

# REPORT

**Project Acronym:** Europeana Newspapers  
**Grant Agreement number:** 297380  
**Project Title:** A Gateway to European Newspapers Online

---

## D6.3.1 Report on Europeana Newspapers Information Day at the Bibliothèque nationale de France, Paris

---

**Revision:** 1.0  
**Author:** Joséphine Renard (BnF)  
**Contributions:** Jean-Baptiste Vaisman, Marion Ansel, Ioannis Anagnostopoulos,  
Elisabeth Freyre (BnF)

Project co-funded by the European Commission within the ICT Policy Support Programme		
Dissemination Level		
P	Public	x
C	Confidential, only for members of the consortium and the Commission Services	

## Revision History

Revision	Date	Author	Organisation	Description
0.1	09-12-2014	Joséphine Renard, Jean-Baptiste Vaisman	BnF	First draft
0.2	10-12-2014	Jean-Baptiste Vaisman, Ioannis Anagnostopoulos	BnF	Proofreading
0.3	12-12-2014	Marion Ansel	BnF	Revisions and addition of annexes
0.4	15-12-2014	Elisabeth Freyre	BnF	Revisions
0.5	16-12-2014	Friedel Grant	LIBER	Review and further editing
1.0	17-12-2014	Clemens Neudecker	SBB	Internal review and final version

### Statement of originality:

This deliverable contains original unpublished work except where clearly indicated otherwise. Acknowledgement of previously published material and of the work of others has been made through appropriate citation, quotation or both.

## Table of Contents

1. Executive Summary .....	4
2. Event Overview.....	5
3. Presentations and Discussions .....	7
4. Audience and Impact .....	17
ANNEX I: Agenda .....	20
ANNEX II: List of Participants.....	23
ANNEX III: Media Coverage .....	26
ANNEX IV: Survey results.....	31

## 1. Executive Summary

This report concerns the tenth and final Information Day of the Europeana Newspapers Project, organised and held by the National Library of France (BnF) in Paris. Information Days are a chance for partners to highlight the work they have been doing in the project at national level.

The Information Days increase overall awareness of the project. They should also convey the following key messages:

1. The value of collaborating at a national and European level to make newspaper content available online (social, economic, cultural, research, technical etc).
2. How to address the technical issues associated with making digitised newspaper content available.
3. The value of online digitised newspaper collections (i.e. highlight available content).

The event was attended by a considerable audience: 88 people in total, among them researchers, librarians, digital content creators, cultural heritage professionals, informatics professionals, archivists, audio-visual professionals and teachers.

The speakers at the Information Day were equally diverse. They included librarians from the Berlin State Library and the University Library of Innsbruck, digital and semantic content creators from CCS (Content Computer Specialists) and the company Syllabs, archivists, historians, publishing and IT professionals and, of course, researchers.

Thanks to this broad panel of speakers, the audience learnt more about many aspects of the Europeana Newspapers project and its possible applications. The day was also an opportunity for those within the project to gather more information on the expectations of researchers with regard to digitised historic newspapers.

Overall, the day was rated a success by those who attended. A survey circulated after the event showed that over 90% of participants rated the presentations as Good or Very Good, and said they intended to explore the content made available by the project (see Annex IV for the survey results).

## 2. Event Overview

The French Information Day took place on 27 November 2014 in Paris, in the Grand Auditorium of BnF.

The aims of the Europeana Newspapers Information Day were to introduce the project to the invited audience, and to learn how researchers and companies could use the mass of information generated (full-text, named entities, structured documents) by the project.

Consequently, there were two parts to the day (see Annex I for the full agenda):

1. After a general introduction, the morning session focused on the technical results of the project. These include improved access to content thanks to OCR (Optical Character Recognition), OLR (Optical Layout Recognition) and Named Entity Recognition (NER). Improvements regarding online access to content, text mode enrichment due to OCR and OLR as well as new browsing tools were introduced. The identification of sections, titles and subtitles as well as named entities (names of places, persons and organisations) and their implementation in relation to indexing were also presented.
2. The afternoon session focused on highlighting the impact of the Europeana Newspapers project for historic and scholarly research. Researchers who already use Gallica's newspaper collections and other digital libraries discussed their user experiences with this digital media. They were also invited to express their expectations with regards to the new access and research methods enabled by developments within the project and by its digital portal. Finally, a private company (Syllabs) explained how the outcomes of Europeana Newspapers could be worked into its services.

Throughout the day, the Europeana Newspapers project and the digitised historic newspaper content now available on Europeana and The Europeana Library was promoted.

Of the 123 people who registered, 88 attended the event (see Annex II for the list of attendees). Their backgrounds spanned a range of educational, non-profit and commercial organisations. They included librarians from prominent institutions such as the University Library of the Sorbonne, archivists from the Archives of the Département of the Gironde, teachers from CELSA (a Graduate School within the University of Paris Sorbonne) and several digitisation companies such as Syllabs, I2S, ECA, Arkhênum and Immanens.

All attendees received promotional materials, including a leaflet about the project and a promotional postcard printed with a French newspaper. Both were provided in French.

The event was mainly aimed at a French speaking audience and a translation service from English to French was available. The whole event was also recorded on video and, after editing, will be placed on the BnF website in early 2015.

This Information Day was organised within a thematic framework on digitisation issues and online access to newspapers. In addition to the Europeana Newspapers Information Day, two other events took place at BnF on the following day (28 November 2014):

- A workshop on *Newspapers of the WWII resistance at the time of digitisation: new readings and research perspectives*<sup>1</sup>
- The final conference of the European project SUCCEED named *SUCCEED in Digitisation, Spreading Excellence*<sup>2</sup>

It was noted that some participants attended both the ENP Information Day and either the workshop on Newspapers of the WWII resistance or the conference SUCCEED in Digitisation. To a certain extent, these three events had a combined audience.

---

<sup>1</sup>[http://www.bnf.fr/fr/evenements\\_et\\_culture/auditoriums/f.journaux\\_resistance.html?seance=1223917800203](http://www.bnf.fr/fr/evenements_et_culture/auditoriums/f.journaux_resistance.html?seance=1223917800203)

<sup>2</sup>[http://www.bnf.fr/fr/professionnels/anx\\_journees\\_pro\\_2014/a.jp\\_141128\\_succeed.html](http://www.bnf.fr/fr/professionnels/anx_journees_pro_2014/a.jp_141128_succeed.html)

### 3. Presentations and Discussions

The Information Day kicked off with a welcome and introduction from Sylviane Tarsot-Gillery, General-director of the BnF, who outlined the reasons for the BnF's involvement in this major initiative for research and access to primary sources such as newspapers.

She stressed that, thanks to the OCR and OLR technologies applied to the corpus, researchers will now have improved access to the digitised press and will find it easier to browse within the text content. Tarsot-Gillery noted that the work done on named entities in French had helped to promote the French language on the web, and that the BnF was particularly pleased with this fact.



Sylviane Tarsot-Gillery opening the Information Day in Paris

The programme then continued with 3 separate Round Table discussions, each with its own set of speakers and special area of focus.

## Round Table #1: Why Europeana Newspapers?

Moderated by **Elisabeth Freyre** from the BnF, this first round table set the scene for the rest of the day through three presentations, which gave a detailed view of the project, its activities and goals.



