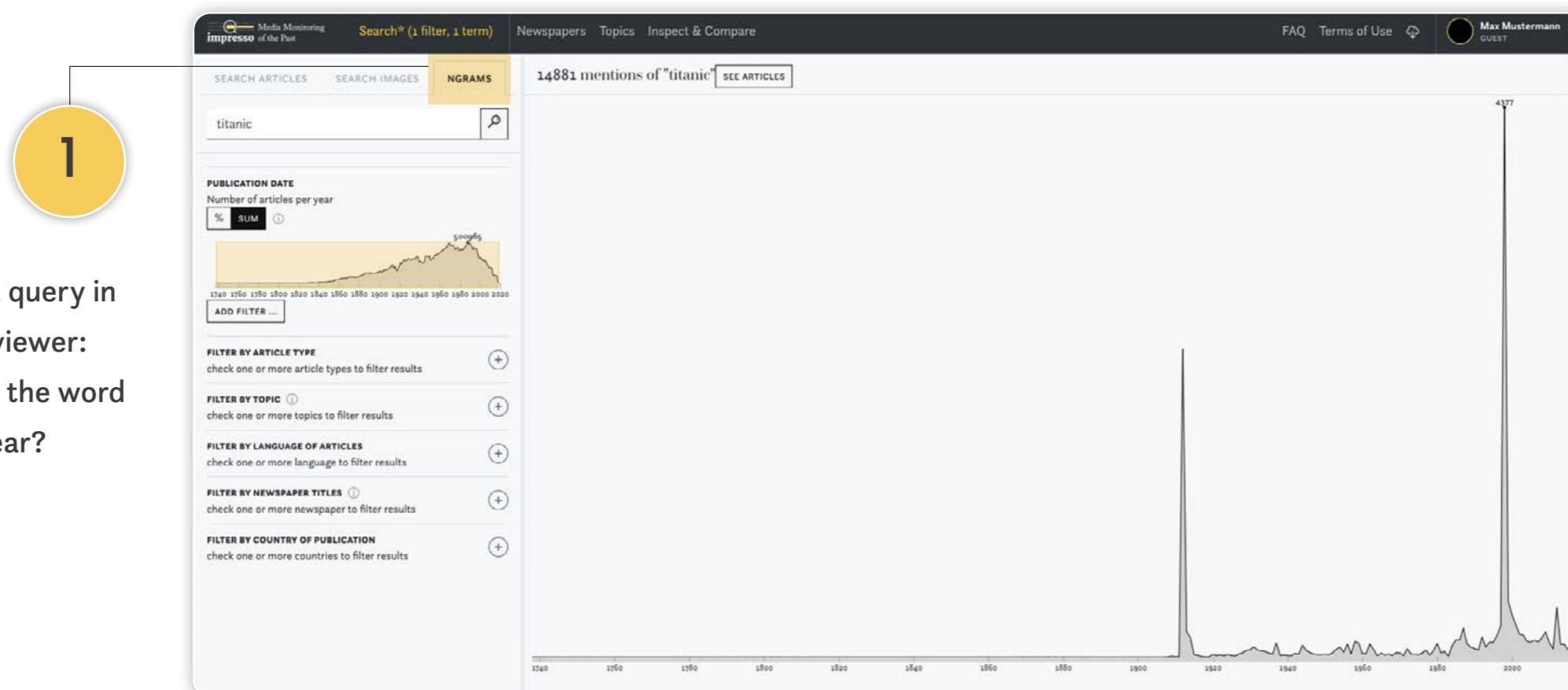


01 | TELL AN EVENT APPART FROM ITS AVATARS: THE TITANIC CATASTROPHE AND ITS MOVIES

In this challenge, we invite you to get acquainted with the features of the impresso interface: how to read the visualisations, how to search, how to use the topic modelling and the Named Entities to narrow down or expand your research. We will also guide you to use the intermediary analysis tools such as the inspect and compare page and the creation of collection. All the answers to the questions can be found in the FAQ page or directly on the interface. Are you set? Let's start!

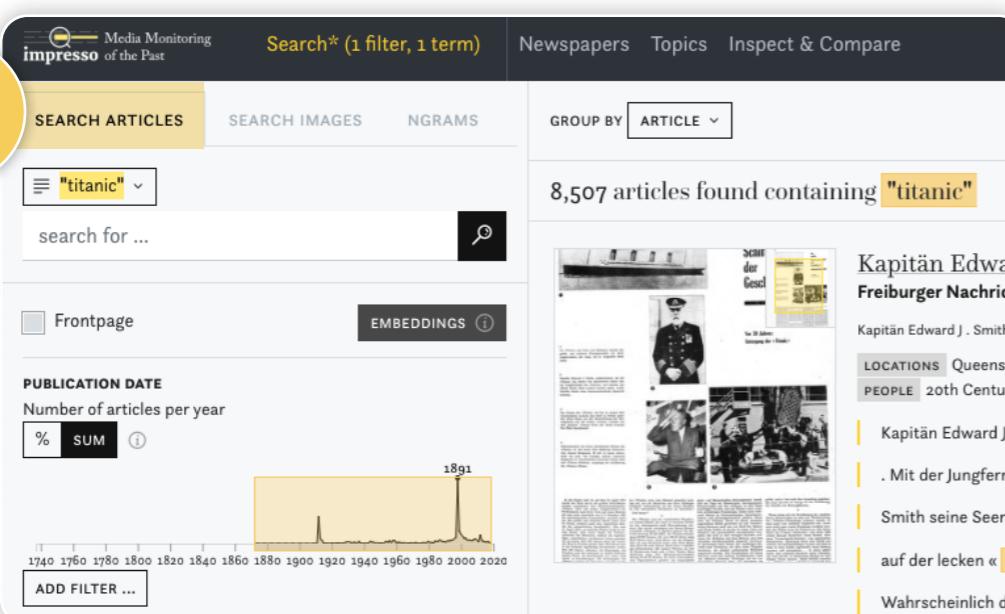
GOAL 1 collect articles on the event and on the movie in separate collections

GOAL 2 find movies dedicated to the event other than the one from 1998



in contrast, search for the word titanic in the general search page: what is the difference between the number of hits shown in the ngram viewer and the search page?

2



describe what you see:
when are the highest hits for the word
“Titanic” and how can you explain it?

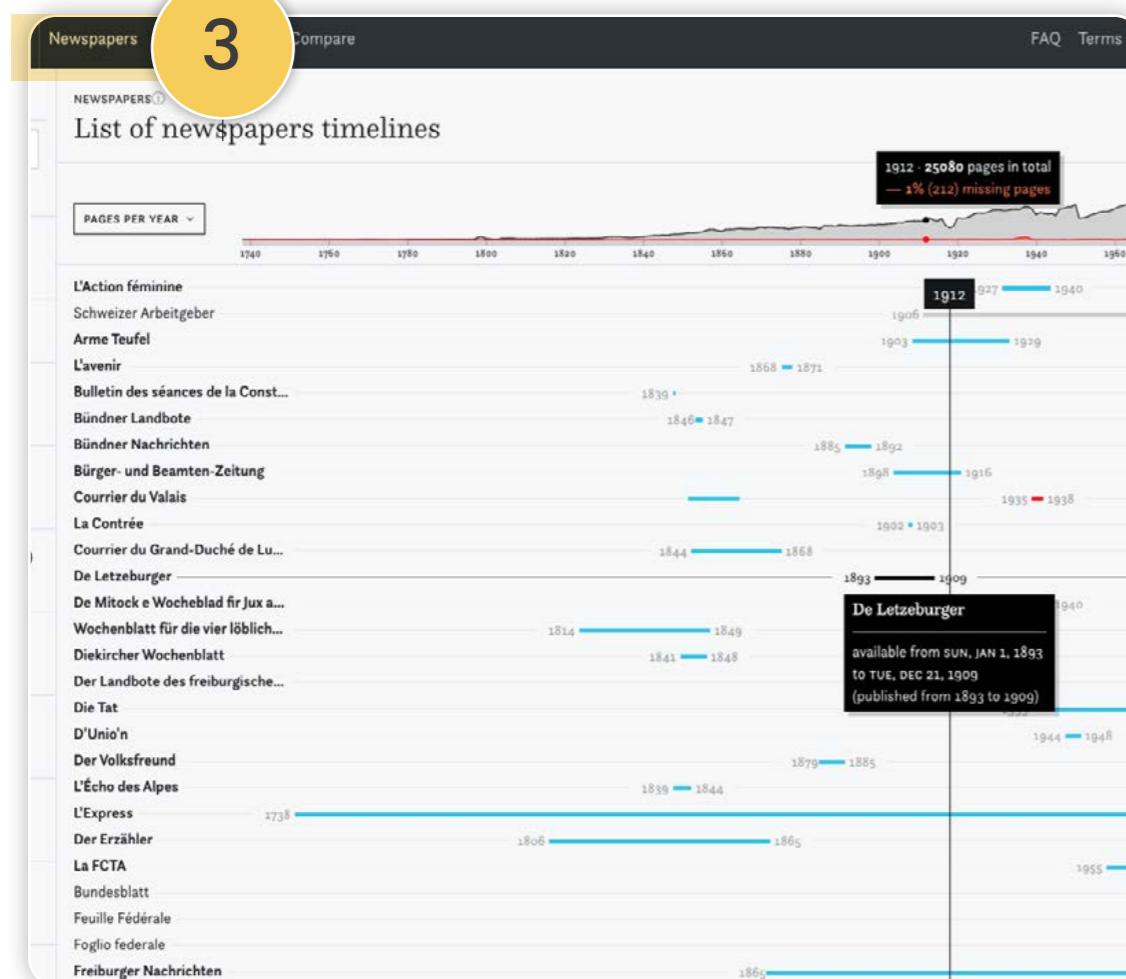
change the frequency line from SUM to %
what changes?

