

# DELIVERABLE

**Project Acronym:** Europeana Newspapers

**Grant Agreement number:** 297380

**Project Title:** A Gateway to European Newspapers Online

---

## D6.2.1 Europeana Newspapers: Report on Workshop on Refinement and Quality Assessment

---

**Revision:** 1.0

**Authors:** Friedel Grant, LIBER – Association of European Research Libraries

**Contributions:** Marieke Willems, LIBER – Association of European Research Libraries  
Vesna Vuksan, University Library “Svetozar Marković”  
Ulrike Kölsch, Berlin State Library

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
1	11.07.2013	Friedel Grant Vesna Vuksan	LIBER	Initial draft
2	14.07.2013	Marieke Willems	LIBER	Additions and revision
3	29.07.2013	Friedel Grant	LIBER	Further revision
0.8	02.08.2013	Friedel Grant	LIBER	Draft final version
0.9	05.08.2013	Ulrike Kölsch	SBB	Revision draft final version
1.0	06.08.2013	Friedel Grant	LIBER	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This deliverable contains original unpublished work except where clearly indicated otherwise.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. Background.....	4
2. Workshop Summary.....	5
3. Methods of Stakeholder Engagement .....	6
3.1 Ask Me Almost Anything (AMAA) .....	6
3.2 Break-out sessions.....	6
3.3 Meet & Greet.....	6
3.4 Results.....	6
4. Participant Feedback .....	7
4.1 Democracy Wall.....	7
4.2 Survey.....	7
5. Suggestions for Future Workshops .....	9
6. Media Coverage.....	10
Appendix I: Published account.....	12
Appendix II: Agenda.....	14
Appendix III: Democracy Wall.....	16
Appendix IV: Survey Results.....	17
Appendix V: Signed List of Participants.....	20

## 1. Background

The Europeana Newspapers Project is aggregating 18 million digital newspaper pages for The European Library<sup>1</sup> and Europeana<sup>2</sup>. Ten million of these pages will also be refined through techniques such as Optical Character Recognition (OCR) and Optical Layout Recognition (OLR), which will allow users to fully search the text of the refined newspaper pages and to identify specific articles and images.

In addition to improving the usability of digital newspapers, the aggregation and refinement work of Europeana Newspapers allows the partners involved in the project to learn new skills and share knowledge related to newspaper digitisation. These professional competencies are then shared with the broader library and cultural heritage community through workshops and information days.

This report concerns the first of three planned workshops run by the project: a two-day event that was held in Belgrade, Serbia in June, 2013.

---

<sup>1</sup> <http://www.theeuropeanlibrary.org>

<sup>2</sup> <http://www.europeana.eu>

## 2. Workshop Summary

The *Europeana Newspapers: Refinement and Quality Assessment Workshop* took place on June 13-14, 2013 in Belgrade, Serbia. It was hosted by project partner *University Library "Svetozar Markovic"* and publicised via the Europeana Newspapers newsletter<sup>3</sup>, emails to the network, blog posts on the Europeana Newspapers<sup>4</sup> and LIBER websites<sup>5</sup> and activity on a variety of social media channels.

This resulted in a high level of interest for the workshop, with approximately 50 people attending from libraries across Europe. Many had recently joined Europeana Newspapers as an Associated<sup>6</sup> or Networking<sup>7</sup> partner. The workshop was therefore not only a chance to build skills. It was also an important chance for new and longstanding partners to network with each other.



In terms of topics covered, the workshop began with a general introduction to the project and then moved to a series of presentations on the technical methods used within Europeana Newspapers. Topics included:

- Introduction to refinement (technologies, requirements, workflows, datasets)
- Evaluation and interpretation of digitised newspapers, including the role of Ground Truth in evaluating quality
- Optimisation of digitisation workflows
- Transition from unstructured to structured newspaper data

There were also several hands-on sessions, break-out groups and demonstrations, where participants had the chance to actively discuss issues such as refinement techniques and layout evaluation.

A copy of the signed list of participants, the full workshop agenda and the full account of the workshop with links to presentations and videos are included as appendices to this document. In addition, all presentations are available on the Europeana Newspapers Slideshare account: [http://www.slideshare.net/Europeana\\_Newspapers](http://www.slideshare.net/Europeana_Newspapers).

---

<sup>3</sup> <http://us5.campaign-archive2.com/?u=d6917fe9aedb5cab842024bb9&id=e1f7433e2e>

<sup>4</sup> <http://www.europeana-newspapers.eu/inside-view-on-refinement-to-be-presented-in-belgrade>

<sup>5</sup> <http://www.libereurope.eu/blog/europeana-newspapers-workshop-on-digital-newspaper-refinement>

<sup>6</sup> <http://www.europeana-newspapers.eu/consortium/associated-partners>

<sup>7</sup> <http://www.europeana-newspapers.eu/network>

### 3. Methods of Stakeholder Engagement

The workshop included multiple events designed to encourage stakeholder engagement. These activities were planned by the LIBER office, as part of the overall organisation of the workshop structure and agenda.