From left to right: Ioannis Anagnostopoulos, Hans-Jörg Lieder, Elisabeth Freyre and Philippe Mezzasalma

The first presentation from **Philippe Mezzasalma**, Head of the Press department at the Department of Law, Economics and Politics at the BnF, was entitled *The digitised newspapers collections at BnF: success and limits of the first digitisations*<sup>3</sup>.

Mezzasalma spoke about how the BnF became involved in Europeana Newspapers. He noted that before this project, the Gallica digital library contained mainly only major daily newspapers and offered a relatively shallow user experience. It did not allow content to be easily browsed, and searches in Gallica were often unstable and their relevance random.

<sup>3</sup> [http://www.slideshare.net/Europeana\\_Newspapers/pmezzasalma-enp-paris-27112014](http://www.slideshare.net/Europeana_Newspapers/pmezzasalma-enp-paris-27112014)

The work done via Europeana Newspapers had enabled an improvement in the OCR of the digitised newspapers as well as in the segmentation of the articles within each periodical. It had also helped to meet major expectations of the public, such as:

- Making it possible to identify images and captions more easily;
- Providing more reliable results for full-text searches;
- Giving greater relevance in terms of search results by article or type of articles;
- Allowing direct access to the article which is being searched;
- Enabling the ability to browse more easily within the content.

In addition, the project had allowed the BnF to prioritise certain titles that had not previously been available in a digital format. For Europeana Newspapers, the BnF had selected titles that no longer exist, those published between the 19th and 20th centuries, as well as national and regional daily and weekly newspapers. The BnF endeavoured to pick papers which would reflect all political and cultural shades of opinion in France. In other words, press coverage with an international theme, in tune with all the European countries involved in the project.

The second presentation from **Hans-Jörg Lieder**<sup>4</sup>, Head of the department of bibliographic services at the Berlin State Library, reminded participants of the scale of the Europeana Newspapers project. Lieder highlighted the 3-year duration of the project, the budget of over 5 million euros and the geographic spread of the partners, encompassing nearly all of Europe. Lieder pointed out that millions of pages have been digitised and refined using OCR, OLR and NER technologies and in total 950 titles have been digitised in 20 different languages. The oldest newspaper in the collection dates from 1618 and the most recent one is from 1955.

---

<sup>4</sup> [http://www.slideshare.net/Europeana\\_Newspapers/presentation-42715138](http://www.slideshare.net/Europeana_Newspapers/presentation-42715138)



Map showcasing the project partners, associated partners and network partners

He noted the uniqueness of the project's own content browser<sup>5</sup> in that it enables researchers to search through the full text of many newspapers, as well as to search by geographic area or browse by title. From early 2015, work will continue to increase the number of newspapers that appear when a search is performed by date. Work would also be carried out on the quality of automatic OCR, which was not yet 100% error-free.

**Ioannis Anagnostopoulos**, Coordinating manager of the Europeana Newspapers project at the BnF, set out the contribution of the BnF to the project<sup>6</sup>. He noted that Europeana Newspapers had:

- Enabled the aggregation of historical European newspapers;
- Raised visibility of the Europeana portal and of the portals of partner libraries;
- Brought in a notable improvement of search features, thanks to the semantic enhancement of the data and metadata relating to the selected corpus.

Anagnostopoulos then outlined the technical work of various project partners in some detail, including how the BnF has processed more than 2 million pages using OCR and OLR. He also described how the BnF had worked with other scientific and technical partners of the project (e.g. the National Library of the Netherlands, and the LIP6 laboratory at the Pierre and Marie Curie University) on Named Entity Recognition (NER).

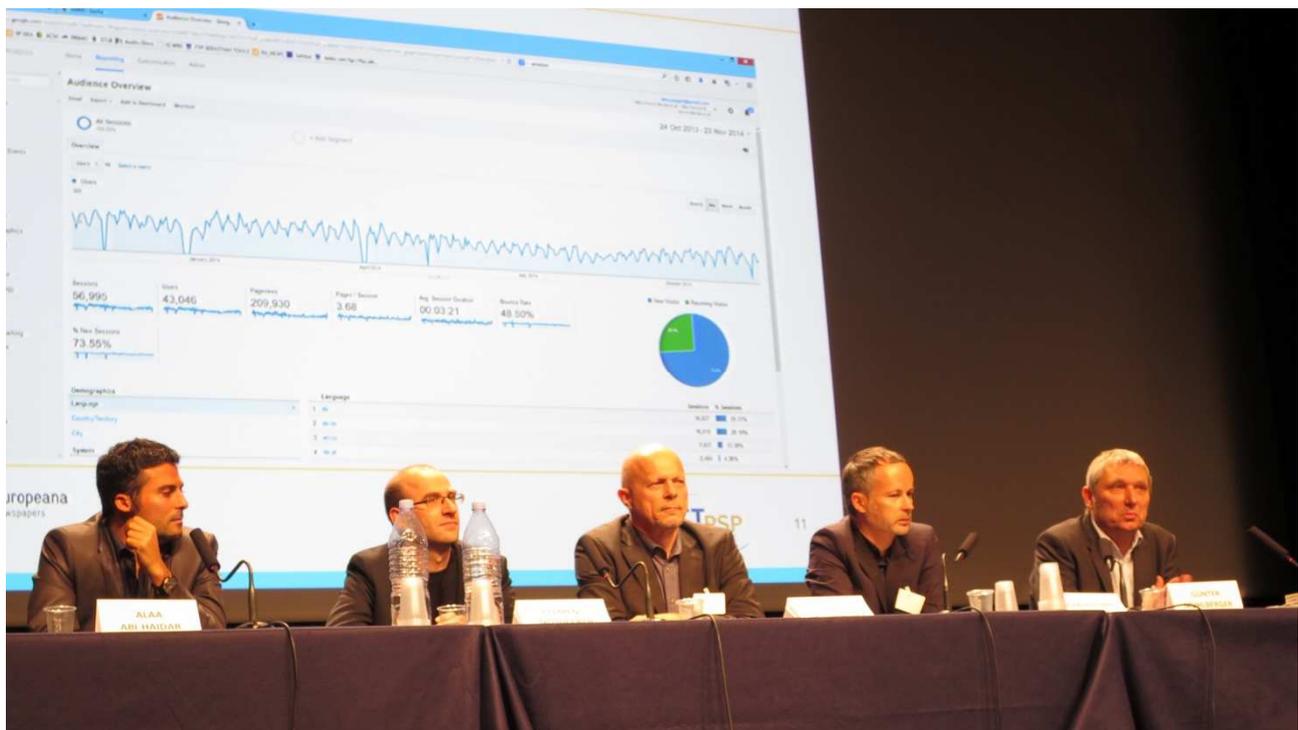
<sup>5</sup> [www.theeuropeanlibrary.org/tel4/newspapers](http://www.theeuropeanlibrary.org/tel4/newspapers)

<sup>6</sup> [http://www.slideshare.net/Europeana\\_Newspapers/presentation-of-the-bnf-at-their-information-day](http://www.slideshare.net/Europeana_Newspapers/presentation-of-the-bnf-at-their-information-day)

Finally, Anagnostopoulos stated that early studies showed the project represented an improvement to the research and viewing functionalities of the digitised press and would increase the number of visitors to Gallica.