3



go to the newspaper page: how many newspapers pages are available in the impresso collection for the years 1912 and 1998? what is the potential impact on the previous result list?

3



4

we will start now with searching for articles about the movie **Titanic**, with the help of **NAMED ENTITIES**:
Filter by **PERSONS** related to the 1998 movie

save the result list as a collection titled “Titanic 1998”

Saving items in your collection
FEB 28, 2020, 10:56 (#4839)

CAMERON OR LEONARDO DICAPRIO
containing titanic mentioning James Cameron or Leonardo DiCaprio or Kate Winslet

PROGRESS — 13.33 %

QUARY 18, 1998 p.30

export search results as csv
FEB 28, 2020, 10:28 (#4838)

REMOVED

saved in Robert Schumann- radio; mentioning Robert Schumann

DONE — 100 %

Saving items in your collection
FEB 25, 2020, 16:44 (#4834)

REMOVED

Robert Schumann- radio with topic "soir · matin · midi · nuit · journal...", mentioning Robert Schumann

DONE — 100 %

Media Monitoring
impresso of the Past

Search* (3 filters, 3 terms)

SEARCH ARTICLES SEARCH IMAGES NGRAMS

"titanic" film · cinéma · semaine · jean · joh

Titanic-movie 1998

search for ...

Frontpage EMBEDDINGS

PUBLICATION DATE

Number of articles per year

% SUM

1740 1760 1780 1800 1820 1840 1860 1880 1900 1920 1940 1960 1980 2000 2020

ADD FILTER ...

FILTER BY COLLECTION

results are filtered when:

NOT SAVED IN

Titanic-movie 1998
@max.mustermann
LAST MODIFIED FRI, FEB 28, 2020

MORE...

FILTER BY ARTICLE TYPE

check one or more article types to filter results

FILTER BY PERSON

check one or more persons to filter results

FILTER BY LOCATION

check one or more locations to filter results

FILTER BY TOPIC

results are filtered when:

To complement the current list of results, look for topics which relate to movies. EXCLUDE the collection you created in the previous step to avoid duplicates

save that result list in a new collection

open the **INSPECT & COMPARE** page. In column A on the left side, click to add the current search query. Find column B on the right and select the “Titanic 1998” collection.

The screenshot shows the Impresso platform interface. At the top, there's a navigation bar with links for 'Media Monitoring', 'Search* (3 filters, 3 terms)', 'Newspapers', 'Topics', 'Inspect & Compare' (which is highlighted), 'FAQ', 'Terms of Use', and a user profile for 'Max Mustermann GUEST'. Below the navigation is a search bar with a dropdown set to 'titanic', a search button, and a link to 'OPEN IN SEARCH PAGE... (1,121 RESULTS)'. To the right of the search bar is a section titled 'no results in common' with a link to 'OPEN IN SEARCH PAGE... (EMPTY, NO RESULTS)'. On the far right, there's a collection card for 'TITANIC-MOVIE 1998' containing 1,413 results, with a link to 'OPEN IN SEARCH PAGE... (1,413 RESULTS)'. A large yellow circle with the number '7' is overlaid on the top center of the interface.

Scroll down and take note of the similarities and differences in columns A and B.

The screenshot shows a person search results page for 'Barbara Stanwyck'. On the left, there's a list of names under 'PERSONS', with 'Barbara Stanwyck' highlighted. On the right, a modal window titled 'CURRENT SELECTION' displays information about Barbara Stanwyck, including her birth year (1907), death year (1990), and profession (HUMAN Brooklyn, 1907 - Santa Monica). It also shows a timeline from 1953 to 1966 with a peak around 1955. A checkbox for 'apply current search filters (3 filters)' is checked, and a note says '32 results from 1953 to 1966, within current search: containing "titanic" not saved in ; with (MORE ...)'. Buttons for 'ADD AS SEARCH FILTER' and 'EXCLUDE FROM CURRENT SEARCH' are present. Below the modal is a small image of Barbara Stanwyck and the text 'Barbara Stanwyck (Q94002) HUMAN Brooklyn, 1907 - Santa Monica, 1990 [W] American actress SOURCE: WIKIDATA W/Q94002 MORE...'.

can you recognise information in those filter that will help you identify movies that came out before 1998 on the Titanic?

inspect the person list: what does it tell you?

RESULTS IN COMMON: 1,644 results in common

YIELDS:

- Column A:** "titanic" (2,852 results)
- Column B:** "film" (2,820 results)
- Column A & B:** "titanic" AND "film" (1,644 results)

compare the following queries:

COLUMN A: search for the keyword “titanic” and filter with the **TOPIC** “film, cinéma...”

COLUMN B: search for the keywords “titanic” then “film”

TOPICS	Count	TOPICS	Count	TOPICS	Count
FR film - cinéma - semaine - jean - john	2,852	FR film - cinéma - semaine - jean - john	1,644	FR film - cinéma - semaine - jean - john	1,644
FR journal - monde - vie - direct - série	665	FR journal - monde - vie - direct - série	384	FR journal - monde - vie - direct - série	1,095
FR bateau - mer - lac - bord - port	650	FR bateau - mer - lac - bord - port	366	FR bateau - mer - lac - bord - port	443
FR dimanche - samedi - lundi - vendredi - soir	585	FR dimanche - samedi - lundi - vendredi - soir	345	FR dimanche - samedi - lundi - vendredi - soir	423
FR scène - théâtre - spectacle - pièce - public	489	FR scène - théâtre - spectacle - pièce - public	325	FR scène - théâtre - spectacle - pièce - public	391
FR enfant - famille - père - vie - mère	465	FR enfant - famille - père - vie - mère	277	FR enfant - famille - père - vie - mère	381

look at the frequencies and the topics: what do you notice? where do the queries A and B overlap? what is the difference?

PERSONS	Count	PERSONS	Count	PERSONS	Count
Barbara Stanwyck	45	Barbara Stanwyck	44	Beverly Hills%2c California	69
Michèle Morgan	45	Michèle Morgan	44	Santa Barbara%2c California	69
Kate Winslet	37	Celine Dion (album)	29	Hercule Poirot	67
Nicolas Hulot	37	Charlie Chaplin	91.304%	Barbara Stanwyck	63
Charlie Chaplin	34	Sacha Guitry	29	Lucky Luke	61
Jean Dujardin	34	Alain Chabat	28	James Bond	53

scroll down to the persons and exclude on both sides the main actor of the 1998 movie and the director: what do you notice? what names do appear in column A and B?

use the metadata filters and curate the left column to focus on movies produced before 1998 on the Titanic catastrophe

8

PERSONS

Barbara Stanwyck	32
Jean-Claude Pascal	24
Michèle Morgan	21
Bernard Blier	20
Clifton Webb	20
Charlie Chaplin	18
Françoise Arnoul	18
Jean Servais	18
Sacha Guitry	18
Daniel Gélin	16

LOCATIONS

France	168
--------	-----

PERSON

Keanu Reeves PERSON

CURRENT SELECTION

apply current search filters (0 filters)

4,168 results in total (from 1965 to 2017)

LOCATION

Keanu Pelvis (Q43416)
HUMAN Beirut, 1964 -

Canadian actor, producer and musician
SOURCE: WIKIDATA W/Q43416

MORE...

PERSONS

James Cameron	1,195
Leonardo DiCaprio	900
Kate Winslet	829
Woody Allen	368
Taylor Hackford	329
Don Bluth	322
Demi Moore	299
Billy Zane	289
Sabine Azéma	284
Keanu Reeves	274

