

Introductie tot de CLARIAH Media Suite

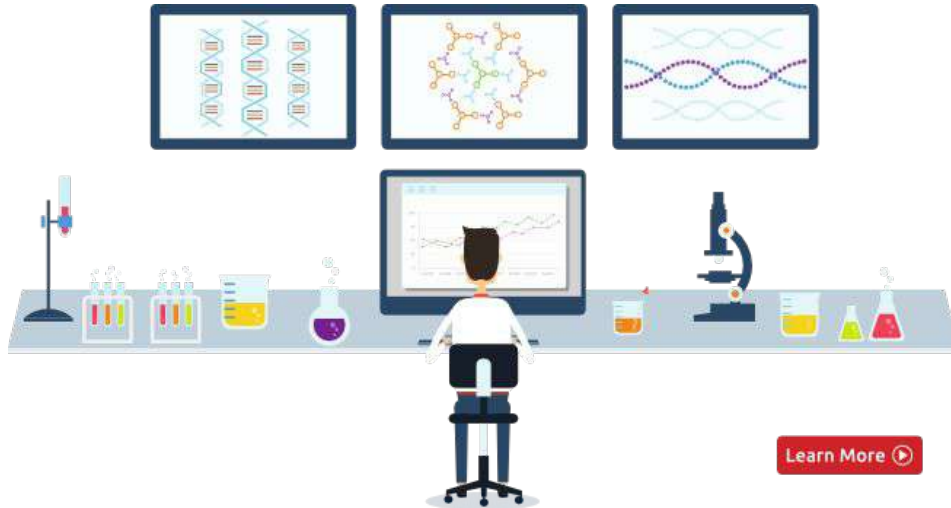


Context: CLARIAH

- Deel van Nationale Roadmap Grootschalige Infrastructuren (NWO)
- Nederlandse bijdrage aan European research infrastructures for the Humanities and Social Sciences (CLARIN en DARIAH)
- CLARIAH-CORE (2014-2018)
- CLARIAH-PLUS (2019-2024)

Wat doet CLARIAH?

Een infrastructuur (data, tools, community) die toegang biedt tot grote collecties en ook de tools aanbiedt om hiermee onderzoek te doen in de kunsten, geesteswetenschappen en sociale wetenschappen



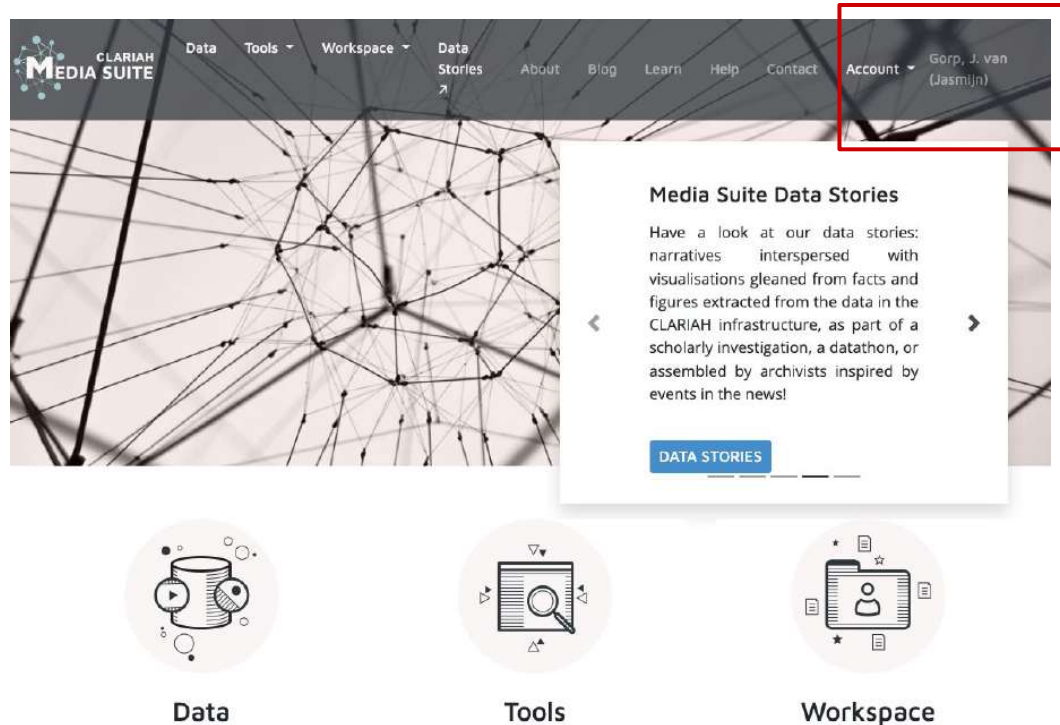
Audiovisuele data als complexe, grootschalige en beschermde data