## Round Table #2: Technical innovation helping users

The aim of the second part of the morning session, moderated by **Jean-Philippe Moreux** from the BnF, was to take an in-depth look at the technical innovations of the project.



From left to right: Alaa Abi-Haidar, Clemens Neudecker, Claus Gravenhorst, Jean-Philippe Moreux and Günter Mühlberger.

**Günter Mühlberger** is professor and project director at the University of Innsbruck, the university responsible for processing the 8 million pages of the Europeana Newspapers project using OCR. In his presentation *How can Optical Character Recognition technology help users in their research?*<sup>7</sup>, He described how the University of Innsbruck and the BnF have been working together for 15 years already, and how within the project they had collaborated on the metadata, OCR and the ALTO standard. He noted that currently the OCR had a rate of 83.4% of correct words in French, despite earlier fears that it would never be reliable enough to be used. Future new technologies may allow this error rate to be further reduced.

Mühlberger next focused his attention on the users. Some, he said, were occasional users motivated by curiosity but others such as researchers had a deep interest in the content and could spend hundreds of hours reading, downloading, analysing and sharing with other researchers. They were interested in more control and understanding of the content being searched. This was why the University of Innsbruck had done work to ensure that researchers could, for example,

<sup>7</sup> [http://www.slideshare.net/Europeana\\_Newspapers/presentation-gunter](http://www.slideshare.net/Europeana_Newspapers/presentation-gunter)

approve or reject the results of the OCR engine and had linked content to external sources such as Wikipedia.

**Claus Gravenhorst**, Director of strategy at CCS (the company which has processed 2 million pages of the project using OLR from 5 partner libraries), then spoke of the work done by CCS as part of the Europeana Newspapers project in his presentation ***Optical Layout Recognition: From unstructured to structured newspaper data***<sup>8</sup>.

Gravenhorst talked about the different stages involved in the digitisation of newspapers. According to the CCS workflow, once the digitisation is complete, the layout and structure of each page are analysed in order to determine its content. Then an examination of the data followed by a conversion of metadata enables the recognition of the structural details of the document. This data is then returned to the project partners. All of these steps are necessary to enable the user to obtain a maximum amount of information on the content of press articles for which they are searching.

**Clemens Neudecker**, coordinator of the Europeana Newspapers project at the Berlin State Library, then presented ***Development of Named Entities Recognition for French Newspapers***<sup>9</sup>. He noted that the main advantage for users is to be able to link named entities to other online resources such as catalogues from other libraries. He also spoke of ongoing challenges, for example, NER tools have to be adapted for each language in order to be able to interpret the sentences, grammar, context and structure. It is also possible that place names have changed over time, and this makes it necessary to use additional modules to map from ancient to modern forms.

**Alaa Abi Haidar**, a researcher in data mining and complex systems at the BnF and the LiP6 laboratory, rounded off this section of the Information Day with a presentation that focused further on the purpose and applications of Named Entity Recognition<sup>10</sup>. He noted that some issues such as the ambiguity of certain terms have not yet been technically resolved. Some words have, for example, more than one meaning (eg. 'Paris' could be the name of a city and a person), and some words are not contained in dictionaries. This gives an idea of some of the challenges which still need to be overcome.

Abi Haidar also explained the use of the UNERD method used to classify each word in a grammatical class. This method helps to determine the reliability of named entities and has been found to be very effective.

### **Round Table #3: What are the uses, practices, and expectations for the users?**

The afternoon session of the Information Day focused on the use of the digitised corpus in the Europeana Newspapers project and the expectations of users. The moderator was **Marc Minon**, director of Cairn, a platform for French-speaking periodicals. The objective of this round table was to demonstrate how the digital environment has changed the work methodologies of researchers and to show what role Europeana Newspapers can play in this changing environment.

---

<sup>8</sup> [http://www.slideshare.net/Europeana\\_Newspapers/enp-infoday2014-pptgravenhorst](http://www.slideshare.net/Europeana_Newspapers/enp-infoday2014-pptgravenhorst)

<sup>9</sup> [http://www.slideshare.net/Europeana\\_Newspapers/presentation-42715189](http://www.slideshare.net/Europeana_Newspapers/presentation-42715189)

<sup>10</sup> [http://www.slideshare.net/Europeana\\_Newspapers/presentation-ala-a-haidar](http://www.slideshare.net/Europeana_Newspapers/presentation-ala-a-haidar)



From left to right: Olivier Hamon , Sophie Kurkdjian, Philippe Tétart, Marc Minon and Laurent Martin

**Laurent Martin** holds many titles: among them a professor and lecturer at institutions such as the University of Paris III Sorbonne-Nouvelle, a researcher at the Integration and Cooperation laboratory for European environments (ICEE) and member of the History Committee of the Ministry of Culture. He talked about the methods used to produce his thesis on the history of the weekly publication *Le Canard enchaîné* and reflected on how current technology could have helped or even changed the approach to his subject. His project of producing a dictionary of the press in Europe in the 20th century will be an opportunity to test the usefulness of the tools currently made available by the project.

**Philippe Tetart** is a senior lecturer in history (University of Maine), a researcher and an expert consultant to the National Sports Museum. He works on the relationship between sport and media, sport journalism and the social applications of sport. He explained that the current technology was, for him, a turning point in research methods. The search process was now much easier than before and the digitisation of the press had allowed him to explore research subjects from a whole new stand point.

**Sophie Kurkdjian**, a doctor of history at the University of Paris I, is a researcher working with the BnF and an associate researcher with the IHTP-CNRS. She works on the history of the illustrated press, and particularly on the women's press of the beginning of the twentieth century. She explained the new opportunities that the digitisation of illustrated newspaper titles offered to her, including the ability to work on several titles simultaneously. This gives her a better understanding of the formal development of women's' illustrated newspapers and the issues related to this development.

**Olivier Hamon**, an engineer and a specialist in web semantics, works in the research and development unit of the company Syllabs. His contribution contrasted with that of the researchers, in that it was more focused on the proposed uses of the digitised press. He drew some early conclusions on the work carried out by Syllabs as part of another project on the historic press.

In addition to the views of the individual speakers, this round table provided the opportunity to better understand user expectations. It was felt that users would like to:

- Be able to perform statistical processes, such as to calculate numbers of words or occurrences of a word within a text;
- View this scientific research in the form of a graph;
- Retrieve isolated texts;
- Create hypertext themselves in order to be able to compile these statistics.

It was also noted that current tools and applications remain text-focused and do not help to analyse images. This is a pity for scholars with an interest in the satirical press or the illustrated women's press, where drawings are as important as text. Associating text with images may be something that can be achieved in the future but it is a complex process, which will require a large time investment.

A few people from the audience wanted to talk about the notion of errors that could be caused by digitising the press. It is a fact that the quality of the OCR has a considerable impact on searches. A user can have access to a text but if that text contains errors, this can hinder overall understanding of the text. Sometimes it is not the OCR itself which is at fault but rather a poorly digitised copy of the original paper, which makes it difficult to identify characters when OCR technologies are applied.