LOCATIONS

La Chaux-de-Fonds	274
-------------------	-----

open the curated list in the
SEARCH PAGE

9

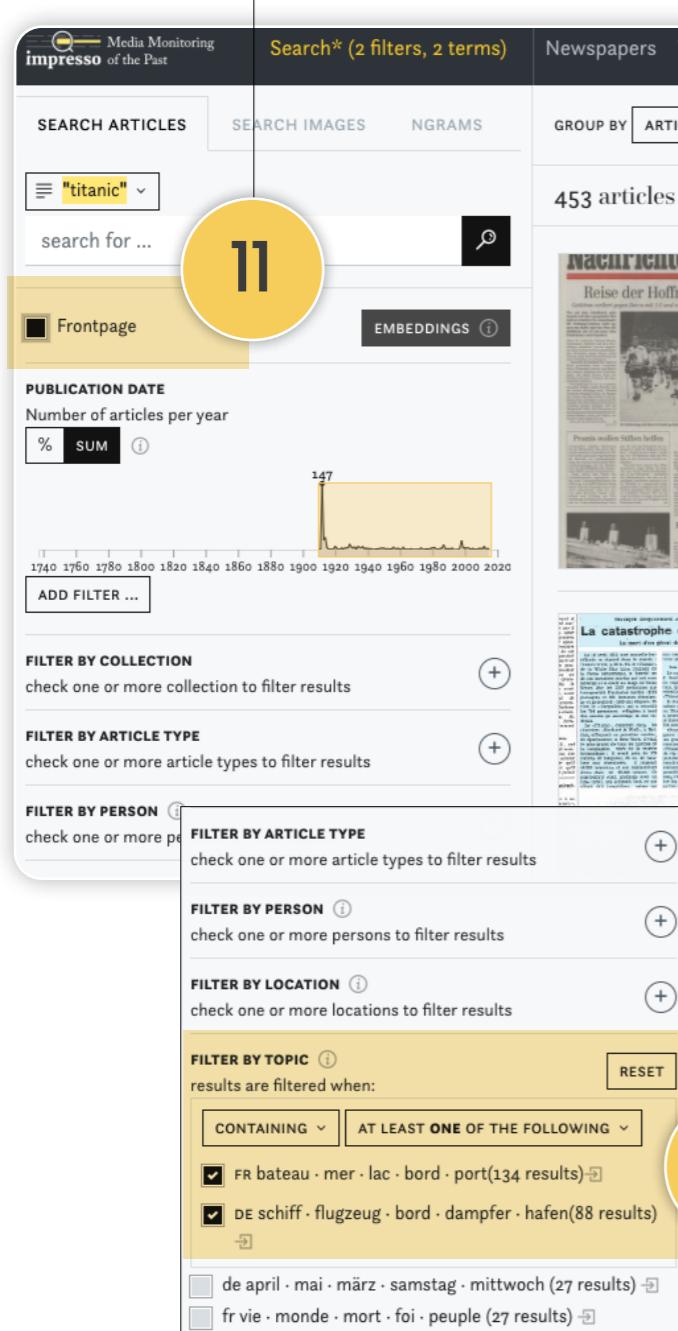
The screenshot shows the impresso search interface. At the top, there's a search bar with the placeholder "Search* (5 filters, 8 terms)". Below it, there are three tabs: "SEARCH ARTICLES", "SEARCH IMAGES", and "NGRAMS". Under "SEARCH ARTICLES", several filters are applied: "titanic" (selected), "film · cinéma · semaine · jean · john", "Titanic-movie 1998", "Celine Dion (album) OR James Camer", and "From Jan 1, 1912 to Oct 16, 1992". There's also a search input field "search for ..." and a "SEARCH" button. On the left, there's a "Frontpage" link and a "PUBLICATION DATE" section with a chart showing article counts per year from 1740 to 2020. Below that is a "PUBLISHED BETWEEN" section with date inputs for "1912-01-01" and "1992-10-15", and a "FILTER BY COLLECTION" section where "Titanic-movie 1998" is checked.

The screenshot shows the search results page. At the top, it says "Newspapers Topics Inspect & Compare". On the right, there are sections for "Saving items in your collection" (with a progress bar at 83.33%), "Saving items in your collection" (with a progress bar at 100%), and "export search results as csv" (with a "DOWNLOAD CSV" button). The main area displays several newspaper articles. One article from "Gazette de Lausanne" dated September 8, 1958, is shown with the headline ". CINÉMAS". Another article from "Journal de Genève" dated July 26, 1980, is also visible. Both articles have "VIEW" and "ADD TO COLLECTION ..." buttons.

10

create a new COLLECTION titled “titanic movies pre 1998”
based on this result list, to collect articles about other
movies than the 1998 Titanic movie

now we will look for articles dealing with the catastrophe itself:
search for the word **titanic** and select articles published in the FRONT PAGE



11

The screenshot shows the Impresso platform interface. On the left, there's a sidebar with filters for 'SEARCH ARTICLES', 'SEARCH IMAGES', 'NGRAMS', and a 'Frontpage' button (which is highlighted). Below it are sections for 'PUBLICATION DATE' (a chart showing a single peak at 1912), 'FILTER BY COLLECTION', 'FILTER BY ARTICLE TYPE', 'FILTER BY PERSON', and 'FILTER BY LOCATION'. A large yellow circle labeled '11' is overlaid on the sidebar area. The main content area shows search results for 'titanic' on the front page. It includes a summary: '453 articles found containing "titanic" appearing on the front page'. Below this are two news snippets from 'Freiburger Nachrichten' (Wednesday, March 25, 1998) and 'L'Impartial' (Saturday, April 14, 1962).

Newspapers Topics Inspect & Compare

FAQ Terms of Use Max Mustermann GUEST

GROUP BY ARTICLE ORDER BY RELEVANCE DISPLAY AS LIST TILES

453 articles found containing "titanic" appearing on the front page

Elf Oscars entern die «Titanic»
Freiburger Nachrichten WEDNESDAY, MARCH 25, 1998 p.1

Elf Oscars entern die «Titanic» James Camerons Mega-Melodrama «Titanic», der mit rund 200 Millionen Dollai Produktionskosten teuerste und kommerz

PEOPLE James R. Cameron, Helen Hunt

VIEW ADD TO COLLECTION ...

La catastrophe du Titanic
L'Impartial SATURDAY, APRIL 14, 1962 pp.1,2 (2 pages)

La catastrophe du Titanic TRAGIQUE CINQUANTIEME ANNIVERSAIRE La mort d'un géant des mers Le 16 avril 1912. une nouvelle terrifiante se répand dans le

LOCATIONS Sol Star, Belfast, Province of New York, Queenstown%2c New Zealand

PEOPLE Edward Smith, Thomas Andrews (shipbuilder)

La catastrophe du Titanic TRAGIQUE CINQUANTIEME ANNIVERSAIRE La mort d'un géant des mers Le 16

avril 1912. une nouvelle terrifiante se répand dans le monde : l'avant-veille, à 23 h. 40, le « Titanic »

12

look for the relevant topics in DE and FR

open the WORD EMBEDDING feature to look for ocr mistakes frequently associated to the word “titanic” and other words connected to it, in FR and DE.

What do you notice? Why do the words “naufrage” and “hollywood” appear?

The screenshot shows the impresso platform interface with two floating windows displaying word embeddings.

Top Window (French Embedding):

- Search bar: "titanic" (highlighted), "French", 50, GO!
- Results: titanic, terror, exodus, naufrage, naufragés, drake, leviathan, harbor, cameron, abyss, empress, atlantis, hollywood, britannic, whale, persia, fantastique, engloutie, mystery, léviathan, glorious, thirteen, twelve, naufragée, amistad.

Bottom Window (German Embedding):

- Search bar: "titanic" (highlighted), "German", 50, GO!
- Results: titanic, schifskatastrophe, queen, carpathia, nantucket, feuerschiff, rettungsboot, eisberg, lusitania, empreß, rettungsbootes, riesenschiff, schiffbrüchigen, strandung, glorious, arabic, wrack, rettungsboote, dampfers, gestrandet.

Yellow Circles:

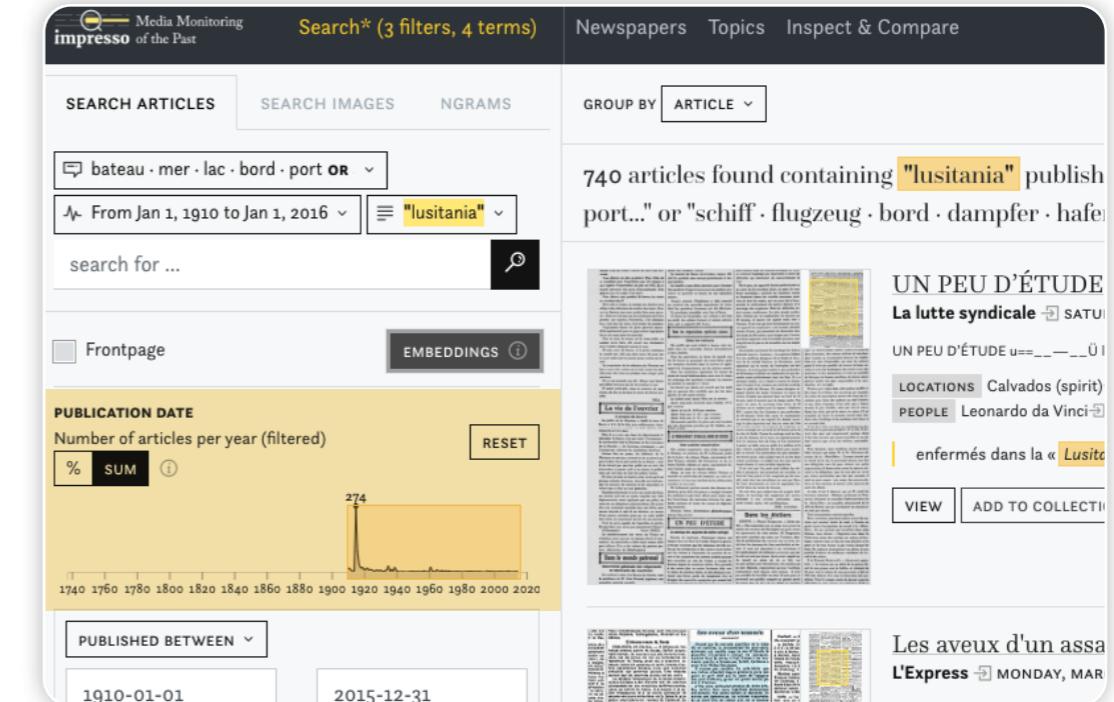
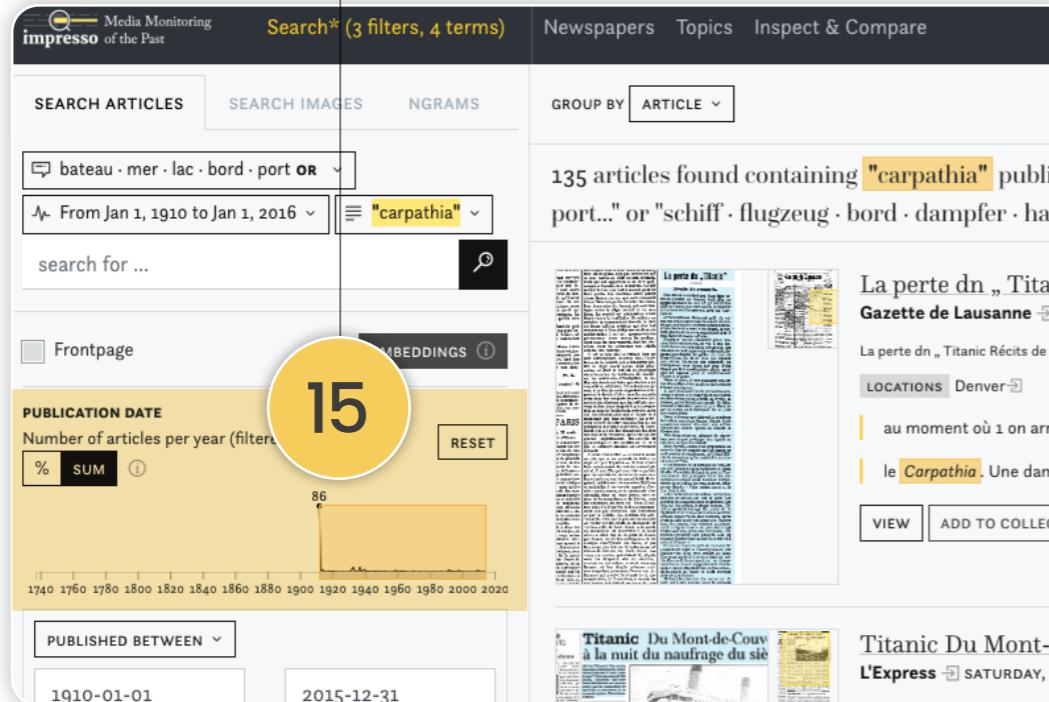
- 13**: Located on the French word embedding window.
- 14**: Located on the German word embedding window.