- Multimodaal
 - Beeld
 - Audio
 - Tekst
 - Montage/narratief
- Grootschalig
 - Zware bestanden
 - Grote hoeveelheden gedigitaliseerd
 - Continue stroom van 'digital born'
- Auteursrechten en Open Beelden



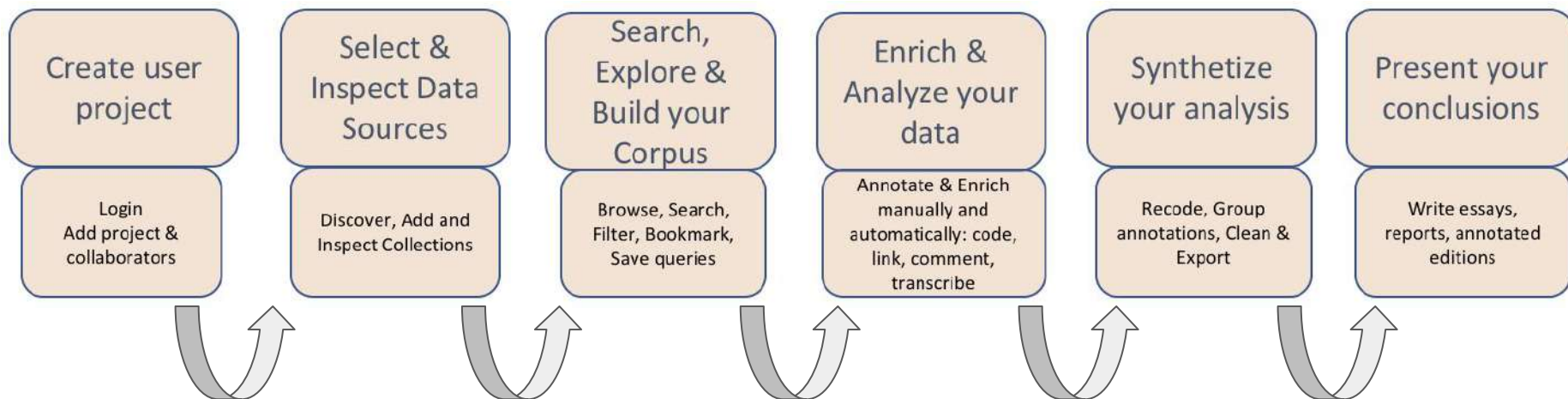
Vandaag De Dag (WNL) via CLARIAH Media Suite

Geïntegreerde en afgesloten onderzoeksomgeving voor wetenschappelijk onderzoek (institutionele login via SURF)