**Pascal Sanz**, director of the Department of Law, Economic, Politics at the BnF, used the Q&A session to point out that some famous French newspapers such as the *Dépêche du Midi* and the *Canard enchaîné* were initially willing to participate in the digitisation programs of the BnF but later withdrew due to copyright issues. For example, the illustrators of the *Canard enchaîné* were opposed to the digitisation of their drawings which meant that digitising the newspaper was completely pointless. Negotiations are currently underway to find other solutions.

## Conclusions

At the end of the day, **Olivier Piffault**, Director of the department of Preservation at the BnF, and **Pascal Sanz** went over the major points raised through the various talks and discussions.



Pascal Sanz and Olivier Piffault

They summarised the main points as follows:

- The launch of the Europeana Newspapers project in 2012 was a step forward, which will encourage other libraries to continue their efforts. Specifically, it was a turning point with regard to the quality, relevance and content of newspaper digitisation.
- The quality of the OCR, OLR and NER work done during the project is key and in some cases represents a real step forward for the project partners. At the BnF, for example, the alignment of named entities with the major repositories, thus removing ambiguities, amounts to a real new innovation for the BnF.
- The developments carried out by Europeana Newspapers will make it possible to continue improving the quality of OCR, OLR and NER.
- The project has developed a new standard in terms of the digitisation of newspapers: the Europeana Newspapers METS/ ALTO Profile (ENMAP) which will be published in January 2015.
- The project may give the impression of having focused on digitisation but the ultimate consideration for the work done so far has been the user's needs and the use that can be made of digitisation.



Bruno Racine

To conclude, **Bruno Racine** talked about the challenges represented by the digitisation of the press collections, which are a highly important source for historians. He reminded the audience that despite all that has been achieved, much work remains to be done. National libraries must comply with financial constraints, and much of the digitisation in France is focused on books.

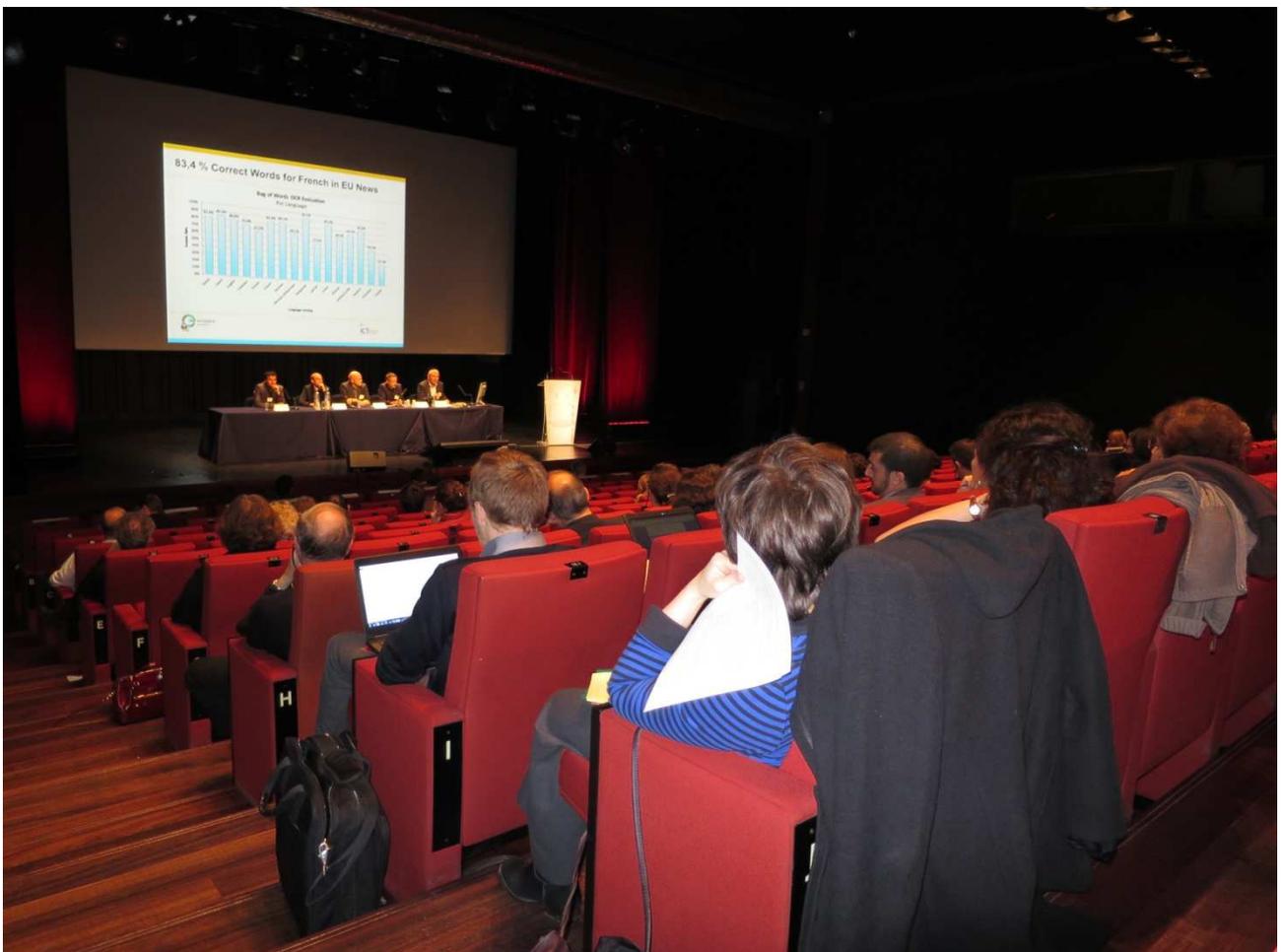
Racine also pointed out that the involvement of the BnF in the Europeana Newspapers project was a sign of its commitment towards Europeana, a commitment which can also be seen through its participation in several European projects, all of which aim to enrich the content of the service provided by Europeana.

Launched by the European Commission in 2008, Europeana remains to this day the only cultural initiative which is being undertaken throughout Europe, and an initiative which brings together not only the thousands of cultural institutions but also research centres, laboratories, the creative industries, researchers and developers.

### 3. Audience and Impact

The event was extremely successful and generated great interest in the project which will run beyond the Information Day. This is shown by the fact that of the 88 participants, 74 said they wanted to stay informed via the project newsletter.

The main target group for the Information Day were professionals and researchers. Looking at the list of participants, it is clear that the event succeeded in reaching these groups since there was a strong presence of people from both sectors in the audience.



Attendees of the French Information Day

Similarly to the Dutch Information Day, the main aims of the French one were to:

1. Inform end users interested in newspaper research of the available content from the project, and the search functionalities in both the content browser hosted by The European Library and Europeana, notably thanks to the technical developments applied to digitised content;

2. Increase overall awareness of the Europeana Newspapers project;
3. Increase awareness of the value of historical newspapers as a resource for researchers in the humanities;
4. Learn about the changes brought out by these new methods of content consultation and gather end-users expectations.

We believe that all of the above were achieved. We raised critical awareness of the project and the content, engaged users and discussed the value of digitised historic newspapers.

Throughout the day the audience engaged with the speakers and welcomed the opportunity to learn about the project. During and after the presentations, they gave their feedback as questions and comments.

They also engaged through social media. On Twitter, there were over 30 tweets using the tag [#eurnewsFR](#), including the examples below:





Finally, the surveys (see Annex IV) confirm that the event was very much appreciated. Most of the participants expressed particular interest in the technological innovations session and the panel session dedicated to the use of digitisation.

