history of gastronomy in Luxembourg, for instance the history of bars, bistros, hotels, pubs and restaurants from an architectural, social and cultural point of view.



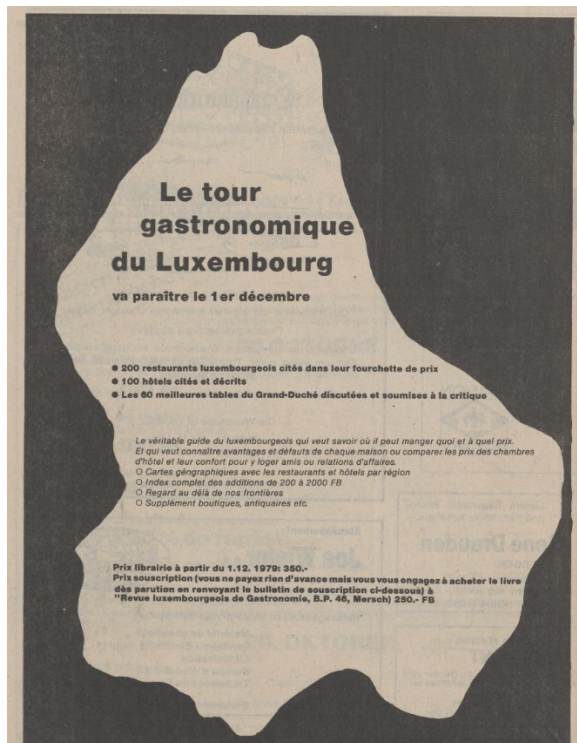
(BnL)¹

Such a question is neither unusual nor exceptional, but let's just imagine how a researcher is approaching the National Library to write a Luxembourgish counterpart to Éric Birlouez' *Histoire de la cuisine et de la nourriture : du menu des cavernes à la gastronomie moléculaire*, to Peter Peters *Kulturgeschichte der österreichischen Küche* or *From Kebab to Čevapčići. Foodways in (Post-)Ottoman Europe*. Ed. by Arkadiusz Blaszczyk and Stefan Rohdewald.

¹ https://luxemburgensia.bnl.lu/cgi/luxonline1_2.pl?action=fv&sid=cpluxvim&vol=16&page=7365r&zoom=3 & [Luxemburgensia online \(bnl.lu\)](https://luxemburgensia.bnl.lu)

What can a national library do if someone wanted to write such a study on the history of gastronomy, especially if it does not have a specific food history collection like our colleagues from Scotland?

Some necessary documents that are traditionally found in national libraries, as they are subject to the legal deposit, are restaurant guides and food journals. There may also be specialist literature. Of course, advertisements and notices in the daily and weekly press will also be informative.



(BnL)²

But you will soon realise that you need different documents, which you will only find to a limited extent in national libraries because they do not appear in traditional collections. I am convinced that two types of documents and data deserve our special attention: printed ephemera and web pages.

Let me start with printed ephemera as a data source of the future. They are already collected by some national libraries, for example by the British Library, the BnF, the BNE and at the Bibliotheca Narodowa. Especially in Anglo-American libraries, ephemera collections are common. Moreover, some ephemera, such as posters, postcards and calendars, are already collected on a broader scale as they are an integral part of collection policies via the legal deposit. So, depending on the definition of legal deposit, researchers will find in their respective national library posters of gastronomic fairs, restaurants, hotel management schools, etc., or postcards of restaurants, refectories and canteens, for example of hospitals, schools, prisons, etc.

² <https://persist.lu/ark:70795/3srg9sgzt/pages/48/articles/DIVL3249?search=gastronomie> & <https://viewer.eluxemburgensia.lu/ark:70795/ktcmx9/pages/8/articles/DTL807?search=gastronomie>



(BnL)³

But beyond that, there are other printed documents that are less common in national libraries at this point in time: bar receipts, restaurant letterheads, order coupons, menu cards or even just advertising leaflets like this lovely one. It is a small booklet you can flip through

A young man and his wife love each other tenderly, but one day they had an argument. You see their sorrow was immense and they shed bitter tears. Was this the end of a great love? But no, look! Turn the page. Riesling Perlé was their match. A renewal of joy animated their union.