<http://mediasuite.clariah.nl>

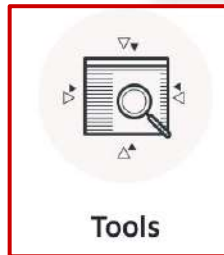
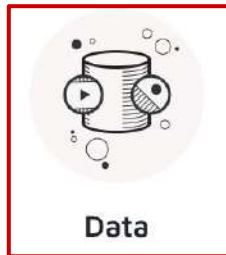
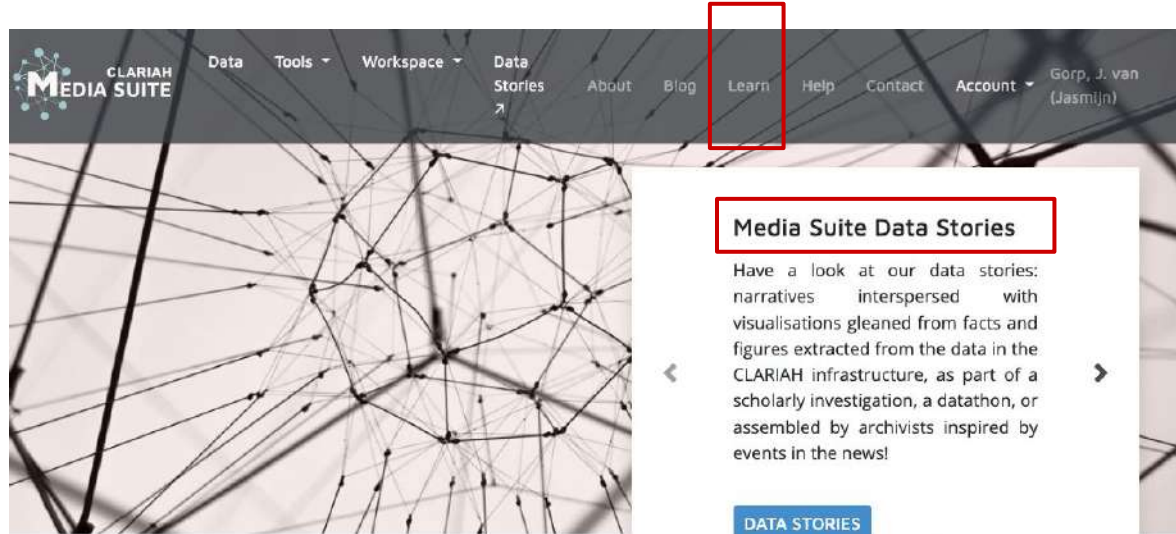
Ondersteuning van het onderzoeksproces



Figuur gebaseerd op:

Melgar Estrada, Liliana, Marijn Koolen, Hugo Huurdeman, en Jaap Blom. "A process model of time-based media annotation in a scholarly context". In *CHIIR 2017: ACM SIGIR Conference on Human Information Interaction and Retrieval*. Oslo, 2017.

Media Suite: Data - Tools - Workspace - Learn - Datastories



CKAN: registry voor de collecties

 **Clariah**

Datasets Organizations Groups About Search

Organizations

What are Organizations?

CKAN Organizations are used to create, manage and publish collections of datasets. Users can have different roles within an Organization, depending on their level of authorisation to create, edit and publish.

62 organizations found

Order by: Name Ascending

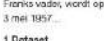
**ANP press agency**
The Algemeen Nederlands Persbureau (ANP) was founded 11 December 1934 by the...

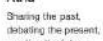
0 Datasets

**Amsterdam Museum**
0 Datasets

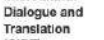
**Ange Wieberdink Productions**
1 Dataset

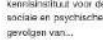
**Calidoscoop Film**
<http://calidoscoopfilm.nl>
1 Dataset

**Anne Frank Stichting**
in nauwe samenwerking met Otto Frank, Anne Frank's vader, wordt op 3 mei 1957...
1 Dataset

**Atria**
Sharing the past, debating the present, creating the future. Onder dat motto...
5 Datasets

**Cultuur & Co.**
1 Dataset

**Center for Intercultural Dialogue and Translation (CIDT)**
Located in the busy suburb of Maasdi in Osnabrück, CIDT is a premier translation...
1 Dataset

**Cogis**
Cogis is het kennisinstituut voor de sociale en psychische gevolgen van...
2 Datasets

**DANS**
Data Archiving and Networked

Dataset Groups Activity Stream

Kijk en Luistercijfers

Data format This collection contains reports on TV and radio ratings in the Netherlands - 'kijk- en luistercijfers'. There are 3347 reports indexed. Each report contains multiple pages of ratings. When you search, you will get reports as search results. If you have searched on a particular word, then when you open the result, that word will be highlighted on the pages. Most of the fields are manually entered. The OCR fields and fields in 'blocks' are automatically generated.

OCR The OCR of the pages is found in the 'pages' field, in the fields 'words' (all text in the page) and 'blocks.words' (block level).