## ANNEX I: Agenda

### *French version*

- 9h00 – 09h30 Accueil et enregistrement des participants
- 9h30 – 09h40 Quelques mots d'introduction par **Sylviane Tarsot-Gillery** | directrice générale | BnF
- 09h40 – 11h00 Pourquoi Europeana Newspapers ?  
Modérateur : **Elisabeth Freyre** | BnF
- Les collections de journaux numérisés de la BnF : les réussites et les limites des premières numérisations par **Philippe Mezzasalma** | BnF
  - Présentation du projet, ses objectifs, ses résultats par **Hans-Jörg Lieder**, Bibliothèque d'Etat de Berlin
  - L'apport de la BnF dans le projet par **Ioannis Anagnostopoulos** | BnF
  - Questions / réponses
- 11h00 – 11h30 Pause
- 11h30 – 13h00 Des innovations techniques au service des utilisateurs.  
Modérateur : **Jean-Philippe Moreux** | BnF
- Comment les techniques de reconnaissance de caractères utilisées aident la recherche des utilisateurs ? par **Günter Mühlberger** | Bibliothèque universitaire d'Innsbruck
  - Comment les techniques basées sur la reconnaissance des articles peuvent apporter de nouveaux outils aux utilisateurs ? par **Claus Gravenhorst** | CCS
  - Le développement des entités nommées en français par **Clemens Neudecker** | Bibliothèque d'État de Berlin et **Alaa Abi-Haidar** | BnF
  - Questions / réponses
- 13h00 – 14h30 Déjeuner libre
- 14h30 – 16h00 Quels usages, quelles pratiques, quelles attentes pour les usagers ?  
Modérateur : **Marc Minon** | Cairn
- Table ronde avec :
- **Philippe Tétart** | Université du Maine
  - **Sophie Kurkdjian** | Université Paris I
  - **Laurent Martin** | Université Paris III Sorbonne-Nouvelle
  - **Olivier Hamon** | Syllabs
- Les intervenants témoigneront de l'évolution de leurs usages (navigation à l'intérieur du texte, nouveaux champs d'étude et mode de lecture, classification différente des contenus) quant à la presse numérisée et exprimeront leurs attentes vis-à-vis de ces nouveaux modes de consultation.
- 16h00 – 16h15 Pause

- 16h15 – 16h30 Vidéo sur le projet
- 16h30 – 17h • Conclusions et recommandations par **Pascal Sanz** et **Olivier Piffault** | BnF
- 17h00 – 17h30 • Clôture de la journée d'étude par **Bruno Racine** | président | BnF

### *Translation*

- 9h00 – 09h30 Welcome and registration
- 9h30 – 09h40 A few words of introduction by **Sylviane Tarsot-Gillery** | Director-general | BnF
- 09h40 – 11h00 Why Europeana Newspapers?  
Moderator: **Elisabeth Freyre** | BnF
- The digitised newspapers collections at BnF: success and limits of the first digitisations **Philippe Mezzasalma** | BnF
  - Presentation of the project, its objectives, results by **Hans-Jörg Lieder**, Staatsbibliothek zu Berlin
  - The contribution of BnF to the project by **Ioannis Anagnostopoulos** | BnF
  - Q & A
- 11h00 – 11h30 Break
- 11h30 – 13h00 Technical innovations for the users  
Moderator: **Jean-Philippe Moreux** | BnF
- How to use character recognition techniques to help users searches? by **Günter Mühlberger** | Innsbruck University Library
  - How do the techniques based on the recognition of items can bring new tools to the users? by **Claus Gravenhorst** | CCS
  - The development of named entities in French by **Clemens Neudecker** | Staatsbibliothek zu Berlin and **Alaa Abi-Haidar** | BnF
  - Q & A
- 13h00 – 14h30 Lunch break
- 14h30 – 16h00 What uses, what practices, what expectations for the users?  
Moderator : **Marc Minon** | Cairn  
Round table with:
- **Philippe Tétart** | Université du Maine
  - **Sophie Kurkdjian** | Université Paris I
  - **Laurent Martin** | Université Paris III Sorbonne-Nouvelle
  - **Olivier Hamon** | Syllabs

The speakers will testify on the evolution of their uses (navigation within the text, new fields of study and new reading methods, different content classification) on the digitised press and they will express their expectations towards these new consultation methods.

16h00 – 16h15 Break

16h15 – 16h30 Video on the project

16h30 – 17h • Conclusions and recommendations by **Pascal Sanz** and **Olivier Piffault** | BnF

17h00 – 17h30 • Closing of the day by **Bruno Racine** | président | BnF

## ANNEX II: List of Participants

Out of 123 Eventbrite registrations, 88 persons attended the Information Day (including the speakers). 74 agreed to be subscribed to the project newsletter.

#	Name	Surname	Organisation	Subscription to the project newsletter
1	Abi Haidar	Alaa	BnF/Lip6/CNRS	yes
2	Anagnostopoulos	Ioannis	Bibliothèque nationale de France	yes
3	Angjeli	Anila	Bibliothèque nationale de France	no
4	Ansel	Marion	Bibliothèque nationale de France	yes
5	Baryla	Christiane	Bibliothèque nationale de France	yes
6	Bellamy	Dorothée	Cairn.info	yes
7	Bergeret-Cassagne	Axelle	Bibliothèque nationale de France	yes
8	Bertrand	Claire	Bibliothèque nationale de France	yes
9	Blackburn	Charles		yes
10	Blum	Catherine	Bibliothèque nationale de France	yes
11	Bouchard	Isabelle		yes
12	Bourdon	Françoise	Bibliothèque nationale de France	yes
13	Brault	Rachel	Bibliothèque du Patrimoine de Clermont Communauté	yes
14	Camus	Alice		yes
15	Cantie	Philippe	Bibliothèque nationale de France	yes
16	Coffin	Isabelle		yes
17	Constantin	Michel		yes
18	Coste	Marie-Madeleine	Bibliothèque nationale de France	yes
19	Dechavanne	Sylvie	Archives départementales du Val-d'Oise	yes
20	Delauney	Else	Bibliothèque nationale de France	yes
21	Denou	Nathalie	Service historique de la Défense	yes
22	Desbouchages	Beatrice	Sénat	yes
23	Descharrières	Benoît		yes
24	Desnoyer	Lydie	BSG	yes
25	Doumerc	Gaëlle		yes
26	Dutertre	Julien	Archives départementales de la Gironde	yes
27	Essard-Busail	Bruno	Centre du livre et de la lecture en Poitou- Charentes	yes
28	Floc'hlay	Catherine	Bibliothèque nationale de France	yes
29	Floiras	Jacky	Education nationale	yes
30	Freyre	Elisabeth	Bibliothèque nationale de France	yes
31	Gilles	Arnaud	Bibliothèque nationale de France	yes
32	Gravenhost	Claus	CCS	no

