

LNB Attīstības departamenta 2019./2020.gada aktualitātes

Augstskolu, koledžu un speciālo bibliotēku apaļā galda diskusija

26.02.2020, Silva Vucena

- **Latvijas Nacionālā digitālā bibliotēka**
 - Akadēmiskais repozitorijs, datu repozitorijs un akadēmiskais godīgums
 - **Digitālā pētniecība**
- Par aktuālo īsumā

Viss digitālais saturs vienā meklētājā

690 135 periodiskie izdevumi, 16 999 grāmatas, 77 382 attēli, 2202 kartes, 12 835 skaņu un video ieraksti



Izvērstā meklēšana

Digitālie objekti

LNB katalogs

Viss krājums

digitalbiblioteka.lv

Kolekcijas



Rādīt visas

Latvijas tīmekļa vietņu arhīvs

LATVIJAS NACIONĀLĀ DIGITĀLĀ BIBLIOTĒKA

VIENOTAIS MEKLĒTĀJS >

GRĀMATAS >

PERIODIKA >

ZUDUSĪ LATVIJA >

AUDIO >

CITI RESURSI >



Latvijas tīmekļa vietņu arhīvs

Rasmošana - Interneta vietņu saturs automatizēta vākšana, vadoties pēc vietnes adreses (angļu val. *Harvesting* - ražas novākšana)

140 997 224

Objektu skaits arhīvā

9.5 TB

Kopējais izmērs

Tīmekļa lapas

89 532 258
2.85 TB

Attēli

29 923 798
2.29 TB

Dokumenti

2 063 572
1.42 TB

Audio

88 048
826.64 GB

Video

29 598
953.28 GB

Piekluve arhivētajām vietnēm tiek nodrošināta no Latvijas Nacionālās bibliotēkas telpām [Pieklūt Web arhīvam](#)



Akadēmiskais repozitorijs

Mērķis – uzglabāt vienkopus un padarīt pieejamus Latvijā izstrādātos akadēmiskos darbus.

2020. gada uzdevums:

Akadēmiskā repozitorija ACADEMIA metadatu atbilstība OAI-PMH protokolam un satura agregācija nozīmīgākajos repozitoriju agregatoros - OpenAIRE, BASE un CORE

Pētniecības **datu** repozitoriju tīkls

FAIR principi:

- atrodams (Findable)
- pieejams (Accessible)
- sadarbspējīgs (Interoperable)
- atkārtoti izmantojams (Reusable)

Akadēmiskais godīgums

- izglītojošie un metodiskie resursi
- plaģiātisma kontroles IT risinājumi
- akadēmiskā godīguma un ētikas komisija
- akadēmiskā godīguma politika, normatīvais regulējums

Datu repozitorijs

Akadēmiskais godīgums

Digitālā pētniecība



Digitālu avotu, digitālu pētniecības metožu, digitālas publicēšanas un saglabāšanas risinājumu izmantojums pētniecības mērķu sasniegšanai.

Datizrace ir lielu datu kopu analīze, kas tiek veikta ar mērķi atrast sakarības un attiecības, kas nebūtu identificējamās bez tehnoloģiju palīdzības.

Tekstizrace ir datizrace teksta veida datos.

2019. un 2020. gadā

Tekstu korpusi:

- Latviešu senāko romānu korpus 1879 – 1920
- Laikrakstu korpusi – «Latvijas Vēstnesis», «Jaunākās Ziņas»
- Korpusu veidošana pēc pieprasījuma (pētnieks var pasūtīt nepieciešamo korpusu)

Jaunas, datu analīzei un vizualizācijai paredzētas saskarnes prototips LNDB:

- Iespēja atlasīt objektus no LNDB vai strādāt ar gataviem korpusiem
- Vārdu lietojuma statistika, konkordances, korpusu salīdzinoša analīze, vizualizācijas
- Pieejamība plašam lietotāju lokam bez programmēšanas iemaņām

2nd Baltic Summer School
of Digital Humanities

Essentials of Coding and Encoding

23 – 26 July 2019
National Library of Latvia, Riga
digitalhumanities.lv/bssdh/2019

Latvijas
Nacionālā
bibliotēka
simts

LMI
LĪTERRĀTU
POLIĀRĀ
UN AKUSTRĀ
MĀKSLA

LATVIJAS
UNIVERSITĀTE
1919



Nordplus

CLARIN

Digitālā pētniecība LNB

DHN2020 Conference
17–20 March 2020, Riga

SEARCH ARTICLES SEARCH IMAGES NGRAMS

search for ...