what are the difference between the french and the german words surrounding Titanic?

in the word embedding suggestions in German, select **carpathia** or **lusitania** with the same topics as filter

as an overall conclusion: if you go back to the first expectations formulated based on the frequency line and unigram viewer of the titanic query:

what has changed? what did you learn about what that result list means?



CHALLENGE COMPLETED!

You have learned:

- ✓ how to read the ngram and frequency lines
- ✓ what the impresso corpus contains
- ✓ how to query the corpus with key-words, persons (named entities)
- ✓ how to create and curate collections
- ✓ how to inspect and compare queries and collections
- ✓ how to curate a result list with topics
- ✓ how to expand a query with word embeddings, by exploring synonyms and surrounding words and identifying OCR mistakes

You understand how to explore the newspapers collections based on the filters and the visualisations, as a mean to search for an event and its avatars with other means that adding keywords. We have learned how the filters inform us about the content of the result list and help understand what to expect in them, before further close inspection.

NEXT CHALLENGE: understand the limits of these tools!

02 | HOW TO FIND THE RIGHT PERSON? USES AND LIMITS OF NAMED ENTITIES TO NAVIGATE DIGITISED NEWSPAPERS

Now, we propose to test the strengths and limits of the named entities that have been recognised in the impresso corpus, with the challenge of differentiating Robert Schuman - the politician, from Robert Schumann, the composer

GOAL 1 create two collections of articles, one dealing only with the politician Robert Schuman and one dealing only with Robert Schumann, the composer

GOAL 2 Create a sub collections with articles mentioning only Robert Schumann's pieces, played at concerts (create a timeline/frequency graph with the concerts)

The screenshot shows the Impresso search interface. At the top, there are three tabs: 'SEARCH ARTICLES', 'SEARCH IMAGES', and 'NGRAMS'. Below the tabs, the search term 'robert schu' is entered. A yellow circle with the number '1' highlights the search results list. The first result, '...robert schu', is highlighted with a blue background and has a 'Search in article contents' button. The second result, '...robert schu', has a 'Search in article titles' button. Below these, there are several other results including 'M. Robert Schuman', 'Robert Schumann', and 'Robert Schuman'. A yellow circle with the number '2' highlights the 'Robert Schumann' result under the 'Persons' section. This result has a 'More Topics' button. Below the results, there is a preview of a historical newspaper page with various articles and advertisements. At the bottom of the interface, there is a footer with the text 'select entity Robert Schumann'.

The screenshot shows the Impresso search interface with the following filters applied:

- SEARCH ARTICLES**: Robert Schumann
- SEARCH IMAGES**
- NGRAMS**
- FRONTPAGE**
- EMBEDDINGS**
- PUBLICATION DATE**: SUM (line chart showing article counts from 1740 to 2020, with a peak around 1960)
- FILTER BY COLLECTION**: check one or more collection to filter results
- FILTER BY ARTICLE TYPE**: check one or more article types to filter results
- FILTER BY PERSON**: MENTIONING, ALL OF THE FOLLOWING
 - Robert Schumann (743 results)
 - Franz Schubert (540 results)
 - Ludwig van Beethoven (459 results)
 - Richard Strauss (437 results)
 - Johannes Brahms (419 results)
- FILTER BY LOCATION**: check one or more locations to filter results

3

look at the search filters:
who do you see?

what other named entities are
mentioned in the result list?

4

The screenshot shows the Explore interface with the following filters applied:

- LOCATION**
- COUNTRY**
- PERSON** (selected)
- LANGUAGE**
- TOPIC**
- NEWSPAPER**
- COLLECTION**
- YEAR**
- MONTH**

The results list includes:

- Robert Schumann (7,448 results)
- Franz Schubert (540 results)
- Ludwig van Beethoven (459 results)
- Richard Strauss (437 results)
- Johannes Brahms (419 results)
- Johann Sebastian Bach (320 results)
- Franz Liszt (315 results)
- Maurice Ravel (307 results)
- Richard Wagner (294 results)
- Wolfgang Amadeus Mozart (292 results)

can you spot the second Robert
Schuman? there seems to be an
overlap between the two entities...

the problem at stake : entity link
attribution errors, misspellings by
journalists

How to deal with it?

Go to **INSPECT & COMPARE** and compare the queries for each entity: **Robert Schumann**—the composer (in column A) and **Robert Schuman**—the politician (in column B).

5

Media Monitoring
impresso

Search* (1 filter, 1 term)

Newspapers Topics Inspect & Compare

FAQ Terms of Use Max Mustermann GUEST

QUERY * COLLECTION

Robert Schumann ADD FILTER... robert OPEN IN SEARCH PAGE... (7,448 RESULTS)

YEARS

1860 1880 1900 1920 1940 1960 1980 2000

226

NEWSPAPERS

Gazette de Lausanne	981
Freiburger Nachrichten	980
Journal de Genève	957
Die Tat	941
L'Express	812
La Liberté	810
L'Impartial	689
Luxemburger Wort	251
Neue Zürcher Zeitung	206
Le Peuple, La Sentinelle	170
Confédéré	149
L'indépendance luxembourgeoise	123
d'Letzeburger Land	105
Escher Tageblatt	85
Obermosel-Zeitung	61
Luxembourg (1935)	37

6

62 results in common

Lists of newspapers, named entities and topics for articles which appear both in A and B.

OPEN IN SEARCH PAGE... (62 RESULTS)

YEARS

1860 1880 1900 1920 1940 1960 1980 2000

8

NEWSPAPERS

Luxemburger Wort	16
Die Tat	7
La Liberté	6
Neue Zürcher Zeitung	6
Freiburger Nachrichten	5
D'Unio'n	5
Gazette de Lausanne	4
d'Letzeburger Land	4
Journal de Genève	3
L'Express	2
L'Impartial	2
L'Union	1
Escher Tageblatt	1

Max Mustermann GUEST

QUERY * COLLECTION

Robert Schuman ADD FILTER... robert schuman OPEN IN SEARCH PAGE... (4,575 RESULTS)

YEARS

1860 1880 1900 1920 1940 1960 1980 2000

576

NEWSPAPERS

La Liberté	1,009
Luxemburger Wort	703
Die Tat	483
Journal de Genève	435
d'Letzeburger Land	329
Gazette de Lausanne	317
L'Express	279
Freiburger Nachrichten	243
Neue Zürcher Zeitung	242
L'Impartial	195
Le Peuple, La Sentinelle	139
D'Unio'n	54
Escher Tageblatt	53
Obermosel-Zeitung	34
Confédéré	31
L'indépendance luxembourgeoise	8

which topics do you see related to each entity?
what are the topics of the overlapping articles?

open the results in the search page

V 1.2 . MOD. APRIL 9 , 2020 9 : 54 AM

7 with the help of the topics, distribute the articles between the correct historical figures

8 add all relevant articles to a collection dedicated to **Robert Schuman**—the composer and the others to a **Robert Schumann**—the politician collection.