(BnL)⁴

³ BnL. Collection of postcards

⁴ BnL, Collection of postcards

A researcher will only be able to gather documents to a limited extent. They will hardly find an essential source for the history of gastronomy: I am talking about menu cards like this one of the Court of the Grand Duchy from a wedding of Alix de Luxembourg with Prince Antoine de Ligne or that of the dinner at the occasion of the visit of the President of Estonia.



(BnL)⁵

Actually, a menu represents a central object of scientific research, for example, if you want to find an answer to the question of when an internationalisation of gastronomy took place in Luxembourg and what forms it took. Even if it seems obvious to correlate migration processes, for example, with the enrichment of menus and the specialisation of restaurants according to national cuisines, it is not enough to evaluate trade registers and statistics on the labour market and economic performance. At some point, you cannot avoid evaluating the menus themselves. These are not subject to legal deposit and will not appear in your library collection, unless the archiving of the web is one of your tasks. In fact, webarchiving will give you information about food and drinks, about prices and price trends, about nutritional developments, for example towards vegetarian or vegan food, about the connection between nutrition and health, for example if allergens are indicated, but also, of course, about the history of a restaurant as well as about graphic presentation methods, for example, by answering the question of whether internationalisation on a menu takes place via translations or

⁵ BnL Collections of menu cards

photo illustrations. Here you see a threeilingual menu of a Spanish restaurant and a photo menu of a Turkish restaurant.

Tapas

Para empezar o compartir
Pour commencer ou partager
To start or to share

PULPO A LA GALLEGA
POULET A LA GALLEGA
OCTOPUS A LA GALLEGA
13,7€

PULPO AL AJILLO
POULET A L'AIL
GARLIC OCTOPUS
13,8€

POLO AL AJILLO
POULET A L'AIL / GARLIC CHICKEN
12,9€

ALMEJAS CON CHORIZO Y AJO
PALOURDES AU CHORIZO ET L'AIL
CLAMS WITH CHORIZO & GARLIC
13,7

GAMBAS AL AJILLO
SCAMPI A L'AIL
GARLIC SHRIMPS
12,8€

PAN CON TOMATE
PAIN A LA TOMATE / TOMATO BREAD
0,9€

CHAMPIÑONES AJO Y TOMILLO
CHAMPIGNONS A L'AIL ET THYM
MUSHROOMS WITH GARLIC & THYME
9,6€

ALBONDIGAS CON SALSA DE TOMATE
BOULETTES DE VIANDE AVEC SAUCE TOMATE
MEATBALLS WITH TOMATO SAUCE
10,4€

CHORIZO A LA PLANCHA
CHORIZO SAUTE
SAUTEED CHORIZO
9,6€

ENSALADILLA RUSSA
SALADE RUSSE
RUSSIAN SALAD
9,8€

CHIPIRONES CON AIOLI
MINI CALAMARS FRITS
SMALL FRIED CALAMARI
12€

GAZPACHO DEL BARRIO
SOUPE FROIDE ESPAGNOLE
SPANISH COLD SOUP
1,8€

PISTO DEL DIA CON HUEVO FRITO
RATATOUILLE DU JOUR AVEC OEUF FRIT
COOKED CHEF'S VEGGIES WITH FRIED EGG
8,6€

HUEVOS ROTOS
OEUFS CASSES
BROKEN EGGS
12,9€
Add truffle + 75€



NAVAJAS A LA PLANCHA
COQUELLES A LA PLANCHA
RAZOR CLAMS A LA PLANCHA
13,9€

TACOS CON WAGYU BEEF (2 pcs)
ONION, CORIANDER, JALAPENO & TRUFFLE MAYO
OIGNON, CORIANDRE, JALAPENO
ET MAYO AUX TRUFFES
23,4€
*Add truffle + 75€

MONTADITO DEL MAR
ATUN ROJO, TOMATE Y AGUACATE
THON ROUGE, TOMATE ET AVOCAT
RED TUNA, TOMATO & AVOCADO
12,70€ (2PCS)