Media Monitoring of the Past

Digitalizēto laikrakstu korpusu
pētniecības rīks

IMPRESSO DATA RUNDOWN

76 newspapers collected,
600,919 issues,
5,429,656 pages scanned,
47,798,468 content items identified,
3,462,799 images,
12,493,358,703 words.

2 countries of publication
530,086 named entities disambiguated

More? Check on our [blog](#)

info @ impresso-project [dot] ch
project website: impresso-project.ch
github: [impresso](https://github.com/impresso)
twitter: [@impressoproject](https://twitter.com/impressoproject)

Mining 200 years
of historical newspapers ⓘ

How can newspapers help understand the past? How to explore them?

For legal reasons not all content is available in Open Access.
To gain full access, simply sign this [Non-Disclosure-Agreement](#) and email it to info@impresso-project.ch

... then here are a few examples to get you started!

LINES: OFF

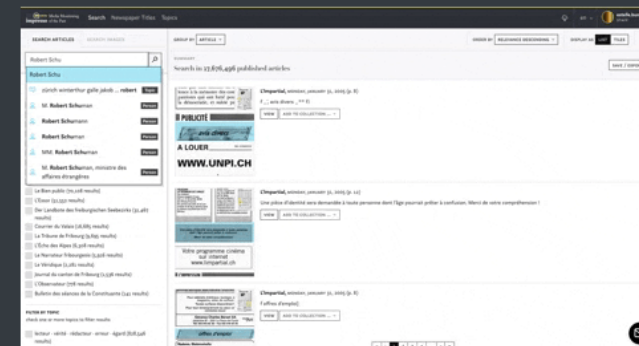
DARK MODE: ON

Reconstruct the coverage of the composer Robert Schumann

Explore the newspapers' mentions of specific persons or locations

PERFORM THIS QUERY

WATCH THE VIDEO

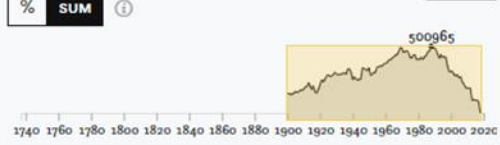


SEARCH ARTICLES SEARCH IMAGES **NGRAMS**

24451 mentions of "lettonie" [SEE ARTICLES](#)

lettonie

PUBLICATION DATE
Number of articles per year (filtered) [RESET](#)



PUBLISHED BETWEEN ▾

1900-01-01 2018-12-31

Monday, January 1, 1900 Tuesday, January 1, 2019

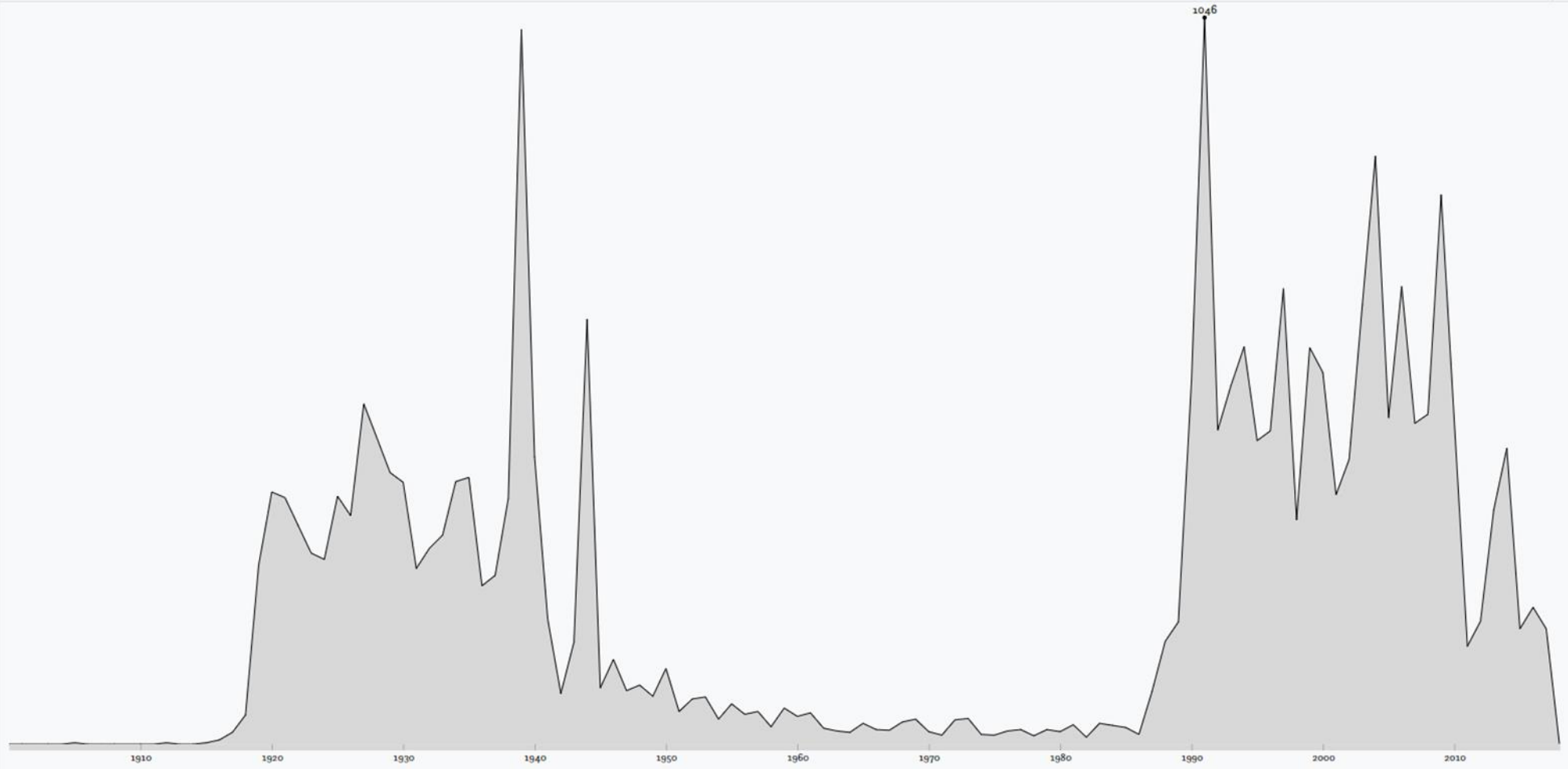
FILTER BY ARTICLE TYPE ⓘ
check one or more article types to filter results (+)