The screenshot shows the impresso platform interface. At the top, there are navigation links for 'Newspapers', 'Topics', and 'Inspect & Compare'. On the right, there are 'FAQ', 'Terms of Use', and a user account section for 'Max Mustermann GUEST'. Below the header, there are search filters: 'SEARCH ARTICLES', 'SEARCH IMAGES', 'NGRAMS', 'GROUP BY ARTICLE', 'ORDER BY RELEVANCE', and 'DISPLAY AS LIST/TILES'. The search bar contains the query 'Robert Schumann AND Robert Schuman'. The results summary indicates '4 ITEMS SELECTED' and 'CLEAR SELECTION'. The main content area displays five articles found with the topic 'april · mai · märz · samstag · mittwoch...'; mentioning ROBERT SCHUMANN and ROBERT SCHUMAN'. The first article is titled 'Eine Rede Robert Schumans in Genf' from 'Die Tat' on MONDAY, MARCH 23, 1953. It includes a thumbnail of the newspaper page, a brief summary, and links to 'VIEW' and 'ADD TO COLLECTION ...'. The second article is from 'Panorama' on FRIDAY, JUNE 13, 1986, titled 'Panorama'. Both articles have a checked checkbox in their respective 'ADD TO COLLECTION ...' boxes.

with the help of the topics, distribute the articles between the correct historical figures

add all relevant articles to a collection dedicated to **Robert Schuman**—the composer and the others to a **Robert Schumann**—the politician collection.

SAVE / EXPORT >

Create Collection from Search Results

Media Monitoring
impresso of the Part

Search* (2 filters, 5 terms)

Newspapers Topics Impact & Companies

FAQ Terms of Use Max Mustermann GUEST

SEARCH ARTICLES SEARCH IMAGES NGRAMS

Robert Schumann AND Robert Schuman

musik · konzert · radio · beromünster

search for ...

results are filtered when:

CONTAINING AT LEAST ONE OF THE FOLLOWING

- DE musik · konzert · radio · beromünster · programm(6 results) [dropdown]
- DE musik · konzert · orchester · abend · programm(5 results) [dropdown]
- DE ausstellung · kunst · künstler · galerie · museum(4 results) [dropdown]
- de verlag · buch · band · geschichte · werk (7 results) [dropdown]
- de welt · leben · mensch · wort · art (3 results) [dropdown]
- de regierung · afp · sowjetunion · reuter · moskau (3 results) [dropdown]
- de new · york · spanien · dollar · amerika (2 results) [dropdown]
- de antrag · rat · kommission · regierungsrat · stadtrat (2 results) [dropdown]
- de wirtschaft · entwicklung · industrie · arbeit · frage (2 results) [dropdown]
- de mann · hand · kopf · nacht · gesicht (2 results) [dropdown]

MORE...

GROUP

11 art "musik · konzert · radio · beromünster" (11 results)

programm..." or
kunst · kün (MORE ...)

COMPARE ...

SAVE / EXPORT

ORDER BY RELEVANCE

DISPLAY AS LIST TILES

Robert Schumann geboren. Zu seinem 160. Geburtstag legen Smoking in jacquard-Karo, Switzerland, Italian language, Boden%2c Sweden, Barcelona,

Robert Schuman

Freiburger Nachrichten SATURDAY, APRIL 8, 1950 p.4

Trotz Ermordung der Schwiegermutter freigesprochen: R.-Sp. Ein 60-jähriger Berliner wurde auf freien Fuß gesetzt, nachdem ein Westberliner Gericht

LOCATIONS Asbury Park%2c New Jersey, Rome, Cambridge, USA, Zürich District, Bern, Zofingen

PEOPLE Tony Galento, Robert Schumann, Robert Schuman, Alvar Aalto

VIEW ADD TO COLLECTION ...

9

10

Now that we have curated the list of articles mistakenly connected to both Robert Schuman(n), we will take a step back and look at the other results from the original queries.

Repeat the query for the Named Entity “**Robert Schuman**” and exclude the curated articles from the result list.

Next, look for the named entities that appear in the result list: do you notice any that seem not to fit?

Media Monitoring
impresso of the Past

Search* (1 filter, 2 terms)

Newspapers Topics Inspect & Compare

FAQ Terms of Use Max Mustermann GUEST

SEARCH ARTICLES SEARCH IMAGES NGRAMS GROUP BY ARTICLE ORDER BY RELEVANCE DISPLAY AS LIST TILES

Robert Schuman AND National La ▾

Robert sch 🔎

Frontpage EMBEDDINGS ⓘ

PUBLICATION DATE Number of articles per year

% SUM ⓘ

ADD FILTER ...

FILTER BY COLLECTION ... you haven't saved any result item in your collection +

FILTER BY ARTICLE TYPE check one or more article types to filter results +

FILTER BY PERSON ⓘ results are filtered when: RESET

MENTIONING ALL OF THE FOLLOWING

Robert Schuman (4,575 results) 🔎

National Lampoon's Animal House (1,319 results) 🔎

Albert G. Hill (1 result) 🔎

18 articles found mentioning ROBERT SCHUMAN and NATIONAL LAMPOON'S ANIMAL HOUSE

Luxemburger Wort FRIDAY, APRIL 30, 1943 p.4

Ärztdienst in Düdelingen Am Sonntag steht in dringenden Fällen der Arzt, Herr Dr. Robert Schuman, Adolf Hitler-Straße, Fernrufnummer 15-42, der Düdelinger Bevölkerung zu Diensten. Geöffnet ist die Apotheke Emil Faber, die auch den Nachtdienst versieht.

Das Standesamt Düdelingen meldet In der Woche vom 19. bis zum 26. April 1943 wurden auf dem Standesamt Düdelingen folgende Personenstandsfälle beurkundet: Geburten: 2 Knaben; Heiraten: 2; Sterbefälle: 2 weibliche. Die

PEOPLE Robert Schuman, National Lampoon's Animal House

VIEW ADD TO COLLECTION ...

Aerztlicher Sonntagsdienst in Ofidelingen Luxemburger Wort SATURDAY, APRIL 10, 1943 p.4

Aerztlicher Sonntagsdienst in Düdelingen Am morgigen Sonntag steht in dringenden Fällen der Arzt, Herr Dr. Robert Schuman, Adolf Hitler-Straße, Fernrufnummer 15-42, der Düdelinger Bevölkerung zu Diensten. Geöffnet ist die Apotheke Emil Faber, die auch den Nachtdienst versieht.

Abschluß-Veranstaltung des Düdelinger Kunstkreises Als Abschluß der diesjährigen Kunstkreisveranstaltungen findet am Dienstag, 13. April, 20 Uhr, im Saale des Kameradschaftshauses in Düdelingen, eine sinfonische

PEOPLE Robert Schuman, National Lampoon's Animal House

VIEW ADD TO COLLECTION ...

Sonntagsdienst in Düdelingen Luxemburger Wort SATURDAY, MARCH 27, 1943 p.4

schmerzliche Krankheit hatte Krankenlager gefesselt. Ged ohne Klage ertrug sie ihre I Gevatter Tod sie in ein bes Seite abrief Frau Schneider

11

do you see any odd names that do not fit?
what about the national lampoon's animal house?
select this entity as an additional filter (filter by person, mentioning ALL of the following)

who is the Robert Schuman mentioned in that context? what is his job? are the articles referring to the politician? if yes, let's continue with him! if not, exclude this entity from the search.

what are we looking at now? What do these 18 articles contain?

The screenshot shows a search interface with a sidebar on the left containing 'MS', 'GROUP', '7,448', a magnifying glass icon, a plus sign icon, and a 'RESET' button. The main area has tabs for 'Newspapers', 'Topics', and 'Inspect & Compare'. A yellow circle with the number '12' is overlaid on the top right. The search bar at the top says 'Select among 16678 options to refine your search for type:' followed by buttons for 'LOCATION', 'COUNTRY', 'PERSON' (which is highlighted in black), 'LANGUAGE', 'TOPIC', 'NEWSPAPER', 'COLLECTION', 'YEAR', and 'MONTH'. Below this is a list of names with their result counts: Robert Schumann (7,448 results), Franz Schubert (540 results), Ludwig van Beethoven (459 results), Richard Strauss (437 results), Johannes Brahms (419 results), Johann Sebastian Bach (320 results), Franz Liszt (315 results), Maurice Ravel (307 results), Richard Wagner (294 results), and Wolfgang Amadeus Mozart (292 results). At the bottom are navigation arrows and a 'CLOSE' button.

look at the topics: which could be indicative of concert programs on the radio?

(you can also filter by language, to focus either on **GERMAN** or **FRENCH** to continue)

explore the topics and select the most relevant one to find radio programs containing announces of Robert Schumann concerts, select the topic and create a collection based on that query

12

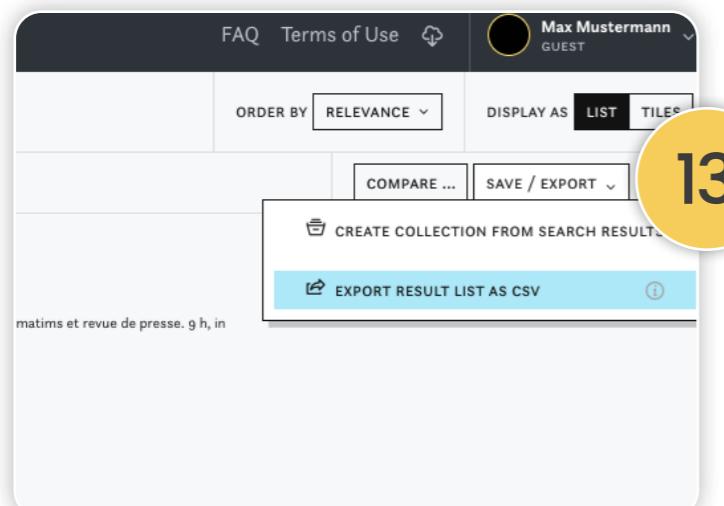
Repeat the same exercice with the query for the named entity **Robert Schumann**.