#### 3.1 Ask Me Almost Anything (AMAA)

The AMAA feature is a way to ensure regular consultation and dialogue between the audience and speakers. It was organised as follows:

1. An email was sent out 2 weeks before the actual workshop. It asked participants to submit questions to be addressed during the workshop.
2. At the workshop, an AMAA box was set up. Questions could be written down on post-it notes and placed in the box. This aimed to:
  - i. Give participants a way to pose questions privately, if they did not want to do so in public.
  - ii. Offer a way to raise questions outside of the formal workshop agenda.

Each presentation finished with an AMAA session, during which participants could ask questions or organisers could highlight questions that had been submitted privately.

#### 3.2 Break-out sessions

Each participating Work Package organised a session where participants had to work with software used in the project and share best practices related to newspaper digitisation with each other.

#### 3.3 Meet & Greet

Participants were split into groups of 5 people. Project, associated and networking partners were equally distributed for effective networking. Discussions of 15 minutes were based on 4 questions:

1. One unique thing about me is...
2. The organisation I work for is...
3. Our connection to historical newspapers is...
4. My relationship to the Europeana Newspapers Project is...

#### 3.4 Results

Participants actively took part in the break-out sessions, meet & greet and most classic form of the AMAA where questions were posed at the end of each presentation. The email and AMAA box were used by only one person each. This workshop provided participants with many opportunities to ask their questions; therefore these two methods were not used.

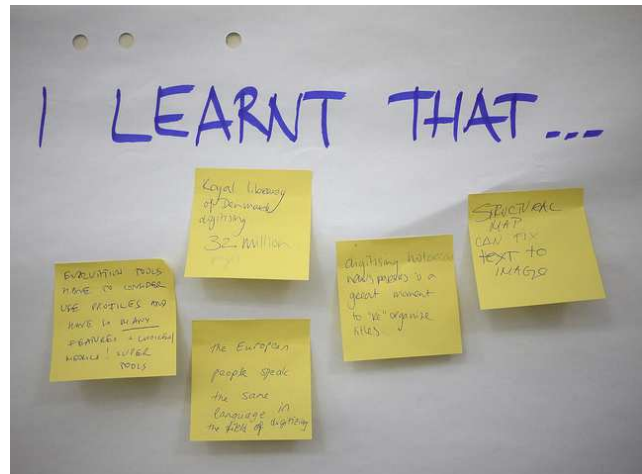
## 4. Participant Feedback

Feedback on the workshop was collected via two main methods:

### 4.1 Democracy Wall

The Democracy Wall asked people to complete one of five sentences and to write their thoughts on a post-it note. The sentences were:

- a. I learned that...
- b. I discovered that...
- c. I felt that...
- d. I noticed that...
- e. I would like to suggest...



Comments ranged from the prosaic – ‘The European people speak the same language in the field of digitising’ – to the practical – ‘Digitising historical newspapers is a great moment to “re” organise titles’.

A full list of comments is available in Appendix II.

### 4.2 Survey

Twenty-four participants answered an email survey about the workshop. The survey asked the following questions:

1. How do you rate the information you received in advance of the workshop?
2. How was the balance between presentations and interactive sessions?
3. How was the balance between the program content and networking/social activities?
4. Did the program fit your professional needs?
5. Any other comments or suggestions?

The results revealed the following:

- **63%** of survey respondents said the information they received was *Very Good*. The remainder rated the information as *Good*.
- **88%** of respondents said the balance between presentations and interactive sessions was about right. Eight percent wanted more participant activities.
- **83%** of respondents said the balance between the program content and networking was about right. Twelve percent wanted more presentations on project activities.

All survey participants said that the workshop had met their professional needs. Some elaborated on this positive evaluation with a comment. These comments included the following:

*“It was a great introduction for me to the project and an excellent opportunity to meet other partners.”*

*“It did fit quite well. It was very interesting to get more information about WP3 and the software Aletheia.”*

*“It was very interesting and the level of discussion was both inspiring and useful in practical ways.”*

*“I discovered new systems which I can test, and met new people, who solve the same problems.”*

Other comments from participants included a mix of further praise and suggestions for future workshops.

*“It was a tight program. Thus it would have been nice with a more strict control of the schedule. I liked the place and the people.”*

*“More written information about using the software Aletheia and making the use scenarios would be highly preferred.”*

*“It would be great if those associated project partners who already have a rather elaborated portal for access to their digitisations (e.g. Luxembourg, Switzerland) and thus can be considered as benchmarks would give us a short introduction to their sites as well as tell us which difficulties they encountered and how they solved them! I would find this very useful.”*

The full results of the survey are available in Appendix III.



## 5. Suggestions for Future Workshops

Based on the feedback we gathered (see section 4) and the general observations of the workshop organisers, we have a number of suggestions for future events.