#	Name	Surname	Organisation	Subscription to the project newsletter
33	Grisoni	Philippe	Sénat	yes
34	Hamon	Olivier	Syllabs	yes
35	Haon	Sandrine	Bibliothèque	yes
36	Hebert-Chaminade	Sophie	Cinémathèque Française	yes
37	Hondet	Elsa	Sénat	yes
38	Hopfner	Cindy		yes
39	Josse	Isabelle	Bibliothèque nationale de France	no
40	Jouve	Jean-Jacques	CEA	no
41	Kurkdjian	Sophie	IHTP	yes
42	Lallement	Nicole		yes
43	Lambert	Sylvaine	Ville de Paris Bibliothèque Historique	no
44	Lame-Bergis	Alix	Bibliothèque nationale de France	yes
45	Langlais	Pierre-Carl	CELSA	no
46	Launay	Patricia	BnF	yes
47	Lavergne	Céline	ARKHENUM	yes
48	Lecoeur	Laurence	INA	yes
49	Lemaitre	Aurelie	Irisa-université Rennes 2	yes
50	Lerouge	Julien	LITIS (Université de Rouen)	no
51	Lervese	Maud	SHD, Département Bibliothèque	yes
52	Lieder	Hans Jörg	Staatsbibliothek zu Berlin	no
53	Likforman	Laurence	Telecom ParisTech	yes
54	Malé	Séverine	Immanens	no
55	Malec	Angélique	Université Paris Sud	yes
56	Mannaz-Dénarié	Christine		no
57	Martin	Laurent	Université Paris III Sorbonne-Nouvelle	no
58	Mauléon	Agnès	i2S	yes
59	Mezzasalma	Philippe	Bibliothèque nationale de France	yes
60	Minon	Marc	Cairn.info	yes
61	Moreux	Jean-Philippe	Bibliothèque nationale de France	yes
62	Muehlberger	Guenter	Innsbruck University	yes
63	Neudecker	Clemens	Staatsbibliothek zu Berlin	yes
64	Nowicki	Frederic		yes
65	Nunes	Eric		yes
66	Ogerau	Anne-Stéphane	BnF-Partenariats	yes
67	Oury	Antoine	ActuaLitté	yes
68	Panayiotou	Hélène	Bibliothèque nationale de France	yes
69	Paquet	Thierry	Université de Rouen	yes
70	Piffault	Olivier	Bibliothèque nationale de France	yes
71	Pluvieux	Isabelle	Centre de recherche du château de Versailles	yes

72	Poirier	Annick	Bibliothèque nationale de France	yes
#	Name	Surname	Organisation	Subscription to the project newsletter
73	Pomer	Nicolas	NPC	yes
74	Rastoul	Hélène	BPI	yes
75	Renard	Joséphine	Bibliothèque nationale de France	yes
76	Rouillard	Maurice		yes
77	Sanz	Marie-Lyse		yes
78	Sanz	Pascal	Bibliothèque nationale de France	yes
79	Steffen	Florian	Bibliothèque Nationale Suisse	yes
80	Tétart	Philippe	Université du Maine	no
81	Tomasi	Anne	B.I.T.	yes
82	Tomasi	Gilbert	B.I.T.	yes
83	Tousni	Slimane	Bibliothèque nationale de France	no
84	Vaisman	Jean-Baptiste	Bibliothèque nationale de France	yes
85	Valérie	Louison-Oudot	Bibliothèque nationale de France	no
86	Valotteau	Hélène	Mediathèque Françoise Sagan, Paris	yes
87	Vayness	Shalev		yes
88	Verdese	Vincent	Archives nationales	yes

## ANNEX III: Media Coverage

Prior to the event, the programme of the Information Day was sent to the Ministry of Culture and Communication, to the directors of major French cultural institutions such as the National Archives, university libraries, those in charge of periodical collections at the Parisian libraries, the digital managers of the major publishing houses and the digital managers of the major French daily newspapers.

The event was also publicised via different media channels such as:

- Official letters of invitation
- Mailing lists
- Distribution of the agenda to events organised by the national cooperation department
- BnF's newsletters
- BnF's social media
- Europeana Newspapers website and social media
- Announcement of the event in the weekly magazine *Livres Hebdo* n°1018 issued on 14 November 2014

### COLLOQUE La presse dans Europeana

Jeudi 27 novembre, la Bibliothèque nationale de France (BNF) accueille une journée d'étude consacrée à la presse historique dans les collections de la bibliothèque numérique Europeana. Depuis 3 ans, la BNF participe à un projet de recherche soutenu par la commission européenne visant à donner accès en ligne à 18 millions de pages via Europeana et The European Library. Cette journée ouverte à tous gratuitement sur inscription, permettra de faire le point sur les collections disponibles et les améliorations techniques à apporter. La deuxième partie du programme montrera l'impact d'un tel projet sur l'historiographie et la recherche scientifique à travers les témoignages de chercheurs et utilisateurs de longue date des collections de presse numérisées. V. H.

Site : <http://bnf.fr>

Extract of *Livres Hebdo* n°1018

- Announcement of the event and a brief article on the project in the magazine *Chroniques* n°71 issued in September 2014

## Europeana Newspapers

Journée d'étude **Jeu. 27 novembre 2014**  
Le projet Europeana **9 h 30 - 17 h**  
Newspapers **Site François-Mitterrand**  
**Grand auditorium**

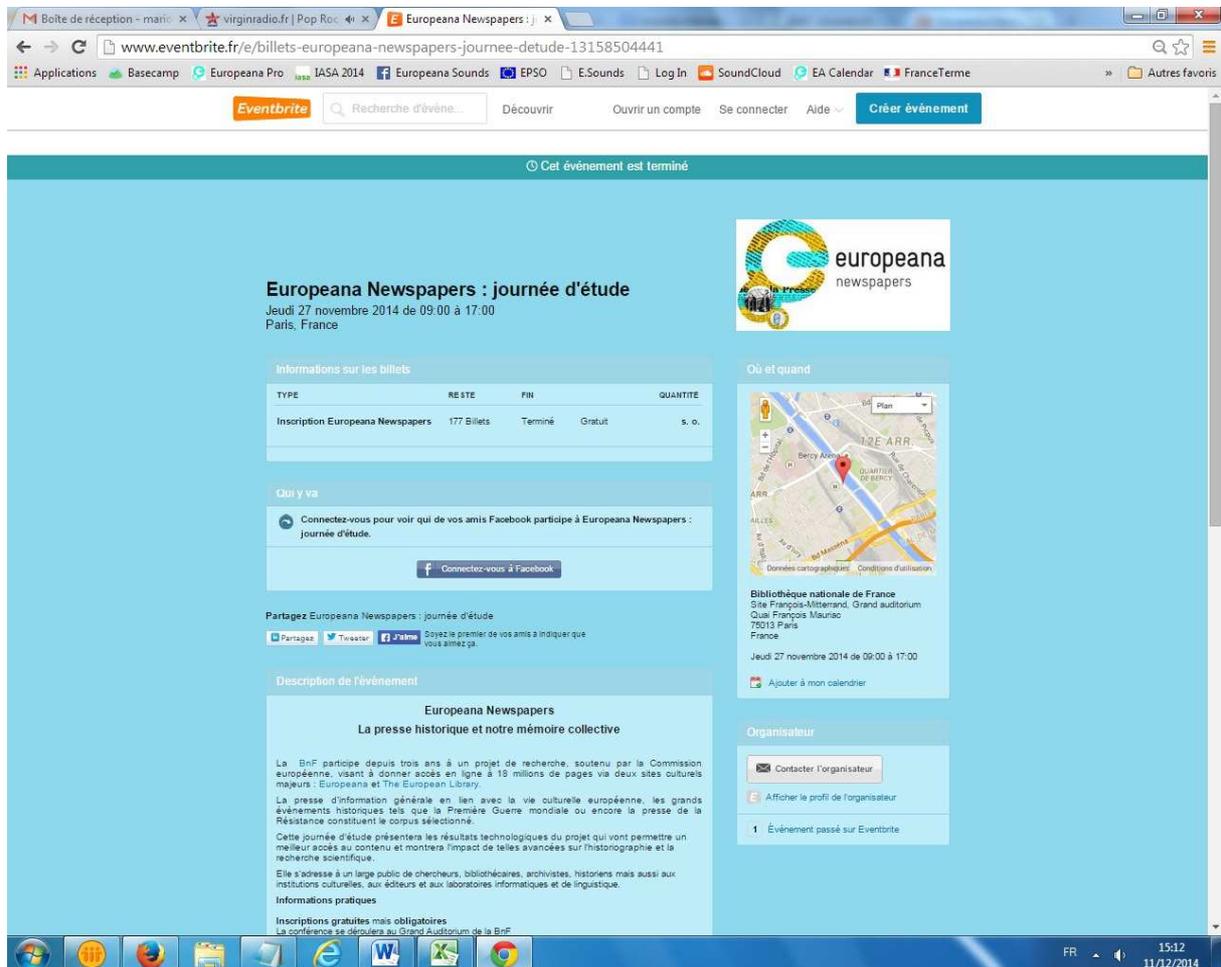
**Lancé en février 2012, Europeana Newspapers vise à rendre accessible en ligne les collections historiques de la presse quotidienne européenne : un outil précieux pour les chercheurs, mais aussi pour les passionnés et les simples curieux.**