FILTER BY TOPIC ⓘ
check one or more topics to filter results (+)

FILTER BY LANGUAGE OF ARTICLES ⓘ
check one or more language to filter results (+)

FILTER BY NEWSPAPER TITLES ⓘ
check one or more newspaper to filter results (+)

FILTER BY COUNTRY OF PUBLICATION ⓘ
check one or more countries to filter results (+)

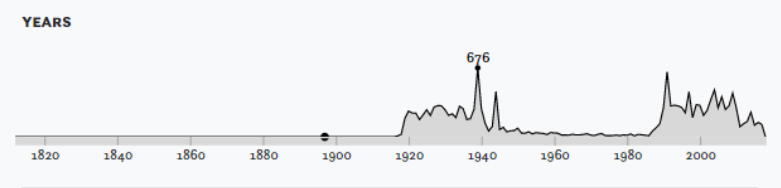


N-grammu vizualizēšanas rīks - meklētā vārda pieminējumu skaita attēlošana uz laika skalas

QUERY * **COLLECTION**

"Lettonie"

lettonie



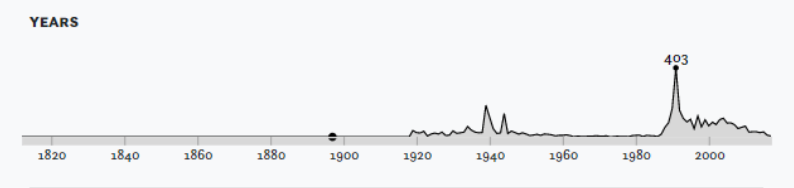
NEWSPAPERS

L'Express	3,386
La Liberté	3,204
Journal de Genève	3,072
L'Impartial	3,069
Gazette de Lausanne	1,783
Le Peuple, La Sentinelle	560
L'indépendance luxembourgeoise	304
Luxembourg (1935)	210
Confédéré	134
Solidarité	55
La lutte syndicale	53
d'Letzeburger Land	47
L'Essor	26
Luxemburger Wort	25
Obermosel-Zeitung	12
Freiburger Nachrichten	8
Neue Zürcher Zeitung	6
Fischer Tageblatt	5