- 1. The basic format can be repeated.** The Belgrade workshop was nearly fully booked and the feedback we received was overwhelmingly positive. This suggests that the core structure of the workshops is solid and can be carried through to future events.
- 2. Supporting printed materials could be helpful.** Because of the highly technical nature of the project and the volume of information, participants would appreciate some printed materials to take away with them. One example of this could be a short factsheet that conveys the core points of each presentation and gives web links to further information.
- 3. Associated Partners should have a more active role.** Some Associated Partners have extensive experience in the digitisation and display of digital newspapers. We should consider inviting them to present their work in future workshops, in addition to presentations from full project partners.
- 4. Networking between participants should continue to be a prominent feature on the agenda.** These exercises (e.g. the Meet & Greet sessions) served several purposes. They provided a break from the more formal presentations gave participants a chance to learn from each other and increased stakeholder engagement by introducing project organisers directly to attendees. In addition, feedback from our Democracy Wall and survey shows that the networking activities were very positively received.
- 5. Local media should be directly targeted for publicity.** The team from *University Library "Svetozar Markovic"* was highly successful in raising additional publicity for the project through local newspapers, radio and television stations. The local organising teams of future events should also proactively contact media outlets to offer interviews and commentary about the workshop they are hosting.

## 6. Media Coverage

The events and highlights of the workshop were shared in many ways with the project and wider cultural heritage community.

Workshop reports were published on the websites of Europeana Newspapers<sup>8</sup> and LIBER<sup>9</sup>, and photographs from the day were placed on the Flickr<sup>10</sup> and Facebook<sup>11</sup> accounts of Europeana Newspapers.

Members of Work Package 6 also used the workshop as a chance to make a video illustrating the broader goals of the project. This video was placed online<sup>12</sup> and will serve as a promotional tool for future events. It has already been shown to the audience of the Europeana Newspapers workshop at LIBER's 42<sup>nd</sup> annual conference (June 2013). In addition, some participants also blogged about the workshop, including the research department from the Dutch National Library (Koninklijke Bibliotheek)<sup>13</sup>.

The organising team in Belgrade was very active and succeeded in having the workshop covered by various local media sources.

Media organisation	Web address
Blic	<a href="http://www.blic.rs/Kultura/Vesti/387588/400000-stranica-cirilicnih-novina-iz-Srbije-prvi-put-dostupno-evropskoj-javnosti">http://www.blic.rs/Kultura/Vesti/387588/400000-stranica-cirilicnih-novina-iz-Srbije-prvi-put-dostupno-evropskoj-javnosti</a>
B92	<a href="http://www.b92.net/kultura/vesti.php?nav_category=272&amp;yyyy=2013&amp;mm=06&amp;d=13&amp;nav_id=722458">http://www.b92.net/kultura/vesti.php?nav_category=272&amp;yyyy=2013&amp;mm=06&amp;d=13&amp;nav_id=722458</a>
RTS	<a href="http://www.rts.rs/page/rts/sr/Dijaspora/story/1518/Vesti/1343176/Kolekcija+istorij+skih+%C4%87irili%C4%8Dnih+novina.html">http://www.rts.rs/page/rts/sr/Dijaspora/story/1518/Vesti/1343176/Kolekcija+istorij+skih+%C4%87irili%C4%8Dnih+novina.html</a>
24 sata	<a href="http://www.24sata.rs/vesti/beograd/vest/pred-javnoscju-400-000-stranica-cirilicnih-novina/93577.phtml">http://www.24sata.rs/vesti/beograd/vest/pred-javnoscju-400-000-stranica-cirilicnih-novina/93577.phtml</a>

<sup>8</sup> <http://www.europeana-newspapers.eu/focus-on-newspaper-refinement-quality-assessment-in-belgrade>

<sup>9</sup> <http://www.libereurope.eu/blog/the-challenges-of-refining-digitised-historical-newspapers>

<sup>10</sup> <http://www.flickr.com/photos/enewspapers/sets/72157634094399977/>

<sup>11</sup> <https://www.facebook.com/EuropeanaNewspapers>

<sup>12</sup> <http://youtu.be/SjtWmMh5hzE>

<sup>13</sup> <http://researchkb.wordpress.com/2013/06/20/europeana-newspapers-refinement-aggregation-workshop>