TORTILLA DE PATATAS
OMELETTE AUX POT
POTATOES OMELETTE
7,9€
*Add truffle + 75€

PIMIENTOS PADRON Y CHIPS DE AJO
POIVRONS ESPAGNOLS ET SES CHIPS D'AIL
SPANISH PEPPERS & GARLIC CHIPS
9,8€

MONTADITO IBERICO
JAMON IBERICO, MANCHEGO, FIGUILLLO Y TOMATE
JAMBON IBERICO, MANCHEGO, FIGUILLLO ET TOMATE
IBERICO HAM, MANCHEGO, FIGUILLLO & TOMATO
9,8€ (2PCS)

MONTADITO BARRIO
FOIE GRAS, COMFOTA Y PIMIENTOS PADRON
FOIE GRAS, COMFOTE ET POIVRON ESPAGNOL
FOIE GRAS, CHUTNEY & SPANISH PEPPERS
15,90€ (2PCS)
*Add truffle + 75€

ESPARRAGOS BLANCOS Y
MAYONESA DE TRUFA
ASPERGES BLANCHES ET MAYO AUX TRUFFES
WHITE ASPARAGUS WITH TRUFFLE MAYO
11,9€

PATATAS BRAVAS Y DOS SALSAS
PATATES BRAVES AUX DEUX SAUCES
FRIED POTATOES WITH TWO SAUCES
8,9€

TACOS CON CARNE ASSADA
TACOS AU BOEUF BRAISE
SLOW COOKED BEEF TACOS
12,9€ (2PCS)

BACALAO FRITO CON SALSA TARTAR
FRITURE DE CABILLAUD ET SAUCE TARTARE
COD TEMPURA & TARTAR SAUCE
10,8€

PIQUILLO CON ENSALADA DE ATUN
POIVRONS FAROIS AVEC SALADE DE THON
STUFFED PIQUILLO PEPPERS WITH TUNA SALAD
11,6€ (3PCS)

Matbah

Kebab & Grillades

[Le Restaurant](#)

[La Carte](#)

[Commandez en ligne](#)

[Contact](#)

Kebab

Köfte

Grillades

Burger

Entrées chaudes

Entrées froides

Salade et Végétarien

Desserts

Boissons soft

Kebab



Sandwich Döner Kebab au poulet

5,50€



Dürüm Döner Kebab au poulet

6,00€



Assiette Döner Kebab au poulet

11,50€

5

Let me give you a random example to illustrate the problem. In 2000, a critic started the column *Surprises around a daily menu* in the weekly newspaper *d'Land* under the pseudonym D'Artagnan with his recommendations, including the *Batucada*, where you can enjoy Brazilian food.⁶ Among other things, he also states the internet address: www.surftel.lu. Since the National Library only started web harvesting in 2017, the references from 2000 cannot be examined. Of course, the website has been shut down, but at least the restaurant and a new address still exist, including a new URL address, so that after 22 years it is at least verifiable that the daily menu costs 12.90€ today, and the equivalent of 8.90€ back then. I also see from the last harvesting that a mojito costs 10.50€ in this restaurant, 11.90€ in a Portuguese restaurant belonging to the same trade group, 12.40€ in the Spanish restaurant, which also has the same owner. As you can see, menus are informative in many ways, and at least in Luxembourg, they are only accessible via the internet.

Surprises autour d'un menu du jour (1)

Les adresses à midi

D'Artagnan

DE NOMBREUX GUIDES, tels le *Ni vu ni connu* 2000, fournissent certes une orientation, mais ne se livrent pas à une critique gastronomique que le consommateur est en droit d'attendre. Ces guides servent en quelque sorte de répertoire, ils sont en général bien faits, l'on retrouve vite les jours d'ouverture et une courte description des spécialités des restaurants.

D'Artagnan étant souvent amené à passer les heures de midi en ville et conscient que beaucoup de gens en font de même, il s'est proposé de faire un tour des tables à midi en essayant de faire un inventaire des bonnes et moins bonnes adresses.

Le Via Sud: ambiance provençale, mais cuisine décevante

Suite à une recommandation d'un ami américain, d'Artagnan essaya l'autre jour le restaurant Via Sud situé au 22 - 24, rue du Curé. Comme entrée, d'Artagnan se composa sa salade selon ses désirs à un prix raisonnable. Comme menu du jour, le chef afficha la brochette de dinde. Pensant qu'à priori, une brochette de dinde devait être correcte et se laissant blinder par la terrasse sympathique, la déception fut d'autant plus grande. D'Artagnan croqua les premiers morceaux et demanda de suite à la serveuse si elle était sûre qu'il s'agissait de dinde. La réponse fut évi-

Les menus et plats du jour proposés par nos restaurants à midi ne se valent pas tous

demment « oui, j'ai demandé à la cuisine, c'est bien de la dinde ». O.K. se dit-il, tout en répondant que c'était franchement pas bon.

