

From: impresso project info@impresso-project.ch
Subject: [Test] Out now: New impresso release with new exploratory tools and corpus expansion
Date: 7 September 2020 at 20:58
To: info@impresso-project.ch



Media Monitoring of the Past

Newspapers Topics Inspect & Compare [FAQ](#) [Terms of Use](#) **Marten Düring**
STAFF

Media Monitoring of the Past

Mining 200 years
of historical newspapers ⓘ

How can newspapers help understand the past? How to explore them?
Some examples to get you started!

Dear colleagues,

we are happy to announce a new release of the *impresso* interface. Over the past months we have been working hard on the expansion of our corpus, new tools for its exploration, as well as a more detailed documentation and many improvements in the interface.

Before we come to the overview of selected new features, we would like to share a few announcements with you.

Upcoming *impresso* events

In the coming months, the *impresso* team will host a number of events:

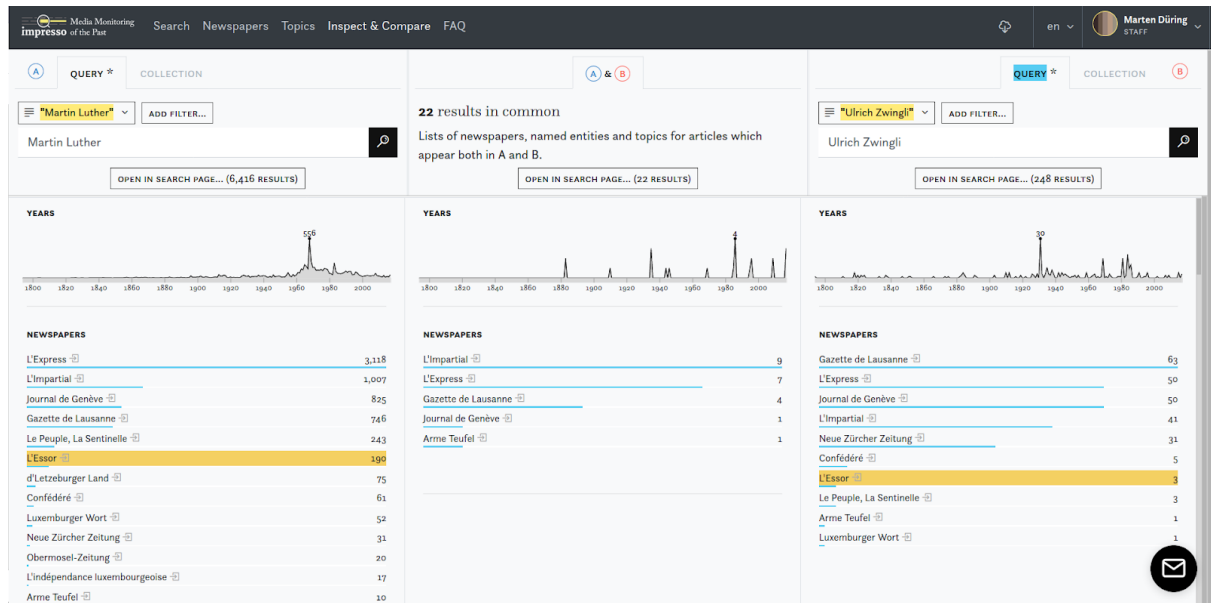
Already by the end of the week, C2DH will host the **Arcadia-Workshop** in Luxembourg to exchange about computational research workflows for historical newspapers.

Between 23rd and 24th April, we will bring together a group of historians for our [Eldorado-Workshop](#) to discuss how digitisation is changing historical research practices. From a total of 60 submissions, 24 papers were selected for presentation. The workshop is co-organised by EPFL and C2DH and will take place in Lausanne. Detailed information on registration will follow in a dedicated email.

We launched our [CLEF 2020](#) shared task [HIPE \(Identifying Historical People, Places and other Entities\)](#), a named entity processing evaluation campaign on historical newspapers in French, German and English.

Finally, on 19th May we will host a public-facing **ForumZ event** dedicated to historical newspapers and titled “**Reading Yesterday’s News in the Digital Age. Text Mining 200 Years of Historical Newspapers**”. The event will take place at the National Library of Luxembourg. More info will follow soon.