Media organisation	Web address
Beta	<a href="http://www.beta.rs/?tip=article&amp;kategorija=vesti&amp;zemlja&amp;ida=2899154&amp;id=&amp;ime">http://www.beta.rs/?tip=article&amp;kategorija=vesti&amp;zemlja&amp;ida=2899154&amp;id=&amp;ime</a>
Glas javnosti	<a href="http://www.glas-javnosti.rs/clanak/kultura/glas-javnosti-13-06-2013/400-000-stranica-cirilicnih-novina-pred-javnoscu">http://www.glas-javnosti.rs/clanak/kultura/glas-javnosti-13-06-2013/400-000-stranica-cirilicnih-novina-pred-javnoscu</a>
Međunarodni radio Srbija	<a href="http://glassrbije.org/kultura/%C4%8Dlanak/novine-iz-srbije-na-%C4%87irlici-u-elektronskoj-formi">http://glassrbije.org/kultura/%C4%8Dlanak/novine-iz-srbije-na-%C4%87irlici-u-elektronskoj-formi</a>
Glas Srpske	<a href="http://www.glassrpske.com/kultura/vijesti/Cirilicne-novine-za-evropsku-javnost/lat/122242.html">http://www.glassrpske.com/kultura/vijesti/Cirilicne-novine-za-evropsku-javnost/lat/122242.html</a>
Radio televizija Vojvodine	<a href="http://www.rtv.rs/sr_lat/kultura/pred-evropskim-citaocima-400.000-strana-cirilicnih-novina_400091.html?utm_source=feedburner&amp;utm_medium=feed&amp;utm_campaign=Feed%3A+RtvSveVesti+(RTV+poslednje+vesti)">http://www.rtv.rs/sr_lat/kultura/pred-evropskim-citaocima-400.000-strana-cirilicnih-novina_400091.html?utm_source=feedburner&amp;utm_medium=feed&amp;utm_campaign=Feed%3A+RtvSveVesti+(RTV+poslednje+vesti)</a>
In4S	<a href="http://www.in4s.net/index.php/magazin/kultura/39829-400000-stranica-irilicnih-novina-iz-srbije-prvi-put-dostupno-evropskoj-javnosti-">http://www.in4s.net/index.php/magazin/kultura/39829-400000-stranica-irilicnih-novina-iz-srbije-prvi-put-dostupno-evropskoj-javnosti-</a>
Kurir	<a href="http://www.kurir-info.rs/vredna-zbirka-400000-stranica-cirilicnih-novina-pred-javnoscu-clanak-838713">http://www.kurir-info.rs/vredna-zbirka-400000-stranica-cirilicnih-novina-pred-javnoscu-clanak-838713</a>
NS reporter	<a href="http://www.nsreporter.rs/vesti/pred-evropskim-citaocima-400-000-strana-cirilicnih-novina/">http://www.nsreporter.rs/vesti/pred-evropskim-citaocima-400-000-strana-cirilicnih-novina/</a>
Novine novosadske	<a href="http://novinenovosadske.rs/pred-evropskim-citaocima-400-000-strana-cirilicnih-novina/">http://novinenovosadske.rs/pred-evropskim-citaocima-400-000-strana-cirilicnih-novina/</a>
Pančevačko čitalište	<a href="http://www.citaliste.com/content/view/4797/148/lang,sr/">http://www.citaliste.com/content/view/4797/148/lang,sr/</a>
Naslovi	<a href="http://www.naslovi.net/2013-06-12/beta/pred-evropskim-citaocima-400-000-strana-cirilicnih-novina/6000989">http://www.naslovi.net/2013-06-12/beta/pred-evropskim-citaocima-400-000-strana-cirilicnih-novina/6000989</a>
Aladin	<a href="http://vesti.aladin.info/2013-06-13/71892017">http://vesti.aladin.info/2013-06-13/71892017</a>

## Appendix I: Published account

### Focus On Newspaper Refinement & Quality Assessment in Belgrade

Techniques for refining and assessing the quality of digitised historical newspapers were the main topic of discussion at a recent Europeana Newspapers workshop at the University Library “Svetozar Markovic” in Belgrade, Serbia.

#### Day One

Workshop participants were given a warm and personal welcome by Dr Alexandar Jerkov, director of the host library. After this official welcome, the workshop began with an introduction from Marieke Willems of LIBER. She gave a brief overview of the project, and explained the structure of the two-day workshop.

Presentation: [http://www.slideshare.net/Europeana\\_Newspapers/enp-ws-belgrademw10062013](http://www.slideshare.net/Europeana_Newspapers/enp-ws-belgrademw10062013)

Video: [http://www.youtube.com/watch?v=p\\_7T1ZTbKsE&feature=youtu.be](http://www.youtube.com/watch?v=p_7T1ZTbKsE&feature=youtu.be)

Hans-Jörg Lieder from the [Berlin State Library](#) (coordinator of the Europeana Newspapers Project) spoke about the challenges involved in digitising and refining historical newspapers. The relatively poor quality of the paper and ink used to print newspapers, for example, means that they often deteriorate faster than other materials. He also explained how the project would add value to historical newspapers by making it easier to search through the content and pinpoint material of interest. Lieder closed his presentation by explaining future ways in which the project might aggregate and refine more content, and extend its network.

Presentation: [http://www.slideshare.net/Europeana\\_Newspapers/enp-workshop-belgradsbb](http://www.slideshare.net/Europeana_Newspapers/enp-workshop-belgradsbb)

Clemens Neudecker from the [National Library of The Netherlands](#) explained the complex process of newspaper refinement. He noted that the Europeana Newspaper project would refine some 10 million pages from libraries across Europe. This work would be done by two project partners: the [University of Innsbruck](#) and [CCS](#).