A & B

3,911 results in common

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



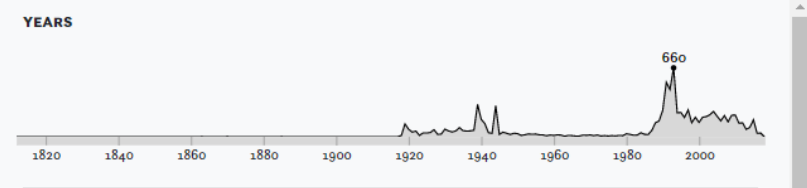
NEWSPAPERS

L'Express	1,014
L'Impartial	832
Journal de Genève	775
Gazette de Lausanne	560
La Liberté	438
Le Peuple, La Sentinelle	143
Luxembourg (1935)	52
Confédéré	26
d'Letzeburger Land	26
L'indépendance luxembourgeoise	19
Solidarité	9
L'Essor	8
La lutte syndicale	8
Luxemburger Wort	1

QUERY * **COLLECTION**

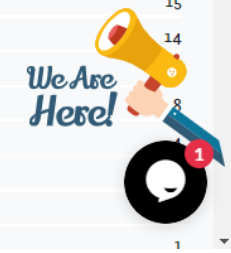
"Estonie"

estonie



NEWSPAPERS

L'Express	2,436
L'Impartial	2,067
Journal de Genève	1,570
Gazette de Lausanne	1,267
La Liberté	1,250
Le Peuple, La Sentinelle	263
Luxembourg (1935)	90
d'Letzeburger Land	58
Confédéré	53
L'indépendance luxembourgeoise	45
La lutte syndicale	15
L'Essor	14
Solidarité	8
Luxemburger Wort	8
L'Union	1
Obermosel-Zeitung	1
Freiburger Nachrichten	1
Die Tat	1



Latvijas un Igaunijas pieminējumi salīdzinājumā

SEARCH ARTICLES SEARCH IMAGES NGRAMS

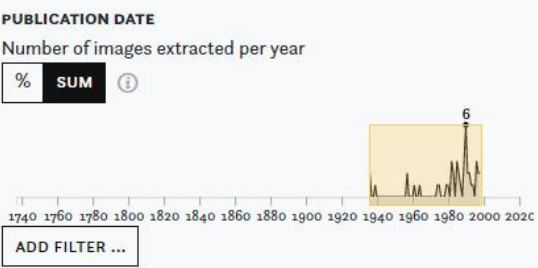
start with random results page

44 images found where title includes

ORDER BY

– "avion"
avion

Frontpage



FILTER BY NEWSPAPER TITLES
check one or more newspaper to filter results



Cet avion américain a encore la chance de voler. Peut-être que plus tard, il rejoindra, lui aussi, la base de Davis-Monthan en Arizona.
Gazette de Lausanne MONDAY, FEBRUARY 2, 1998p.4



Cet avion américain a encore la chance de voler. Peut-être que plus tard, il rejoindra, lui aussi, la base de Davis-Monthan en Arizona.
Journal de Genève MONDAY, FEBRUARY 2, 1998p.4



JacquesChirac lors d'une descente d'avion: une image que l'on ne risque pas de voir en Suisse cette année.
JACQUES WIIT/SIPA PRESS
Gazette de Lausanne FRIDAY, FEBRUARY 14, 1997p.1



JacquesChirac lors d'une descente d'avion: une image que l'on ne risque pas de voir en Suisse cette année.
JACQUES WIIT/SIPA PRESS
Journal de Genève FRIDAY, FEBRUARY 14, 1997p.1



«Un appareil au sol coûte de l'argent Un avion ne rapporte que lorsqu'il vole. Dès lors, il faut



Un voyage aux Etats-Unis sans avoir à déboursier le du billet d'avion

Automātiska attēlu meklēšana laikrakstos



Ģeeweete jaunlaiku sportos.

Sporteem jo deenas jo wairak isplatotees, ari ģeeweete nem winos arween dsihwaku dalibu. Protams, ka lihds ar ģeho parahdibu ţa- zelaš ari dašči jautajumi, kuri no weena waj otra stah- wolka jau wairakfabrt dedšigi pahrrunati. Waj ģeeweete, no- darbodamas ar leetam, kaš rahdaš nodomataš — wišmaj pirmā kahttibā — wiħree- ţheem, nedodaš breefmāš ļaudet daudj no ţawāš daudši- natas „ģeewiţķtibāš“, waj, nododamas ţiţķķeem wingri- najumeem, ar kaħdeem ţporti mehds buht ţaweenoti, wina negrehko pat pret ţawu weţe- libu, neapgruhtina ţawu dsi- muma ihpatnibu apmeerina- ţhanu, neţaboja daiktuma garţchu un eţtetiţko eeţpaidu uš otru dšimumu. Ģaiţča leeta, ka ţchaħdu jautajumu apţpreeţchanai arween eeeš laba teeţa tihri ţubjektiwu juhtu, tā ka no atbildes, kaš wiţuš pahrlēezinatu un no wiţeem deriga padoma te newar buht runāš. Wiš- praktiţkato želu ţchai leetā iţ- raudsiķees frantţchu eewehro- jamaiš ţħurnalš „Rev. de Rev.“: Newiţš liht iţteiktees ţahdai autoritātei weenai ween.

frakturs.lnb.lv

Ziemas talka
vecās drukas labošanai

Ja Tava kaislība ir kļūdu labošana, Latvijas Nacionālā bibliotēka aicina – nāc talkā izlabot 10 000 rindiņu un palīdzī datoram iemācīties lasīt veco druku (fraktūru)!