Le projet, coordonné par la Staatsbibliothek zu Berlin, compte dix-sept partenaires, dont neuf bibliothèques nationales, trois bibliothèques universitaires, deux bibliothèques régionales et la Ligue européenne des bibliothèques de recherche (LIBER). Il vise à rendre accessibles au public, *via* Europeana, les articles des collections de presse quotidienne européenne, soit dix-huit millions de pages. Une attention particulière est portée à ceux publiés pendant la Première Guerre mondiale, en synergie avec le projet Europeana Collections 1914-1918. Europeana Newspapers, c'est aussi bien sûr l'occasion d'optimiser les pratiques en matière de numérisation des fascicules de presse (découpage à l'article, extraction des entités nommées, reconnaissance optique de caractères) et d'enrichir les modes de recherche de ces contenus numériques. Le projet arrivera à son terme début 2015 : consultable gratuitement par tout internaute, Europeana Newspapers jouera alors pleinement son rôle d'outil d'information et de démocratisation de la culture. ●

Corine Koch

Extract of *Chroniques* n°71

In addition, an Eventbrite webpage was created to promote the event and enable registrations.



Cet événement est terminé

**Europeana Newspapers : journée d'étude**  
 Jeudi 27 novembre 2014 de 09:00 à 17:00  
 Paris, France

**Informations sur les billets**

TYPE	RESTE	FIN	QUANTITE
Inscription Europeana Newspapers	177 Billets	Terminé	Gratut s. o.

**Où et quand**

Bibliothèque nationale de France  
 Site François-Mitterrand, Grand auditorium  
 Quai François Mauriac  
 75013 Paris  
 France  
 Jeudi 27 novembre 2014 de 09:00 à 17:00

**Description de l'événement**

**Europeana Newspapers**  
**La presse historique et notre mémoire collective**

La BnF participe depuis trois ans à un projet de recherche, soutenu par la Commission européenne, visant à donner accès en ligne à 10 millions de pages via deux sites culturels majeurs : Europeana et The European Library.

La presse d'information générale en lien avec la vie culturelle européenne, les grands événements historiques tels que la Première Guerre mondiale ou encore la presse de la Résistance constituent le corpus sélectionné.

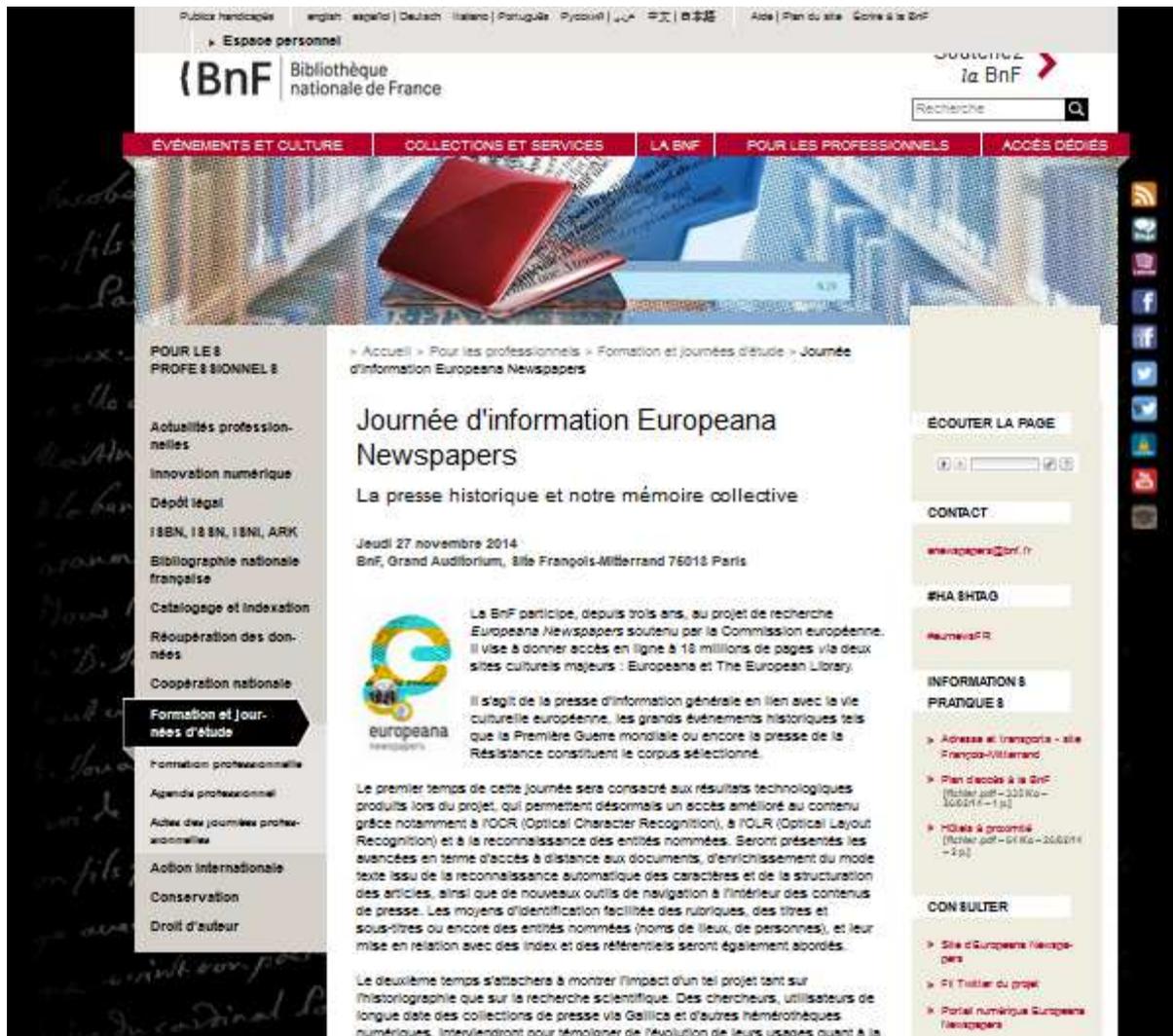
Cette journée d'étude présentera les résultats technologiques du projet qui vont permettre un meilleur accès au contenu et montrera l'impact de telles avancées sur l'historiographie et la recherche scientifique.