Presentation: [http://www.slideshare.net/Europeana\\_Newspapers/enp-wp2refinementintroduction](http://www.slideshare.net/Europeana_Newspapers/enp-wp2refinementintroduction)

Video: <http://www.youtube.com/watch?v=SkHDYmziZR&feature=youtu.be>

This was followed by an explanation of how to establish the quality of newspaper refinement, from Stefan Pletschacher and Christian Clausner of the [University of Salford](#). Quality assessment is important because the poor state of historical newspapers, and the use of fonts that are difficult to read, often leads to errors in the digital version. This can be corrected by using a printed newspaper as the “ground truthed” version of its digital equivalent. By comparing the two versions, the accuracy of the OCR process could be measured and this was in turn helpful when setting requirements for the outsourcing of OCR work.

Video: <http://www.youtube.com/watch?v=XMxLtTx8y1E&feature=youtu.be>

## Meet & Greet

After discussing these technical issues, a 'Meet & Greet' session allowed participants to speak with one another in more depth. In small groups, they spoke about the organisations they represent, their experience with historical newspapers and their role in the Europeana Newspapers Project.

## Forging Local Connections

With their heads full of new insights and technical knowledge, participants enjoyed a sunny walk to the Nikola Tesla Museum, where they learned about his famous inventions such as the induction motor and the remote-controlled boat. There was also a connection to the Europeana Newspapers project because Nikola Tesla gathered newspaper clippings that are now digitised by the museum to make them fully searchable. Read the joint paper from the Nikola Tesla Museum and the University Library "Svetozar Markovic" here: <http://www.inforum.cz/pdf/2013/filipi-matutinovic-stela.pdf>

## Day Two

The second day of the workshop was dedicated to refinement. Lotte Wilms from the [National Library of The Netherlands](http://www.library.nl) kicked off the morning by explaining the task of Named Entities. She demonstrated the rules that the project had decided to use for the naming of geographical places, people and organisations. She also noted that name tagging should be done in a consistent way and common sense should be used to serve the majority of search commands.

Presentation: [http://www.slideshare.net/Europeana\\_Newspapers/what-is-a-named-entity](http://www.slideshare.net/Europeana_Newspapers/what-is-a-named-entity)

NER Hands-On: [http://www.slideshare.net/Europeana\\_Newspapers/enp-wp2nerhands-on](http://www.slideshare.net/Europeana_Newspapers/enp-wp2nerhands-on)

Video: <http://www.youtube.com/watch?v=X8dQ99MHI4g&feature=youtu.be>

Claus Gravenhorst from CCS, the only private partner in the Europeana Newspapers Project, showed how the OCR-program "DocWorks" can bring structure to previously unstructured text.

[http://www.slideshare.net/Europeana\\_Newspapers/enp-ccs-wsbelgradeolrccsfinal](http://www.slideshare.net/Europeana_Newspapers/enp-ccs-wsbelgradeolrccsfinal)

Video: <http://www.youtube.com/watch?v=JqA4Jlc0z10&feature=youtu.be>

[http://www.slideshare.net/Europeana\\_Newspapers/enp-ccs-wsbelgradedwdemomarkupfinal](http://www.slideshare.net/Europeana_Newspapers/enp-ccs-wsbelgradedwdemomarkupfinal)

Günter Mühlberger explained how the project needed to align various types of metadata, in order to make aggregation via [The European Library](http://www.europeana.eu) possible. The Europeana Newspapers Project uses METS/ALTO metadata, and Mühlberger explained its features to the participants. He closed his session with food for thought about structural metadata: How can one assure that the structure of the text is recognised correctly? Which categories are useful and detectable – headline, advertisement, "opinion"?

[http://www.slideshare.net/Europeana\\_Newspapers/enp-wp5-uibkmetadatameetingbelgrade2013](http://www.slideshare.net/Europeana_Newspapers/enp-wp5-uibkmetadatameetingbelgrade2013)

Video: <http://www.youtube.com/watch?v=VXzuOV0EWJc&feature=youtu.be>

Pictures of the workshop can be seen on flickr:

<http://www.flickr.com/photos/enewspapers/with/9024170781/>

## Appendix II: Agenda



AGENDA Workshop Refinement and Quality Assessment Belgrade 2013	
<b>PROJECT NUMBER</b>	<b>297380</b>
<b>PROJECT ACRONYM</b>	Europeana Newspapers
<b>Venue of the Meeting</b>	University Library "Svetozar Markovic" of Belgrade Bulevar kralja Aleksandra 71, Belgrade, Serbia Tel: (+381 11) 3370-509 mail: sofronijevic@unilib.bg.ac.rs ; marieke.willems@kb.nl
<b>13.06.2013</b>	
Hours	Topic
9.00 - 9.15	Welcome - UB & LIBER
9.15 - 9.45	Introduction to the Europeana Newspapers Project (Günter Mühlberger - UIBK)
9.45 - 10.15	Introduction to Refinement: Technologies, Requirements, Workflows, Datasets (Clemens Neudecker - KB)
10.15 - 10.45	Coffee break
10.45 - 12.00	Presentation (Stefan Pletschacher & Christian Clausner - USAL) <ul style="list-style-type: none"> <li>• General evaluation approaches and specific challenges for newspapers</li> <li>• The role of ground truth</li> <li>• Evaluation of individual processing steps and end-to-end systems</li> <li>• Interpretation of evaluation results and how to utilise the findings for optimising digitisation workflows</li> </ul>
12.00 - 13.00	Lunch