Data and Resources

 **Field Descriptions**
Field descriptions for this collection

Explore

clariah_media_es_in... kijkcijfers luistercijfers viewing figures

Additional Info

Field	Value
Maintainer	J. Wassenaar
Version	1.0
Last Updated	October 25, 2021, 10:48 AM (UTC+02:00)
Created	April 8, 2020, 10:09 AM (UTC+02:00)
es_index	kijkenluistercijfers

“Collection selector” tool

The screenshot displays the CLARIAH MEDIA SUITE interface, specifically the 'Collection selector' tool. The left sidebar contains a search bar, a 'Sound and Vision Arch' section, and a 'Results' section showing 2,113,409 items. The main area is titled 'Choose a collection' and features a dropdown menu with the text '-- Select a collection --'. Below this, there are six collection cards, each with a title, institution, logo, and a 'Read more' link.

CLARIAH MEDIA SUITE Data Tools

Choose a collection [X]

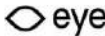
-- Select a collection --

Desmet Film Collection
EYE Filminstituut

[Read more](#)

Desmet Paper Collection
EYE Filminstituut

[Read more](#)

Desmet Poster Collection
EYE Filminstituut

[Read more](#)

Dutch Program Guides (up to 1950)
Nederlands Instituut voor Beeld en Geluid

[Read more](#)

Kijk en Luistercijfers
Nederlands Instituut voor Beeld en Geluid

[Read more](#)

Newspaper Collection
National Library of the Netherlands (KB)
KB nationale bibliotheek
[Read more](#)

Open Images - EYE
EYE Filminstituut


Open Images - Natuurbeelden
Nederlands Instituut voor Beeld en Geluid


Open Images - Sound and Vision
Nederlands Instituut voor Beeld en Geluid


“Search” Tool

 Search

Interstitial Curation 

 Sound and Vision Archive  10.000 

 in  

Lock facet selections ☐

  Max keyword values to display:

Results 3.018

[+ ADD A NEW FACET](#)

 mediaType

Search facet ... 3 terms

Empty field 0

 audio 1949

 video 1069

 Broadcaster

Search facet ... 62 terms

Empty field 170

 akn 4

         NEXT LAST 

 NOVA

als uit de hand gelopen hobby; 3. kernramp **Tsjernoby!** tien jaar geleden; 4. **Tsjernoby!** (3): de hulptroepen; 5. * fragment proefaflevering van een drama-serie over het Nederlandse leger (19" vanaf TC 23.05.14).

Description - Scene Description: KERNRAMP IN **TSJERNOBYL** TIEN JAAR GELEDEN Reportage, ivm de 10-jarige herdenking, over de gevolgen van de ramp in de kerncentrale van **Tsjernoby!**, oprichtster kinderfonds **Tsjernoby!**, over haar zoon bij wie wederom een kankergezwell is geconstateerd; - S. Parashin, directeur kerncentrale **Tsjernoby!**, over, jongeren met lymfeklierkanker; - div. kindertekeningen over de **Tsjernoby!**-ramp; - bewaakt terrein waar materiaal staat dat destijds is gebruikt (alles);

Topic: geographical place - Scene description: **Tsjernoby!**

Summary: bottenverzamelen als uit de hand gelopen hobby; 3. kernramp **Tsjernoby!** tien jaar geleden; 4. **Tsjernoby!** (3): de hulptroepen; 5. * fragment proefaflevering van een drama-serie over het Nederlandse leger (19" vanaf TC 23:05:14).

 1996-04-26 5 matches in archival metadata



“Resource Viewer” voor Annotatie, Segmentering & Speech Transcripts

Resource Viewer

Interstitial Curation

EENVANDAAG

Back to results Bookmark

Metadata Content Annotations

Type: Speech transcript

View: List tsjernoby

Showing 10 of 98 lines

00:00:11 Het kwam libië ineens uit de hoge hoed met een staatskreet vuren japan verhoogt het dreigingsniveau voor een kernramp nadat de straling weer gestegen is rond de kerncentrale van fukushima en het congres vijftienhonderd jaar na tsjernoby in nederland was bedoeld om terug te kijken.