D'Artagnan observa les grimaces de sa voisine en croquant ces délicieux morceaux de dinde qui ressemblaient plutôt à une sorte de viande troisième choix qu'à un bon morceau de dinde et remarqua qu'elle aussi semblait avoir fait le « bon choix ».

D'Artagnan et son accompagnatrice, qui partagea d'ailleurs son appréciation, ne voulaient plus de dessert, ce qui n'était pas une bonne idée puisque d'Artagnan commença déjà à avoir faim vers 16 heures.

Quelle leçon en tirer ? Bien que de nombreux amis l'aient averti de la médiocrité de ce restaurant, d'Artagnan voulut en faire l'expérience personnelle à midi. À déconseiller, en tout cas pour les menus du jour.

Le Batucada : ambiance et cuisine brésilienne

Le Batucada est situé au Plateau du Saint Esprit et pour tous ceux qui ne l'ont pas encore essayé, il gagne à être découvert. Dans un cadre totalement décontracté à la décoration latino-américaine, la serveuse toujours souriante vous sert des menus succulents.

Le Batucada propose tous les jours à midi un seul menu, assorti normalement soit d'une entrée, soit d'un dessert. Les menus sont toujours très copieux, à remarquer que le chef ne se contente pas de les assortir que de frites ou de salade, mais qu'il propose une variété de légumes. Au prix de 350 francs pour le menu à midi, D'Artagnan pense que ce restaurant est certes à ranger parmi les meilleurs en ville au niveau rapport qualité/prix.

La carte, très variée est aussi à essayer quoique les prix sont relativement élevés. À noter que l'ambiance latino-américaine est en quelque sorte cassée par le choix des desserts en provenance de la maison Namur. Regrettons que le chef n'ait pas su faire preuve de plus d'imagination, ce qui ne veut pas dire que les desserts proposés ne soient pas bons ! Il est recommandé de réserver à midi.

Restaurant Batucada ;
1, Plateau du Saint Esprit ;
Tél: 26 20 08 80
www.surftel.lu

⁶ <https://persist.lu/ark:70795/5324mt/pages/11/articles/DTL242?search=gastronomie>

I could give more examples, but let me abstract: Websites and documents on the Internet form the orientational knowledge which we need in our daily life. Researchers, studying the understanding and self-reflection of our time, be it from a sociological, historical, political, cultural, or any other point of view, will not be able to avoid using websites as a primary source of information. I am sure everyone will have to agree that, for the outgoing 20th and the beginning of the 21st century, these sources of information have replaced the encyclopedia in the 18th century. To avoid any misunderstanding: no, the Internet doesn't offer the same framework of structured and verified information as it was the case for the likes of Diderot and d'Alembert, but the importance the Internet has attained as a source of information, seems to me to be comparable to the most eminent knowledge project of the Enlightenment. Scientific research into the post 2000 era might seem unfeasible, even pointless without web sources. As Ian Milligan pointed out: It would mean "ignoring the revolution in communications technology that fundamentally affected how people share, interact, and leave historical traces behind."

Therefore, web archiving is a prerequisite for historiography in general. Nevertheless, as Jane Winters pointed out: „web archives remain largely an unknown, and certainly underused, primary source”.⁷ Actually, legal restrictions, as well as technical challenges seem to be the biggest obstacles in this regard. Most institutions still face the paradox that archived versions of non-commercial websites, which are freely accessible around the world can only be made available from within the walls of the archiving institution. We might therefore have to try new legal approaches: Let us think outside of the box and demand that the contents of the Web should not be subject to a strict interpretation of copyright limitations. Don't get me wrong, copyright should not be abolished and we should continue to respect the rights of creators, but how do we explain to our users that a page, which was available days ago or is even still accessible, can only be accessed in our reading room? If the archiving institutions can't follow the evolution of the services of commercial providers, they will lose the trust and ultimately the legitimacy of the archiving mission. Even leaving aside the legal aspects, the technical challenges remain, as pointed out by Hale, Blank and Alexander: „Confronted with this enormous amount of data, few tools exist to help scholars find information. Furthermore, web pages are not well structured or consistently structured, and they can be extremely difficult to transform into a format that can be used for large-scale quantitative research. In addition, changes in web page format and changes in content often occur simultaneously.”⁸