<b>13.00 - 14.15</b>	Demonstrations (Stefan Pletschacher & Christian Clausner-USAL): <ul style="list-style-type: none"> <li>• Aletheia - a semi-automated ground truth production tool</li> <li>• Layout evaluation (segmentation, classification, reading order)</li> <li>• OCR/text evaluation</li> </ul>
<b>14.15- 14.45</b>	Coffee break
<b>14.45 - 15.30</b>	Break-out session (Stefan Pletschacher & Christian Clausner - USAL): <ul style="list-style-type: none"> <li>• Use scenarios</li> <li>• Technical features</li> <li>• Metrics</li> <li>• Evaluation tools</li> <li>• Specification of requirements</li> <li>• Implementation of new features</li> </ul>
<b>15.30 - 16.00</b>	Reporting back & wrapping up - (Stefan Pletschacher & Christian Clausner - USAL)
<b>16.00 - 16.30</b>	Meet & Greet
<b>17.00 - 18.00</b>	Visit & presentation in Tesla Museum
<b>19:30</b>	Social evening
<b>14.06.2013</b>	
<b>Hours</b>	<b>Topic</b>
<b>9.30 - 10.00</b>	Introduction to NER (Clemens Neudecker -KB)
<b>10.00 - 10.30</b>	OLR at CCS: From unstructured to structured newspaper data and the role of content providers in the overall process (Claus Gravenhorst -CCS)
<b>10.30 - 11.00</b>	Coffee break
<b>11.00 - 12.00</b>	Hands-on session: NER tagging – tools, practical issues and best practices (Clemens Neudecker & Lotte Wilms-KB)
<b>12.00 - 12.45</b>	Hands-on session: (Claus Gravenhorst - CCS) <ul style="list-style-type: none"> <li>• OLR – DocWorks demonstration</li> <li>• Tagging and correction of layout and structural elements</li> </ul>
<b>12.45 - 13.15</b>	Conclusions and wrapping up
<b>13.15 - 14.30</b>	Lunch
<b>15.00 - 19.00</b>	Walking tour city centre or bus tour to Avala monument, details to be confirmed <a href="http://en.wikipedia.org/wiki/Monument_to_the_Unknown_Hero">http://en.wikipedia.org/wiki/Monument_to_the_Unknown_Hero</a> <a href="http://en.wikipedia.org/wiki/Avala_Tower">http://en.wikipedia.org/wiki/Avala_Tower</a>

## Appendix III: Democracy Wall

The feedback from the Democracy Wall was as follows:

*I would like to suggest...*

- That you teach us more about full text search and ranking.
- We all give applause to the organisers. Great!
- Workshop on the outside (such a beautiful day).
- We'd like to hear more about ALTO / METS metadata.
  - Me too!

*I learned that...*

- The European people speak the same language in the field of digitising.
- The Royal Library of Denmark is digitising 32 million newspaper pages.
- Digitising historical newspapers is a great moment to “re” organise titles.
- Structural map can fix text to image.
- Evaluation tools have to consider use profiles and have so many features and choices / metrics! Super tools.

*I felt that...*

- We had a very warm welcome by Dr A. ,Jerkov
- Great event, organisation, delegate packs, friendly and helpful staff at University Library SM
- The atmosphere was very warm.

*I noticed that...*

- Librarians are always good and gentle people.
- Librarians are also very good looking persons!
- The presentations were just super!
- No one fell asleep :)

*I discovered...*

- There are a lot of new things to learn!
- The food is great!
- I have to examine more the ground truthing.
- I like the word count feature in Veridian for article classification
- BnF is piloting crowdsourcing of newspaper content
- Named Entities are not straight forward



## Appendix IV: Survey Results

These are the full results of the feedback survey conducted after the workshop.

1. How do you rate the information you received in advance of the Workshop?		
Answer Options	Response Percent	Response Count
Very Good	62,5%	15
Good	37,5%	9
Fair	0,0%	0
Variable	0,0%	0
Not Good	0,0%	0
Other (please specify)		0
<i>answered question</i>		<b>24</b>
<i>skipped question</i>		<b>0</b>

2. Balance between presentations and interactive sessions:		
Answer Options	Response Percent	Response Count
Need more presentations	4,2%	1
Need more participant activities	8,3%	2
The balance was about right	87,5%	21
Other (please specify)		1
<ul style="list-style-type: none"> <li><i>The participant's activities should be planned better.</i></li> <li><i>Technical issues, rooms, briefing should be better prepared.</i></li> </ul>		
<i>answered question</i>		<b>24</b>
<i>skipped question</i>		<b>0</b>

3. Balance between the program content and networking/social activities:		
Answer Options	Response Percent	Response Count
Need more time for networking activities	4,2%	1
Need more presentations on project activities	12,5%	3
The balance was about right	83,3%	20
Other (please specify)		1
<ul style="list-style-type: none"> <li><i>Museum visit was refreshing!</i></li> </ul>		
<i>answered question</i>		<b>24</b>

#### 4. Did the program fit your professional needs?

All 24 people said the programme had met their professional needs. Some also left the following additional comments:

I had 3 needs.