00:11:12 Congres dat vandaag de nasleep van tsjernoby moesten afsluiten maar waar de congresgangers uiteraard over niets anders konden praten dan over fukushima is gaan we naar het sportjournaal vooruutblikken want het is na eenvandaag zoals u weet henry schut wat zit daar wel in vandaag.

00:11:56 Vandaag kwamen ze bij elkaar om te praten over de eerste kernramp ooit tot nog toe tsjernoby.

00:12:01 Nederlandse bestuurders wetenschappers en ambtenaren die vijftienhonderd jaar geleden aan het roer stonden toen een kernreactor ontplofte in het oekraïense tsjernoby waarvan het congres vandaag kwam welig terecht vandaag ging het uiteraard alleen nog maar over het zwartste scenario over angst die opnieuw werkelijkheid worden kortom over het drama de japanse kerncentrale fukushima.

00:13:17 Iemand met recht van spreken ten tijde van de ramp in tsjernoby is heideeman die nederland moet beschermen tegen de gevolgen van de straling als geen ander weet hij hoe moeilijk het is om aan de juiste informatie te komen om zo goed mogelijk te kunnen reageren.

00:13:50 Ten tweede moet je dus de hele

Search in layers...

+ ADD LAYER

Segments

Speech transcript

Speech transcript Words

Program segments: Libië kondigt wapensollicitatie af

kijken vanavond uitzending verrassende zet kolonel gadafi novo straatjagers scherp stonden no fly zone gaan afwijken kwam libië ineens hoge libische japan eenvandaag hele goede avond truc sluwe woestijn van kolonel khadafi werkelijk plan zon bloedvergieten eind burgersorog land vanmiddag

Workspace

[Home](#) / [Workspace](#) / [Projects](#) / [Interstitial Curation](#) / [Queries](#)



Interstitial Curation

Details

Bookmarks 11

Annotations 1

Queries 16

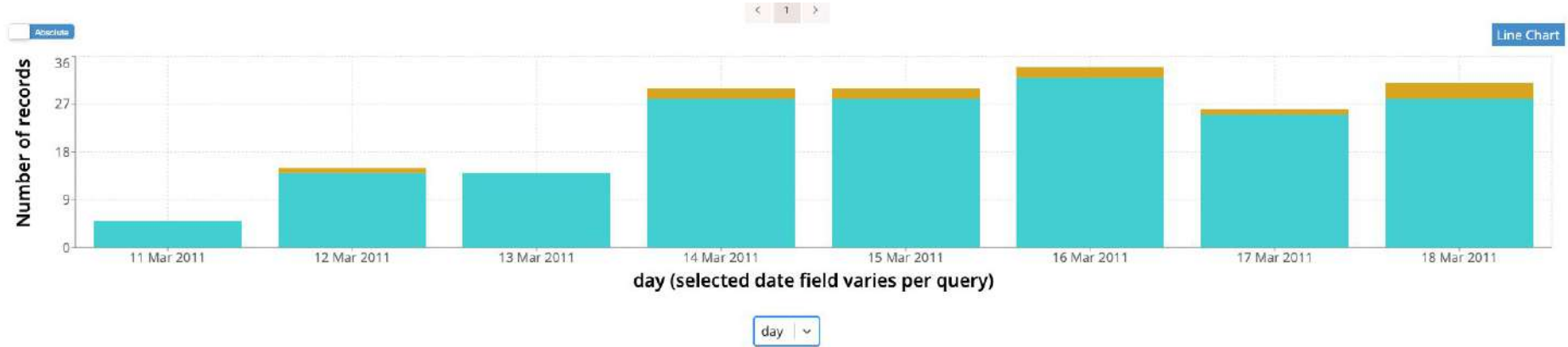
Filters

Search



☐	Name	Query		
☐	All commercials	No search term; # filters: 2	DELETE	OPEN >
☐	All leaders	Search term: Leaders; # filters: 1	DELETE	OPEN >
☐	All Leader_	Search term: Leader; # filters: 1	DELETE	OPEN >
☐	ALL Postbus 51	Search term: "postbus 51"; # filters: 2	DELETE	OPEN >
☐	ALL Postbus 51	Search term: "postbus 51"; # filters: 2	DELETE	OPEN >
☐	announcements in compilations	Search term: aankondiging; # filters: 2	DELETE	OPEN >
☐	Dienstmededeling	Search term: dienstmededeling; # filters: 1	DELETE	OPEN >