Curate the result list with the means of excluding non relevant named entities.

(Bonus: for instance, where do **Gérard Depardieu** and **Robert Schumann** appear together and why?)

Next, we will narrow down the result list with the help of **TOPICS**.

The screenshot shows a dropdown menu for 'Robert Schumann' with a search bar and a magnifying glass icon. Below it is a 'FILTER BY TOPIC' section with a 'check one or more topics to filter results' instruction and a minus sign icon. The list includes: fr musique · concert · piano · orchestre · programme (1,853 results), fr musique · concert · suisse · monde · danse (1,599 results), de musik · konzert · radio · beromünster · programm (1,426 results), de musik · konzert · orchester · abend · programm (975 results), fr soir · matin · midi · nuit · journal (743 results), fr journal · monde · vie · direct · série (657 results), de tagesschau · sport · bild · les · deutschland (574 results), fr vie · monde · auteur · foi · sen (490 results), de theater · oper · sonntag · samstag · stadttheater (489 results), and fr film · cinéma · semaine · jean · john (485 results). At the bottom is a 'MORE...' button.



final step of the challenge: export the curate collection of radio programs announcing concerts with Robert Schumann's music, as csv file

open the zip file and extract the csv file.

Import the csv file in your preferred spreadsheet software, with the parameters of **UTF-8** for the content and set **SEMICOLON** as separator

EXP-1982-11-1fr	ft I RAI	368	CH	EXP	EXP-19	24	1	0.0	1982	FALSE	1982-11-13T00:00:00Z	Luc Du Lausan ft I RAI OpenP
EXP-1982-11-1fr	RADIO	632	CH	EXP	EXP-19	29	1	0.0	1982	FALSE	1982-11-13T00:00:00Z	Philipp Lausan RADIO OpenP
EXP-1982-11-1fr		1819	CH	EXP	EXP-19	29	1	0.0	1982	FALSE	1982-11-24T00:00:00Z	Pierre SUISSE c ove r OpenP
EXP-1976-01-1fr	RADIO	397	CH	EXP	EXP-19	13	1	0.0	1976	FALSE	1976-01-05T00:00:00Z	R. Schumann RADIO OpenP
EXP-1979-08-1fr	RADIO	476	CH	EXP	EXP-19	11	1	0.0	1979	FALSE	1979-08-04T00:00:00Z	Albert Avigno RADIO OpenP
EXP-1955-09-1fr	Un «pé	771	CH	EXP	EXP-19	9	1	0.0	1955	FALSE	1955-09-13T00:00:00Z	M. Edg Rabat Un « p OpenP
EXP-1989-11-1fr		225	CH	EXP	EXP-19	43	1	0.0	1989	FALSE	1989-11-23T00:00:00Z	premie INDE S INDE - OpenP
EXP-1987-08-1fr	; NU	393	CH	EXP	EXP-19	25	1	0.0	1987	FALSE	1987-08-15T00:00:00Z	Robert Zurich ; NUI OpenP
EXP-1987-08-1fr	RADIO	387	CH	EXP	EXP-19	23	1	0.0	1987	FALSE	1987-08-19T00:00:00Z	Robert Zurich RADIO OpenP
EXP-1980-04-1fr	RADIO	492	CH	EXP	EXP-19	25	1	0.0	1980	FALSE	1980-04-12T00:00:00Z	Robert Schum RADIO OpenP
EXP-1980-05-1fr		1601	CH	EXP	EXP-19	21	1	0.0	1980	FALSE	1980-05-19T00:00:00Z	C. Paql Suisse AVEC L OpenP
EXP-1972-02-1fr		772	CH	EXP	EXP-19	23	1	0.0	1972	FALSE	1972-02-19T00:00:00Z	Isabelli Suisse Samed OpenP
EXP-1903-11-1fr		204	CH	EXP	EXP-19	2	1	0.0	1903	FALSE	1903-11-24T00:00:00Z	Nina FALIERO Grand OpenP
EXP-1879-02-1fr	DIXIEN	354	CH	EXP	EXP-18	4	1	0.0	1879	FALSE	1879-02-13T00:00:00Z	TEMPL Bâle E DIXIEN OpenP
EXP-1973-11-1fr	il^] [C	805	CH	EXP	EXP-19	29	1	0.0	1973	FALSE	1973-11-03T00:00:00Z	J. Greg Suisse il^] [OpenP

and now, you can use the metadata to create your timeline

CRITICAL USAGE UNLOCKED!

You have learned to use the features with caution:

- ✓ Named entities are useful to collect relevant articles at the same time, we have to keep in mind that they result from an automated process that has been validated
- ✓ Other automated annotations (other named entities, topics) can help to validate one another: they function as indications rather than evidence
- ✓ Named entities and topics as well as metadata can be used to create an overview and detect potentially relevant or irrelevant results
- ✓ We can store the curated results in collections

NEXT CHALLENGE: transpose historical research questions via the interface's features.

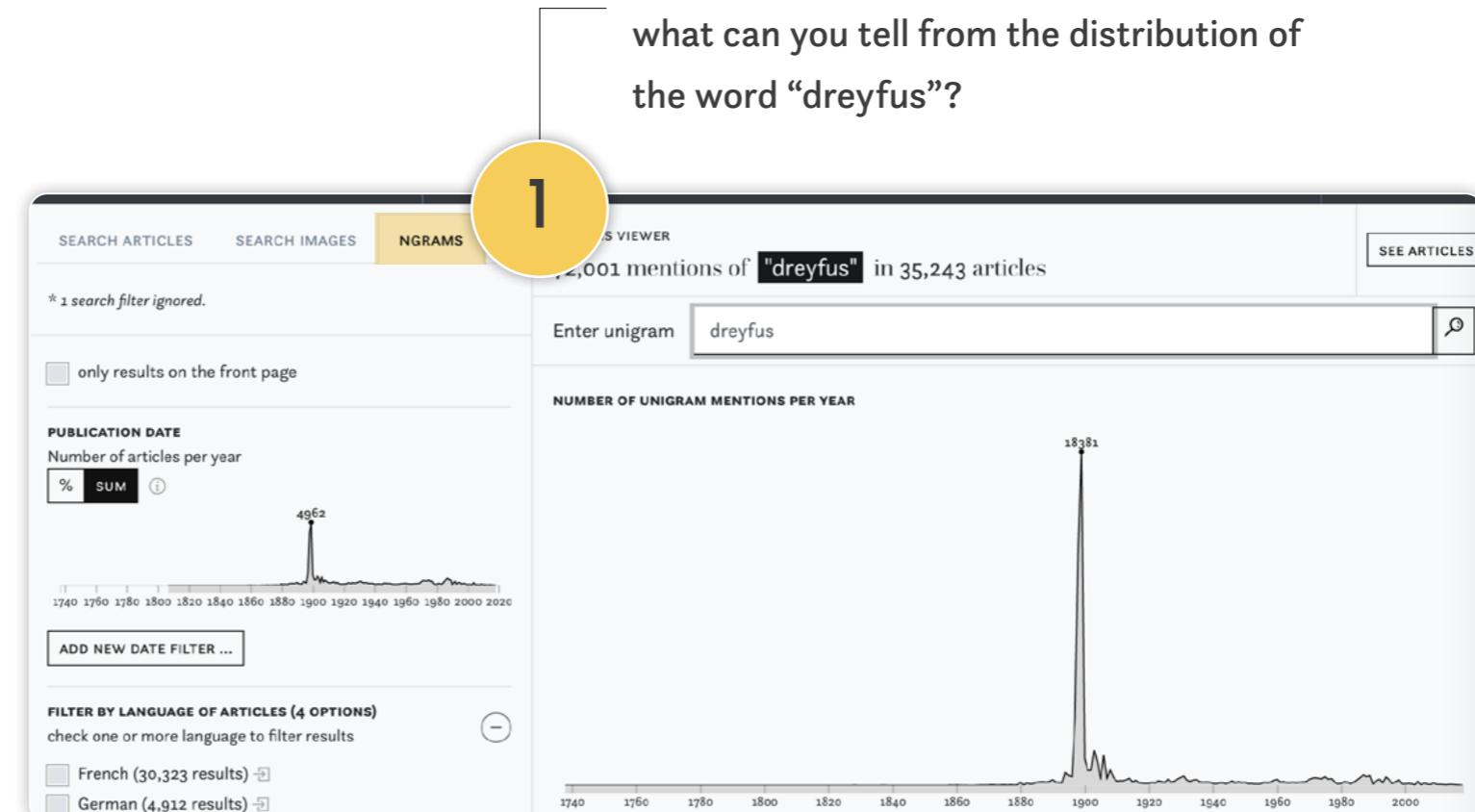
03 | RECONSTRUCT THE MEDIA COVERAGE OF AN EVENT IN DIGITISED NEWSPAPERS: THE DREYFUS TRIAL.