Close and distant reading with Inspect & Compare



How does the coverage of a person or event differ across newspapers? How does it evolve over time? How do user-created collections overlap and where do they differ?

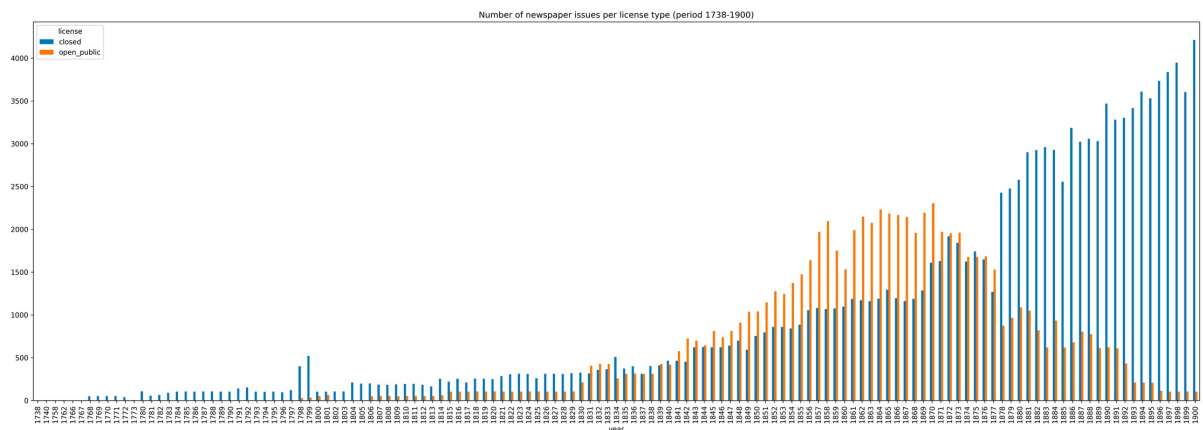
To answer these questions we have developed the Inspect & Compare component.

A high-level **overview of the composition of search queries or collections** appear on the left and right column. It includes all facets you know from search such as e.g. newspapers, persons, locations, topics and country of publication. The middle column indicates their **overlap**.

Do try it out and let us know how you have used it! Click below to compare search results for Martin Luther and Ulrich Zwingli in our corpus.

[Try Inspect & Compare](#)

Corpus expansion and statistics



We are very happy to announce a significant expansion of the impresso newspaper corpus with the addition of the following Swiss titles:

Bündner Landbote ([BLB](#)), Bündner Nachrichten ([BNN](#)), Der Erzähler ([EZR](#)), Der Liberale Alpenbote ([LAB](#)), Der Morgenstern ([MGS](#)), Der Volksfreund ([DVF](#)), Der Wahrheitsfreund ([WHD](#)), Der Zugerbieter ([ZBT](#)), Der helvetische Volksfreund ([HRV](#)), Die Gewerkschaft ([OIZ](#)), Die Tat ([DTT](#)), Freiburger Nachrichten ([FZG](#)), Gazette du Simplon ([GAZ](#)), Gazette du Valais / Nouvelle gazette du Valais ([GAV](#)), Journal de Saxon ([SAX](#)), La Contrée ([CON](#)), La FCTA ([FCT](#)), La Liberté ([LLE](#)), La Voix du Rhône ([VDR](#)), La lutte syndicale ([LLS](#)), Neue Zuger Zeitung (II) ([NZG](#)), Neue Zürcher Zeitung ([NZZ](#)), Neues Tagblatt aus der östlichen Schweiz ([NTS](#)), SMUV-Zeitung ([SMZ](#)), Schweizerische Tag-Blätter ([SRT](#)), Solidarité ([SDI](#)), St.Galler Zeitung ([SGZ](#)), VHTL-Zeitung ([VHT](#)), Wochenblatt für die vier löblichen Kantone Ury, Schwyz, Unterwalden und Zug ([DFS](#))

With this expansion we now give access to a **total of 76 newspapers (50 Swiss and 26 Luxembourgish)** which correspond to 600,919 issues, 5,429,656 pages and 3,462,799 images.

What's next?