1. Now: Networking.
2. Near future: Inspiration on quality control.
3. Future: How to refine our collection.

I got a better idea of my next steps in the project (Experimenting with the evaluation tools and working with DocWorks).

It did fit quite well. It was very interesting to get more information about WP3 and the software Aletheia.

It was very interesting and the level of discussion was both inspiring and useful in practical ways.

A little bit more time to discuss some presentations would be appreciated.

I consider the meeting very useful and the program fit my professional needs.

I would like more information on the technical aspects of NER and the possibility to implement it in existing portals.

I discovered new systems which I can test, and met new people, who solve the same problems...

Depending on participants it could be more intensive in the sense of participation of each person.

#### 5. Any other comments or suggestions?

The following additional comments were submitted:

It was a tight program. Thus it would have been nice with a more strict control of the schedule. I liked the place and the people.

Everything was just perfect :) it was definitely above my expectations

More written information about using the software Aletheia and making these scenarios would be highly preferred.

Short, overall, introduction to the Europeana project at the beginning of the seminar would be nice.

Thank you very much for this very inspiring workshop.

It would be great if those associated project partners who already have a rather

elaborated portal for access to their digitisations (e.g. Luxemburg, Switzerland) and thus can be considered as benchmarks would give us a short introduction to their sites as well as tell us which difficulties they encountered and how they solved them! I would find this very useful. Anyway, the workshop was a great experience - thank you again!




It was very helpful and nice event with technically oriented people, it was better than I expected.

Difficult were discussions between project partners and the not integrating the guests this much. This would have needed more moderation for better understanding.

## Appendix IV: Signed List of Participants

Note: there are some typos in the dates on the images themselves, however the lists in this document are filed under the correct date.

**June 13<sup>th</sup>, 2013**

---

**List of Participants**




Project: A Gateway to European Newspapers Online




Project reference no: 297380




Date: 14 June 2013

Location: Belgade, University Library

No.	Name of Participant	Institution	Signature
1.	Tiina HÖTTÄ	National Library of Finland ALF	Tiina HÖTTÄ
2.	Aleksandra Stojanovic	NUB SK	[Signature]
3.	Andrejuna Stojanovic	University Library 'Nikola Tesla' NIS	Andrejuna Stojanovic
4.	Majaž Kragelj	NUK, Ljubljana	[Signature]
5.	Wesley URDL	AUSTRIAN NATIONAL LIBRARY	[Signature]
6.	Örn Hafnkelsson	National and Univ. Library of Iceland	[Signature]
7.	Krista Kiisa	National Library of Estonia	[Signature]
8.	Slobodanka Komenić	National Library of Serbia	Komenić

  			
9.	PAVLA ŠVAŠTOVA'	MORAVIAN LIBRARY, CZ	Paola Švašková
10.	Jasna Jovanović	National Library of Serbia	Jasna
11.	Ivana Schreiv	Digital Archiving Solutions doo	Sigrid Ivana
12.	Igor Knežević	Digital Archiving Solutions doo	Igor
13.	DRAGAN GULBOSIĆ	DIGITALNI ARHIV INFOBIS	Dragan
14.	Liliane Regamey	Swiss National Library	Liliane
15.	Alexandre Lopes	BCU Lausanne	Alexandre
16.	Karin Pircher	Tepmann Library, Italy	Karin Pircher
17.	Frank Lund	The State and University Library	Frank
18.	Milica Köstlich	State Library of Jek	Milica
19.	Kerstin Wendt	Hamburg State and University Library	K. Wendt
20.	Alan Vaughan Hughes	National Library of Wales	Alan
21.	Joni's Törin	NATIONAL LIBRARY OF CZECH REP.	Joni
23.	Günter Hackl	UIBK	Günter
24.	Rosizza Atanasiou	British Library	RA

  			
25.	ANNA VANDALONI'	NATIONAL LIBRARY OF CZECH REPUBLIC	Anna
26.	LE TALLEC Johann	National Library of France	Le Tallec
27.	ANASTASIOU Jannis	National Library of France	Anastasio
28.	JELENA VOJVODIĆ	NATIONAL MUSEUM BELGRADE	Jelena
29.	ZORAN KRSTIĆ	NARODNA IN UNIVERZITETNA BIBLIOTEKA	Zoran
30.	RENATA PETROVIĆ	National and University Library in Zagreb	Renata
31.	SOFIAKLARIN ZADRAVEC	- II -	Sofia
32.	Andrea Langendoen	National Library of the Netherlands	Andrea
33.	HELENA PATRÍCIO	NATIONAL LIBRARY OF PORTUGAL	Helena
34.	Marieke Willems	LIBER	Marieke
35.	Clemens Nendecke	National Library of the Netherlands	Clemens
36.	Elena Sánchez Nogales	National Library of Spain	Elena
37.	CLAUS GRAVENHORST	CEIS	Claus
38.	Friedel Brent	LIBER	Friedel
39.	Christian Clausner	USAL	Christian