Finally, I would like to give an outlook using the example of the National Library of Luxembourg. We took up web harvesting in 2016. We now run 4 crawls a year. For the moment we are talking about 232860 hosts, 3.5 billion documents and 356 TB. We are currently working on implementing a SOLRwaybackmachine, like some of you, f. e. in Denmark have already done, to provide a research

⁷ Jane Winters: Coda: Web Archives for humanities research – some reflections. In: *The Web as History: Using Web Archives to Understand the Past and the Present*. Ed. by Niels Brügger & Ralph Schroeder, UCL Press 2017. p. 238-248, here, p. 238.

⁸ Scott A. Hale, Grant Blank, Victoria D. Alexander: Live versus archive: Comparing a web archive to a population of web pages. In: *The Web as History: Using Web Archives to Understand the Past and the Present*. Ed. by Niels Brügger & Ralph Schroeder, UCL Press 2017, p. 45-61, here p. 58.

toolkit. Going back to my earlier example, research about gastronomy could be supported by BnL from 2016 onwards.

Concerning ephemera, we will build up a collection only in 2023/2024. There is already a collection of posters and postcards and there are also a few smaller collections such as diplomas and awards, letterheads, invoices and of course menu cards, but they are scattered in different collections, so I'm planning to merge them into one collection. The focus will be on questions concerning the collection development policy and defining the scope of the ephemera collection.⁹ We all know Maurice Rickards *Encyclopedia of Ephemera. A Guide to the fragmentary documents of everyday life* but the hundreds of categories he described must be reduced to some very specific categories of printed ephemera if we want the workflows, the storage capacity and the acquisition policy to remain operational. Advertising gazettes from supermarkets, such as the ones shown here from the same reference period, are, in my view, quite useful from an economic, social and cultural history perspective.



(BnL)¹⁰

We must raise the question to what extent ephemera can be integrated into legal deposit or should it just be part of the e-deposit. Posters and postcards are already mandatory copies. This is not the case for other documents. Therefore, in a first step, and we have already started with this, it is a matter of acquiring historical printed ephemera: contact with private collectors as well as auctions are the most

⁹ Based on the idea of Colleen Holescher and Jullian Ewalt: Separating the wheat from the chaff: Intensive deselection to enable preservation and access, 2016. <http://library.ifla.org/id/eprint/1568/1/100-hoelscher-en.pdf>

¹⁰ BnL

important acquisition ways right now. We have not yet had the good fortune to come across a Luxembourgish John Johnson, whose important collection has been integrated into the British Library. What will be important for the new Ephemera Collections is that library work will change: There will no longer be individual cataloguing, but rather archival-oriented cataloguing at the dossier level; the claim to completeness through legal deposit will be deliberately abandoned in favour of a curated, predominantly commercial acquisition of documents. We do not yet know whether ALMA will be the appropriate cataloguing software. We do not yet know whether and how we will digitise ephemera and make them accessible.

There is only one certainty for me at the moment: Web pages and ephemera belong to the important data of the future, even of the present. Not only will they complement our traditional collections, as some of them already do, but they will also have a special significance in the future. Web pages and ephemera are particularly exposed to oblivion because they are volatile, transient, or dematerialise quickly. Nevertheless, they are rich sources of information which are still underused by researchers because they are not yet systematically collected and made available. I firmly believe that these documents and date are actually at the heart of our society because they orient everyday actions and structure lifeworlds. In that sense, ephemera and web pages are relevant. We will have to take care of them if we, as national libraries, want to remain in the middle of society and remain relevant ourselves too.



(even the small crus give great drunkenness)¹¹

¹¹ https://luxemburgensia.bnl.lu/cgi/luxonline1_2.pl?action=fv&sid=cpvie&vol=3&page=12726r&zoom=3