In the near future, we will ingest two titles (Schweizer Arbeitgeber and the Schweizerische Handels-Zeitung) from our partner the [Swiss Economic Archive](#), as well as the Bundesblatt ([Federal Gazette](#)>). On the long term, we aim at expanding the impresso corpus beyond the Swiss and Luxembourgish borders. To this end, we are actively collaborating with further national libraries

and partners and, depending on time resources, more material in English and French will join the impresso corpus.

Blog post "What's in our Corpus?"

FAQ's and i-buttons

The screenshot displays the 'impresso' interface. On the left, there are three filter sections: 'FILTER BY PERSON', 'FILTER BY LOCATION', and 'FILTER BY TOPIC'. Each section has a sub-header and a list of items with checkboxes. A black tooltip box is overlaid on the 'FILTER BY TOPIC' section, containing the text: 'How to read the topics? The topics created by the topic modeling applied on the whole impresso collection help gather insights into the distribution of themes across the corpus. MORE INFO →'. To the right, there is a search bar labeled 'ALL PEOPLE MENTIONED' with a search icon. Below the search bar, it says 'Search one of 269 available choices'. Underneath, it shows '269 results' followed by a list of names with their respective counts and a small icon: Joe Strummer (32), Jim Jarmusch (8), Tom Waits (5), James Brown (4), Screamin' Jay Hawkins (4), Jane Birkin (3), Mick Jones (The Clash) (3), Nicoletta Braschi (3), Paul Simonon (3), and Ray Gange (3). At the bottom of the results list, there is a pagination bar with the number '1' highlighted.

The impresso interface gives access to many powerful tools for the exploration of historical newspapers. It is important to us to **document** the **methods** we use and the **decisions** we make but also to provide insights into the **techniques** we develop so as to empower you to critically assess the results you obtain and the patterns you discover.

Over the past months we have greatly expanded our set of Frequently Asked Questions and added info buttons with brief explanations of main functionalities to the interface.

[Take me to the FAQ's](#)

Marginalia

- armée · officier · colonel · soldat · service (3)
- journal · annonce · abonnement · bureau · avis (3)
- parti · gouvernement · socialiste · politique · pouvoir (3)
- canton · conseil · loi · droit · gouvernement (2)
- août · juillet · septembre · octobre · voyage (2)
- fond · chaux · fil · fille · marie (2)
- coup · main · tête · cheval · moment (2)
- concert · musique · information · disque · danse (2)
- action · administration · société · conseil · actionnaire (2)
- film · cinéma · semaine · réalisateur · séance (2)

L'IMPARTIAL
JOURNAL QUOTIDIEN et FEUILLE D'ANNONCES
Paraissant à La Chaux-de-Fonds tous les jours, excepté le dimanche

Et fiscalité folle
«C'est étonnant que la loi sur le vote ne soit pas en vigueur...»

La session du Conseil de la S. & N.
M. Land a été élu président de la S. & N.

Les serments communaux
M. Land a été élu président de la S. & N.

Qui gagnera?
M. Land a été élu président de la S. & N.

Les canons de l'Église.
M. Land a été élu président de la S. & N.

- Huldrych Zwingli (1)
- Max Baer (boxer) (1)
- Romano Mussolini (1)

- LOCATIONS**
- La Chaux-de-Fonds (2)
 - Lake Neuchâtel (2)
 - Suisse, Moselle (2)
 - Alexander Polyhistor (1)
 - Biel/Bienne (1)
 - Cantons of Switzerland (1)
 - Gare de Cornavin (1)
 - Italy (1)
 - Japan (1)
 - Jura Mountains (1)

A glance at a newspaper page already conveys a lot of information, for example the number and type of contents and the top headlines. To further enhance the quick browsing of newspaper pages we have introduced marginalia for each page which include the **most prominent topics, persons and locations** detected on a page.

Use your credentials to login and explore the new release. If you would like to gain access to the impresso interface, email us at info@impresso-project.ch

We are curious to hear from you

We are keen to explore with you how impresso's advanced search and discovery features work and to see how they relate to your own project. Do send us an email to info@impresso-project.ch and we will be happy to follow up!





Copyright © 2020 impresso project, All rights reserved.
You are receiving this email because you opted in via our website.

Our mailing address is:

impresso project
EPFL CDH DHI DHLAB
INN 140 (Bâtiment INN) / Station 14
Lausanne CH-1015
Switzerland

Add us to your address book

Want to change how you receive these emails?
You can update your preferences or unsubscribe from this list.