Lūdzam pieteikties,
rakstot uz adresi dh@lnb.lv



Mašīnmācīšanās LNB

Vecās drukas automātiskās atpazīšanas kvalitātes uzlabošana LNDB, izmantojot mašīnmācīšanās pieeju.

Statistika par veiktajiem labojumiem

Pavisam labotas rindiņas: 30619

Veiksmīgi(saprotami) izlabotas rindiņas: 29294

Rindiņas ar pareizu labojumu un vismaz diviem labotājiem: 10673

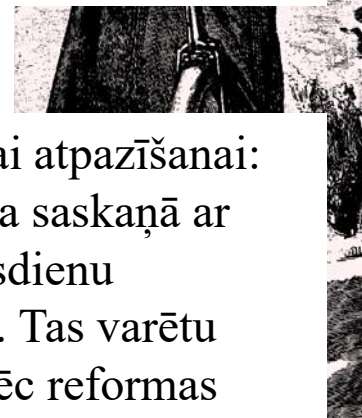
Veiksmīgi sasniegtas 10,000 2x pārbaudītas rindiņas!

Bij Jura deena un jauks, mihliḡs pawaḡara laiks, kurfch pastah-
weja tahds jau no tām deenām, kad šahka šneegs kust. Nemeenu deenu
neatmetās laiks aufstš waj nemihliḡs, bet pat naktiš palika alašch šiltaš
un jaukaš, šaur to nebij lihds peeminetai deenai ne tikween šneegs un
leduš wairš nekur redšams, tā dašchōš šitōš pawaḡarōš kalnainōš apḡa-



Bij Jura deena un jauks, mihliḡs pamasara laiks, kurfch pastah-
roeja tahds jau no tam deenam, kad fahka sneegs kust. Nemeenu deenu
neatmetas laiks aukzts maj nemihliḡs, bet pat naktis palika alasch siltas
un jaukas, zaur ko nebij lihds veeminetai deenai ne tikmeen sneegs un
ledus mairs nekur redsams, ka daschos zitos pamasaros kalnainos apḡa-

Līdz šim atpazītā teksta piemērs
grāmatu portālā: sistemātiski netiek
pareizi atpazīti vairāki burti.



Bij Juḡa diena un jauks, mīliḡs pavasara laiks, kuḡš pastā-
veja tāds jau no tām dienām, kad sāka sniegs kust. Nevienu dienu
neatmetās laiks aukzts vaj nemīliḡs, bet pat naktis palika alaž siltas
un jaukas, caur ko nebij līdz pieminetai dienai ne tikvien sniegs un
ledus vairs nekur redzams, kā dažos citos pavasaros kalnainos apḡa-

Priekšlikums jaunajai atpazīšanai:
teksta modernizēšana saskaņā ar
latviešu valodas mūsdienu
ortogrāfijas reformu. Tas varētu
atvieglot pirms un pēc reformas
iespiestu tekstu automātisku
salīdzināšanu, kā arī atrisinātu s/z
un z/c atveides problēmu.

Par aktuālo īsumā

Aptaujas anketa augstskolu, koledžu un speciālajām bibliotēkām

Kompetenču attīstības centra kursu grafiks – www.lnb.lv

Kompetenču attīstības centra PALDIES!

Pateicība par kvalifikācijas prakses vietām

Prakses vieta	Praktikantu skaits
Latvijas Nacionālā arhīva bibliotēka	1
Latvijas Universitātes bibliotēka	2
Latvijas Valsts mežzinātnes institūta "Silava" bibliotēka	1
Liepājas Universitātes bibliotēka	1
LKA aģentūras Latvijas Kultūras koledžas bibliotēka	1
Rēzeknes Tehnoloģiju akadēmijas bibliotēka	1
RTU Zinātniskā bibliotēka	1
Valsts policijas koledžas bibliotēka	1
Vidzemes Augstskolas bibliotēka	1

Lepojamies - «Saules laiva» 2019