	 <small>Programme co-funded by the EUROPEAN UNION</small>		
25	Lotte Wilms	KB, National Lib. of NL	<i>[Signature]</i>
26	David Wrasna	UB	<i>[Signature]</i>
27	Andrija Vrsai Jelenc	UB	<i>[Signature]</i>
28	ALEXANDRA TRIVAC	UB	<i>[Signature]</i>
29	ADAM ŠAFARŽIČ	UB	<i>[Signature]</i>
30	<i>[Signature]</i>	UB	<i>[Signature]</i>
31			
32			
33			
34			
35			
36			
37			
38			
39			

June 14<sup>th</sup>, 2013



14G.



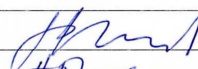
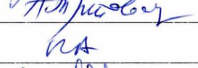
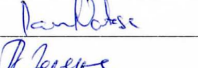
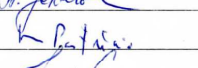
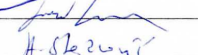



### List of Participants

Project: A Gateway to European Newspapers Online

Project reference no: 297380

Date: 13 June 2013

Location: Belgade, University Library

No.	Name of Participant	Institution	Signature
1.	STANKOVA GABRIJELA	NARODNA BIBLIOTEKA SRBIJE	
2.	ALEXANDRA TRIVIC	UB	
3.	Romya Anastasova	British Library	
4.	ANASTASOVA ROMYIA	UB	
5.	Jelena Andonovski	UB	
6.	HELENA PATRICIO	NATIONAL LIBRARY OF PORTUGAL	
7.	Francine Lind	The State and University Library	
8.	Andrija Stjepanovic	University Library Nikola Tesla	



14.6.





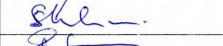
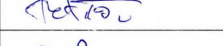


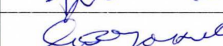

### List of Participants

Project: A Gateway to European Newspapers Online

Project reference no: 297380

Date: 14 June 2013

Location: Belgrade, University Library

No.	Name of Participant	Institution	Signature
1.	Lotte Witms	KB	
2.	Clemens Nendelke	KB	
3.	Sofija Klarin Zadrawec	NSK Zagreb	
4.	Željka Petrović	NSK Zagreb	
5.	Stefan Pletschacher	USAZ	
6.	Ulrike Kötter	State Library of Berlin	
7.	Klausener Fischer	Royal Library of Belgium	
8.	Miroslav Stojanović	NBS	





14.06.2013



### List of Participants

Project: A Gateway to European Newspapers Online

Project reference no: 297380

Date: 13 June 2013

Location: Belgrade, University Library

No.	Name of Participant	Institution	Signature
1.	Dmitry Kurbel'son	Nat. and Univ. lib. of Iceland	
2.	ANAGNOSTOPOULOS Ioannis	National Library of France	
3.	Andrea Langendoen	National Library of the Netherlands	
4.	DEJAN G. LUBSIC	DIGITALM ARHIV INFOSIRO	
5.	Christian Clausner	USAL	
6.	Wesley URDL	ONB	
7.	Karin Pircher	Tegmann Library, Italy	
8.	Tomáš Polák	Česká národní ústřední knihovna	



14.6.



9.	Yves Maurer	National Library of Luxembourg	
10.	Tina Kocica	NLF	
11.	Liliana Repanec	Swiss National Library	
12.	Ivana Sotirov	Digital Arch. SOLUTIONS	
13.	Igor Kvežević	- II -	
14.	Ularina Vasileva	National Library Bulgaria	
15.	LE TALLEC yohann	FRENCH NATIONAL LIBRARY	
16.	Verstin Wendt	Hamburg State and University Library	
17.	Krista Kiisa	National Library of Estonia	
18.	Lopes Alexandre	BCU Lausanne	
19.	DUSLO VITAS	MATEMATIČKI FAKULTET	
20.	Günter Mühlberger	B. Univ. Würzburg	
21.	Vesna Misan	UB	
23.	MARIEKE WILLEMS	LIBER	
24.	FRIEDEL GRANT	LIBER	

14.6.



9.	AANVAUGHAN HUGHES	Nation Library of Wales.	Allyson G
10.	Ekeve Sanchez Nogales	National Library of Spain	EM
11.	Cvetan Rester	Filološki fakultet	Cvetan Rester
12.	JENETA BOSSENT	HAYMOND MUSEO FERROVIA	Jeneta Bossett
13.	Jan Krami	Филозофски факултет	Jan Krami
14.	Gunter Hackel	UIBK	Gunter Hackel
15.	Nim Nilsen	UB	Nim Nilsen
16.	Aleksandar Cernic	NU BSE	Aleksandar Cernic
17.	Stavira Topalovic	MHT	Stavira Topalovic
18.			
19.			
20.			
21.			
23.			
24.			