In this challenge, we propose to combine the features that have been presented earlier to you, in a critical use, to try and analyse the Swiss and Luxembourgish coverage of the Dreyfus affair.

3.1/ How to find articles dealing with the Dreyfus affair? In other words, how much is the word “dreyfus” synonymous with the Dreyfus affair?

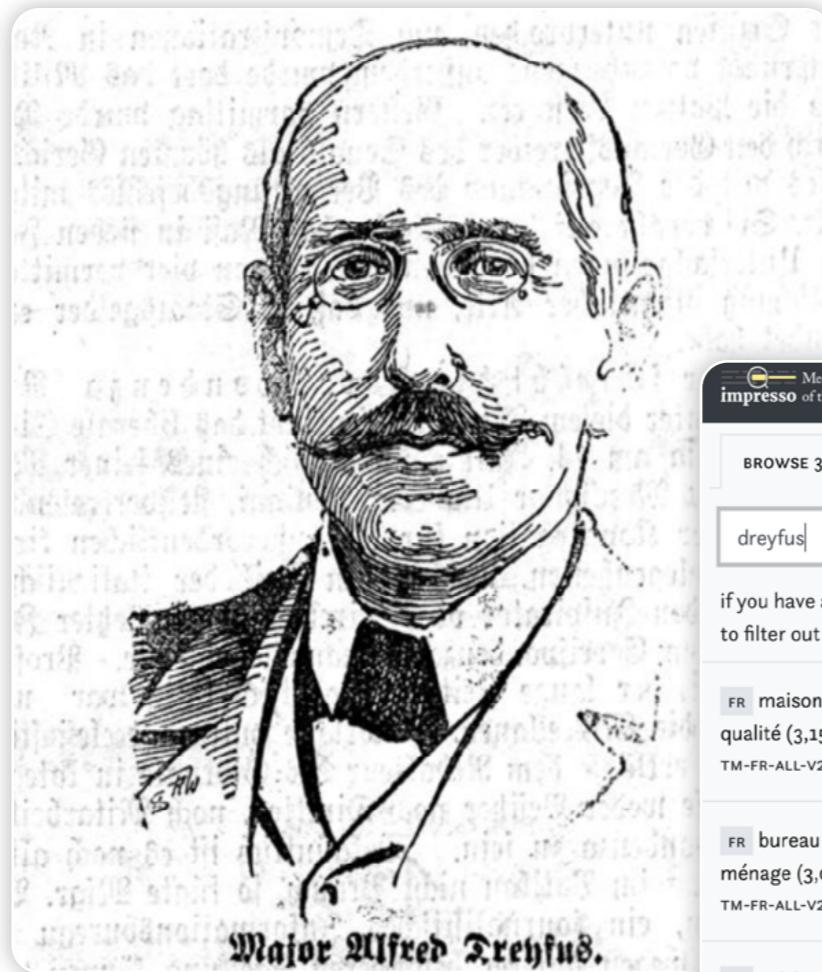
GOAL identify articles dealing with the Dreyfus trial and revision.

Start with the **NGRAM** page:
what can you tell from the distribution of the word “dreyfus”?



The screenshot shows the impresso interface with a search bar at the top containing the query "Search* (4 ...)" and a yellow button labeled "SEARCH IMAGES". Below the search bar are two buttons: "SEARCH ARTICLES" and "SEARCH IMAGES". A yellow circle with the number "2" is overlaid on the interface. At the bottom left, there is a note: "* 1 search filter ignored."

Apply the same query in the IMAGE SEARCH: what can you tell from the information associated with the word “dreyfus”?



The screenshot shows the impresso Topics page with a search bar at the top containing "Search* (3 ...)" and a yellow button labeled "Topics". Below the search bar are buttons for "Newspapers", "Topics", "Inspect & Co...", and "Text reuse". A yellow circle with the number "3" is overlaid on the interface. On the right side, there is a user profile for "Max Mustermann GUEST".

EXPLORE THE LIST OF TOPICS

Visualize 0 topics and their 0 relationships

BROWSE 300 TOPICS 0 VISUALIZED

dreyfus

if you have a search filter, you can use this one to filter out stuff;

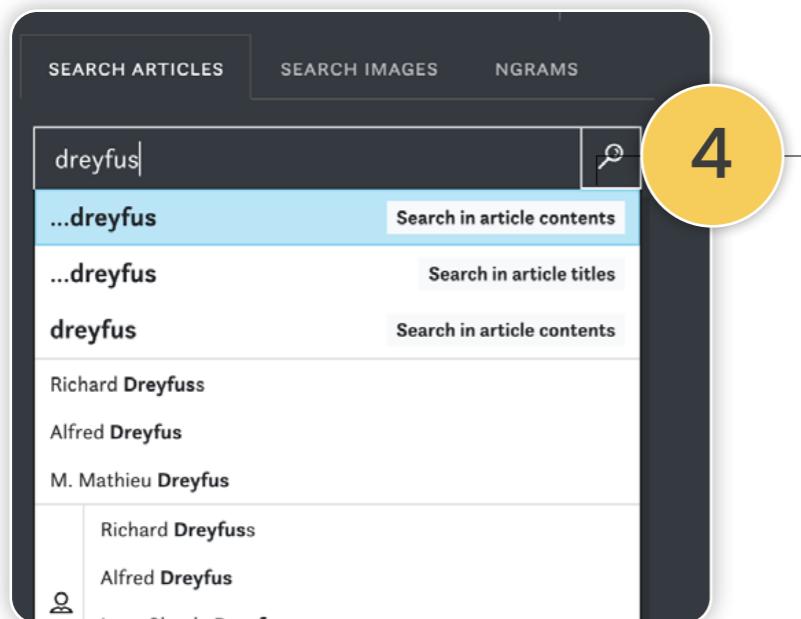
- FR maison · vente · choix · magasin · qualité (3,156,233 articles)
- TM-FR-ALL-V2.0
- FR bureau · place · suite · famille · ménage (3,076,240 articles)
- TM-FR-ALL-V2.0
- FR cuisine · salle · appartement · confort · parc (2,819,261 articles)
- TM-FR-ALL-V2.0
- FR mois · numéro · carte · adresse · poste (2,406,064 articles)
- TM-FR-ALL-V2.0
- FR état · lit · bureau · table · bois

COLORED BY LANGUAGE

CONNECTED SHARE AT LEAST ONE ARTICLE

SIZE BY NUMBER OF ARTICLES

ZOOM & PAN RESET



The screenshot shows a search interface with three tabs: 'SEARCH ARTICLES', 'SEARCH IMAGES', and 'NGRAMS'. The 'SEARCH ARTICLES' tab is active. In the search bar, the word 'dreyfus' is typed. Below the search bar, there are several search suggestions and results:

- ...dreyfus (Search in article contents)
- ...dreyfus (Search in article titles)
- dreyfus (Search in article contents)
- Richard Dreyfuss
- Alfred Dreyfus
- M. Mathieu Dreyfus
- Richard Dreyfuss
- Alfred Dreyfus

Now go to the search page, search again for “dreyfus”: what is the outcome? What are the **TOPICS** connected to that word?

What are the **ARTICLE TYPES** that may not be relevant to find articles dealing with the Dreyfus affair?

What are the **LOCATIONS** that could be relevant to find articles dealing with the Dreyfus affair?



The screenshot shows a search interface with three tabs: 'SEARCH ARTICLES', 'SEARCH IMAGES', and 'NGRAMS'. The 'SEARCH ARTICLES' tab is active. In the search bar, the search term 'dreyfus' is highlighted in a yellow box. Below the search bar, there are search filters and results:

- GROUP BY:** ARTICLE ▾
- ORDER BY:** RELEVANCE ▾
- DISPLAY AS:** LIST (selected) ▾ TILES
- RESULTS:** 35,243 articles found containing **dreyfus**
- FRONTPAGE**
- FIND SIMILAR WORDS** ⓘ
- PUBLICATION DATE:** Number of articles per year (Chart showing a sharp peak around 1899)
- ADD NEW DATE FILTER ...**
- FILTER BY ARTICLE TYPE (6 OPTIONS):** check one or more article types to filter results
 - article (28,664 results) ↗
 - advertisement (2,500 results) ↗

Result Preview:

Die Dreyfus-Affäre.
Bürger- und Beamten-Zeitung ↗ THURSDAY, JUNE 22, 1899 – p.2

Die Akten über Dreyfus sind in Rennes eingetroffen und werden in einem von einem Pfandhouse gemieteten G.ldschranc aufbewahrt. Der Saal, in dem die Verhandlungen stattfinden werden, grenzt an das Militärgefängnis; durch die Mauer wird eine Zelle gekrochen werden, damit Dreyfus zur Vernehmung nicht die Straße zu betreten braucht. Eine Anzahl Polizisten ist am Dienstag hier eingetroffen. Dreyfus wird Ende der Woche erwartet. Eine Zelle steht bereit. Die Zusammensetzung des Kriegsgerichts ist nicht bekannt. In der Stadt herrscht Ruhe. Ferner wird dem „Journal“ aus Rennes gemeldet: Die bevorstehende Ankunft Dreyfus' hat einen ungeheuren Fremdenzufluss veranlaßt. Alle Gasthäuser sind überfüllt. Es wurden umfassende Maßregeln getroffen, um Ausgebungen zu verhindern. Dem „Figaro“ wird aus Brest berichtet, Dreyfus werde wahrscheinlich in St. Malo gelandet werden.