Elle s'adresse à un large public de chercheurs, bibliothécaires, archivistes, historiens mais aussi aux institutions culturelles, aux éditeurs et aux laboratoires informatiques et de linguistique.

**Informations pratiques**  
 Inscriptions gratuites mais obligatoires  
 La conférence se déroulera au Grand Auditorium de la BnF

<http://www.eventbrite.fr/e/billets-europeana-newspapers-journee-detude-13158504441>

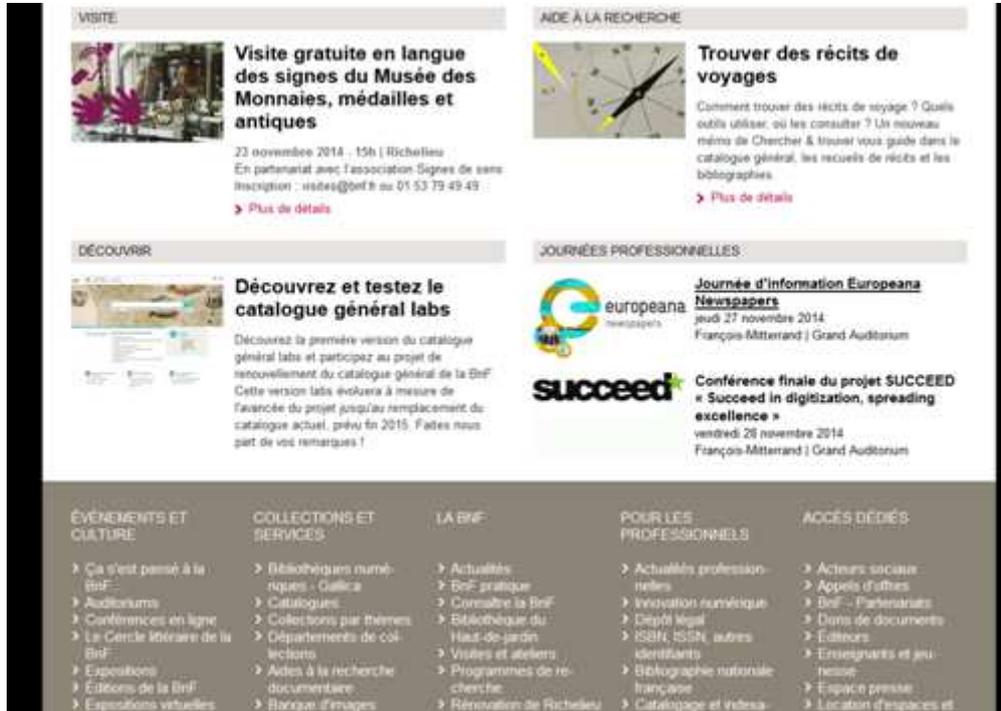
A webpage dedicated to the event was also especially created on [www.bnf.fr](http://www.bnf.fr). It also allowed promotion of the TEL browser.



The screenshot shows the BnF website interface. At the top, there are navigation links for languages (English, Spanish, Deutsch, Italien, Portugais, Pycckий, العربية, 中文, 日本語) and a search bar. Below the header, a red navigation bar contains links for 'ÉVÉNEMENTS ET CULTURE', 'COLLECTIONS ET SERVICES', 'LA BNF', 'POUR LES PROFESSIONNELS', and 'ACCÈS DÉDIÉS'. The main content area features a large image of a red book and a blue folder. The event title is 'Journée d'information Europeana Newspapers' with the subtitle 'La presse historique et notre mémoire collective'. The date is 'Jeudi 27 novembre 2014' at 'BnF, Grand Auditorium, Site François-Mitterrand 75018 Paris'. A sidebar on the left lists various professional services. The right sidebar includes social media links and contact information.

[http://www.bnf.fr/fr/professionnels/anx\\_journees\\_pro\\_2014/a.jp\\_141127\\_europ\\_newspapers.html](http://www.bnf.fr/fr/professionnels/anx_journees_pro_2014/a.jp_141127_europ_newspapers.html)

One week prior to the event, an announcement was made on BnF website's homepage.



The screenshot shows a grid of event announcements on the BnF website. The main announcement is for a free visit to the Musée des Monnaies, médailles et antiques on November 23, 2014. Other announcements include a search guide for travel narratives, a general catalog lab, and professional days for Europeana Newspapers and the SUCCEED project.

Announcement of the upcoming event on [www.bnf.fr](http://www.bnf.fr) homepage

Finally, each attendee received a folder containing:

- the agenda and biographies of the speakers
- the list of participants
- the leaflet and postcard of the project (in their French version)
- a flyer on the international activities of BnF
- the survey to fill in



Participants' folders

## ANNEX IV: Survey results

Every participant received a survey consisting of 11 questions. 40 completed surveys were returned showing positive outcomes since the respondents remarked that the programme was very interesting and the day well organised.

Question 1: How did you hear about Europeana Newspapers?

- Website : 11
- Social media : 1
- BnF : 22
- Others : 6

Question 2: The quality of presentations was:

- Unsatisfactory : 0
- Average : 1
- Good : 28
- Very good : 11

Question 3: For which presentation have you come?

- Why Europeana Newspapers? : 12
- Technological innovations on behalf of the users : 23
- Which uses, which practices, which expectations for the users? : 5

Question 4: Which talk did you find especially helpful?

- OCR, OLR et NER : 12
- Technological innovations : 7
- Users' expectations : 5
- All : 12

Question 5: Of the various topics, which one did you want to be addressed?

- Some attendees would have liked to have a live demonstration of searching with the help of the TEL browser.
- Some attendees would have liked the project to cover the press of other centuries.

Question 6: The interest of panel sessions was:

- Unsatisfactory : 0
- Average : 4
- Good : 26
- Very good : 10

Question 7: Will you be using the browser TEL in the next months?

- Yes : 36
- No, why? 4
- Occasional use : 0

Question 8: The venue (room, facilities, technical infrastructure) was:

- Unsatisfactory : 0
- Average : 0
- Good : 22
- Very good :14

Question 9: The organization of the conference (programme, conference material, coffee & lunch breaks) was:

- Unsatisfactory : 0
- Average : 2
- Good : 20
- Very good : 15

Question 10: In which field are you working?

Most of the attendees have a profession in libraries, archives, computer science, universities and research

Question 11: Please, let us know your comments and suggestions:

Some of the attendees would like to see the presentations online. The slides are available online via the projects [Slideshare](#) account and the BnF will put the videos of the talk online in 2015).