LOCATIONS: Rennes ↗, Brest, France ↗

Die **Dreyfus**-Affäre. Die Akten über **Dreyfus** sind in Rennes eingetroffen werden, damit **Dreyfus** zur Vernehmung nicht die Straße zu betreten braucht. Eine Anzahl Polizisten ist am Dienstag hier eingetroffen. **Dreyfus** wird Ende der Woche erwartet. Seine Zelle steht bereit. Die Zusammensetzung : Die bevorstehende Ankunft **Dreyfus** hat einen ungeheuren Fremdenzufluss veranlaßt. Alle Gasthäuser sind überfüllt berichtet, **Dreyfus** werde wahrscheinlich in St. Malo gelandet werden.

Dreyfus-TM in Rennes X

5

Create a collection of articles containing “dreyfus” in their **TITLES**, that are **NOT ADVERTISEMENTS** and that contain the **LOCATION “Rennes”**.

What do you notice? Which topics and persons appear in the result list?

The screenshot shows the Impresso search interface with the following parameters:

- SEARCH ARTICLES**
- SEARCH IMAGES**
- NGRAMS**
- GROUP BY ARTICLE**
- ORDER BY RELEVANCE**
- DISPLAY AS LIST**

Search filters include:

- Dreyfus in title-affaire in content
- Rennes
- dreyfus
- Frontpage
- FIND SIMILAR WORDS
- PUBLICATION DATE: Number of articles per year (SUM selected)
- ADD NEW DATE FILTER
- FILTER BY ARTICLE TYPE (1 OPTION): article (154 results)
- FILTER BY COLLECTION (6 OPTIONS): results are filtered when

The results section displays:

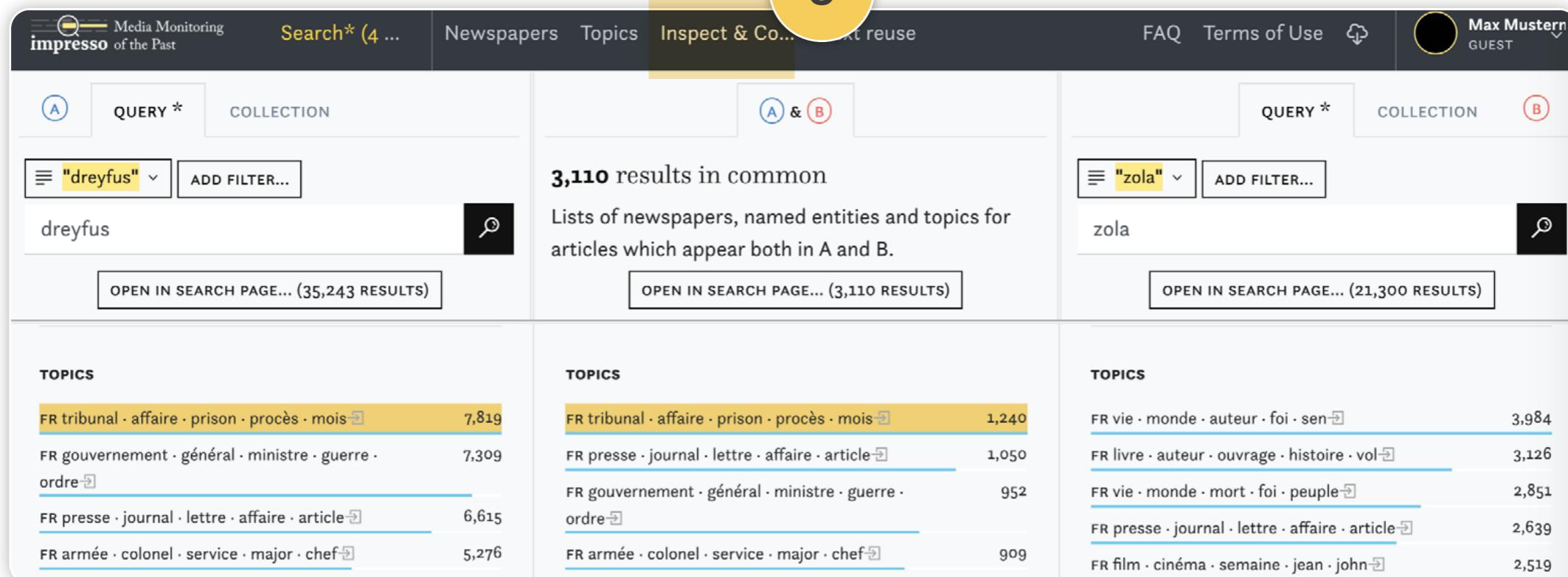
- 154 articles found saved in Dreyfus in title-affaire in content; mentioning RENNES**
- L'AFFAIRE DREYFUS** (L'Express, Friday, December 11, 1959 - p.8)
- LOCATIONS**: French cruiser Algérie, Rennes
- PEOPLE**: Hubert-Joseph Henry
- Dreyfus in title-affaire in content**
- VIEW** and **ADD TO COLLECTION ...**
- L'AFFAIRE DREYFUS** (L'Express, Friday, August 2, 1907 - p.3)
- LOCATIONS**: Vincennes, Rennes, Paris
- Dreyfus in title-affaire in content**

3.2/ Which are the most relevant persons mentioned with Alfred Dreyfus?

How often were Emile Zola and Alfred Dreyfus co-mentioned in the press?

When was Zola mentioned first in the press, as available in the impresso corpus? When was he mentioned the most? When was he mentioned first in the context of the Dreyfus affair?

go to the **INSPECT AND COMPARE** page and compare the coverage of the words “dreyfus” and “zola”



what can you say about the **TOPICS** related to each word?

7

do the same with the named entities [Alfred Dreyfus](#) and [Emile Zola](#)

The screenshot displays three separate search interfaces, labeled A, B, and C, each showing results for a specific named entity.

Interface A (Left): Search query is "Alfred Dreyfus". Results show 1,097 results. The top result is "drey".

TOPICS	
FR tribunal · affaire · prison · procès · mois	450
FR presse · journal · lettre · affaire · article	371
FR armée · colonel · service · major · chef	365
FR gouvernement · général · ministre · guerre · ordre	338
FR vie · monde · mort · foi · peuple	213
FR conseil · commission · projet · loi · rapport	197
DE regierung · paris · frankreich · minister · kammer	164
FR président · fête · ville · lieu · discours	152
FR président · ministre · gouvernement · général · chef	152
FR question · point · pays · conseil ·	149

Interface B (Middle): Search query is "Alfred Dreyfus" AND "Emile Zola". Results show 107 results in common.

TOPICS	
FR tribunal · affaire · prison · procès · mois	44
FR presse · journal · lettre · affaire · article	39
FR armée · colonel · service · major · chef	36
FR gouvernement · général · ministre · guerre · ordre	26
FR vie · monde · mort · foi · peuple	25
FR vie · monde · auteur · foi · sen	18
FR président · fête · ville · lieu · discours	17
FR fille · main · voix · foi · temps	15
FR livre · auteur · ouvrage · histoire · vol	15
FR président · ministre · gouvernement · général · chef	14

Interface C (Right): Search query is "Emile Zola". Results show 5,945 results. The top result is "zola".

TOPICS	
FR film · cinéma · semaine · jean · john	978
FR vie · monde · auteur · foi · sen	972
FR livre · auteur · ouvrage · histoire · vol	859
FR vie · monde · mort · foi · peuple	771
FR presse · journal · lettre · affaire · article	687
FR scène · théâtre · spectacle · pièce · public	667
FR tribunal · affaire · prison · procès · mois	573
FR gouvernement · général · ministre · guerre · ordre	547
FR fille · main · voix · foi · temps	538
FR président · fête · ville · lieu · discours	521

look at the articles which contain both named entities: how many are actually covering the affair?

CHALLENGE COMPLETED!

You have learned:

- ✓ how to read the ngram and frequency lines
- ✓ what the impresso corpus contains
- ✓ how to query the corpus with key-words, persons (named entities)
- ✓ how to create and curate collections
- ✓ how to inspect and compare queries and collections
- ✓ how to curate a result list with topics
- ✓ how to expand a query with word embeddings, by exploring synonyms and surrounding words and identifying OCR mistakes

You understand how to explore the newspapers collections based on the filters and the visualisations, as a mean to search for an event and its avatars with other means that adding keywords. We have learned how the filters inform us about the content of the result list and help understand what to expect in them, before further close inspection.

NEXT CHALLENGE: understand the limits of these tools!

DIGITAL HUMANIST LEVEL UNLOCKED!

You know now how to use Natural Language Processing and visualisations to query and analyse digitised historical material:

- ✓ You learned how to interpret a peak in frequencies:
how to differentiate keyword hits from the news coverage of an event (Dreyfus as a string of character vs. the Dreyfus Affair)
- ✓ You learned how the topic models were computed,
what they entail and what not (nouns vs. names)
- ✓ You learned how to use historical information to find relevant articles (location of an event, persons related to specific dimension of an event)

CONGRATULATIONS! you are now an expert user of digitised newspapers interfaces!