“Compare” tool



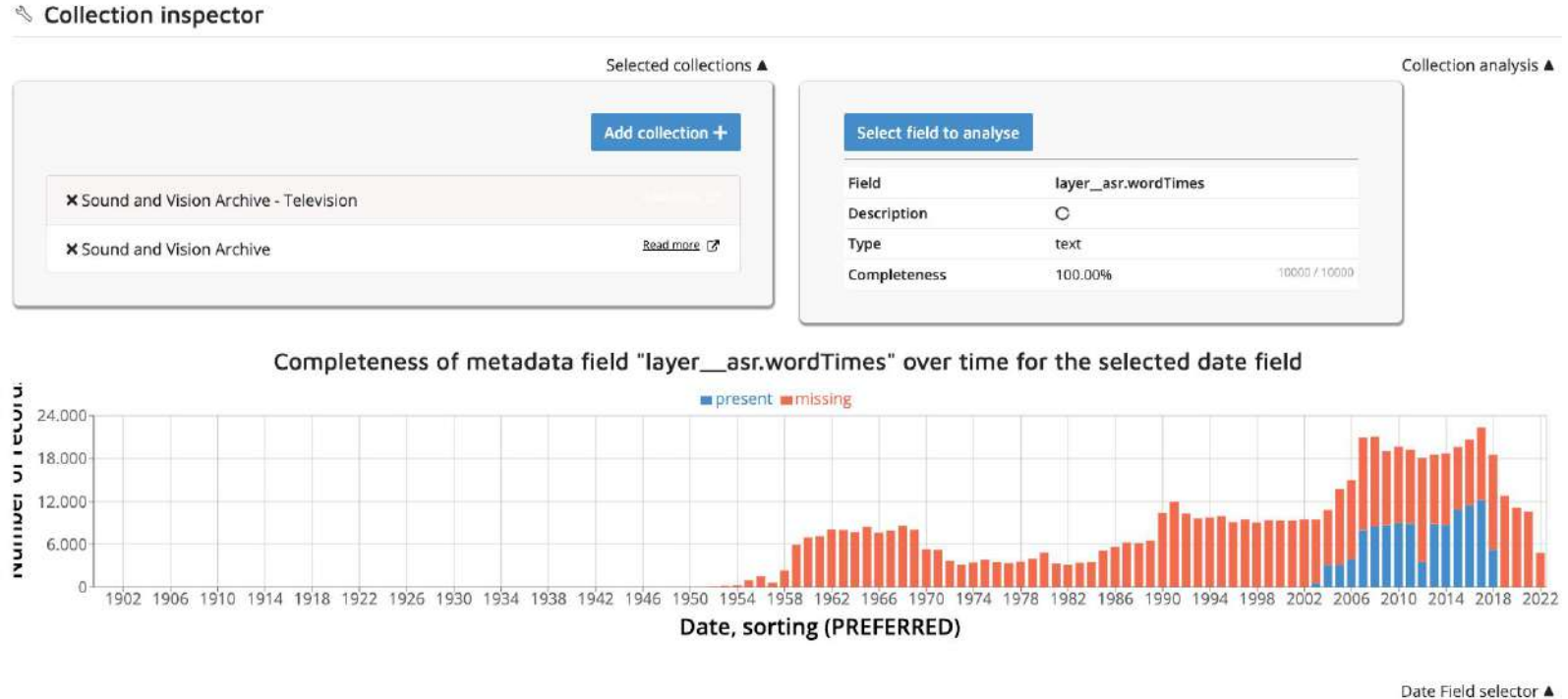
Query #1: Fukushima

COLLECTION: Sound And Vision Archive
QUERY TERM: Fukushima
DATE RANGE:

Query #2: Tsjernobyl

COLLECTION: Sound And Vision Archive
QUERY TERM: Tsjernobyl
DATE RANGE:

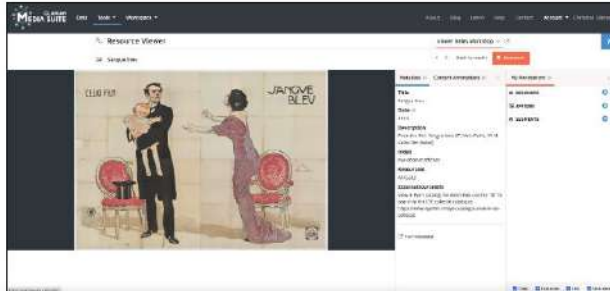
“Collection Inspector” Tool



Learn “tutorials”



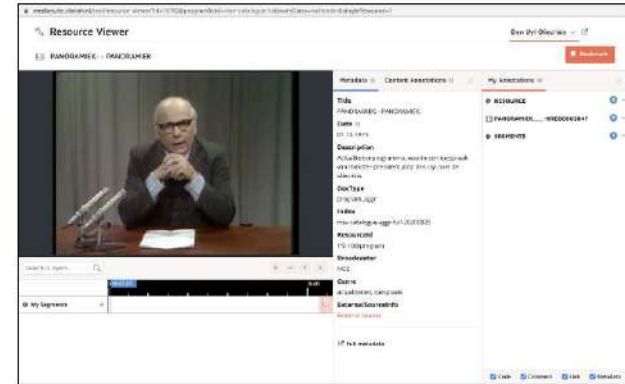
Tutorial: Visual Analysis and Historical Source Criticism of Desmet's Film Posters and Business Documents



Tutorial for CLARIAH Media Suite V 5.3, May 2021

Tutorial Description, Case and Objectives

TUTORIAL: VIEWING RATES FOR TELEVISION HISTORY



Tutorial for CLARIAH Media Suite V 5.3

Tutorial Description, Case and Objectives

This tutorial introduces you to using the CLARIAH Media Suite's search tool and bookmarking functionality for research with the viewing and listening rates collection. The tutorial specifically introduces you to ways of using the Media Suite that support television history teaching and research that aim at critically evaluating processes of viewing rate research and data driven

Learn “curated playlists”

Curated Playlist: Finding Interstitials in the Television Archive

Jasmijn Van Gorp, Utrecht University

Introduction

Interstitials, be it promos, commercials or trailers, are literal bits between two programmes, making up a true treasure trove for television historical research. As Ellis (2011) puts it, interstitials function as ‘user guides’ in how broadcasters envisioned TV consumption by viewers in a given moment. Johnson (2012) illustrates the latter in her own research on branding UK and US television and points out how channels branded themselves, in addition to raising revenues through commercials and providing continuity techniques through interstitials. In this curated playlist, I show that interstitials not only are a treasure trove to shed light on scheduling and branding: they also prove useful for archive studies. I will elaborate this point by showing how researching interstitials can play a role in the understanding of the workings of broadcast archives and its metadata.

Interstitials can be seen as the epitome of what Polan (2013: 347) calls ‘forgettable’ television: “programming designed to be forgotten at virtually the very moment of its original viewing”, therefore rendering them vastly under-researched. Traditionally television historical accounts and broadcast archives alike used to be focused on the programmes rather than the bits between the programmes. Consequently the majority of interstitials were never saved or made accessible for research. Yet in the (rare) case of saved interstitials, Johnson (2012) noticed large numbers of

Category 4: Weekly Recordings

Johnson (2009) used Weekly recordings for her research. Weekly recordings are full day recordings of the entire 24h broadcast-cycle of one channel. NISV also stores these kinds of Weekly recordings of two weeks per year, for each channel from 1987. The best way to reach this subcollection is to conduct a search query for “Integral Recordings” and “Week” which results in about 3032 recordings, of which most can not be played in the Media Suite. The findings also contain recordings of the commercial broadcasters.

For my curated playlist I select a weekend-day (a Sunday), along with the first broadcaster, Nederland 1: NED 1: WEEK 40-1999 1.10.1999 . I watched the first reel and noted all interstitials and their categorization type (see overview, table 1). This shows that during a time period of two hours almost 44 interstitials were shown, which were all left out of the title of the record, the description and of the programme schedule in the newspaper. This alone proves my point earlier made that interstitials are -not only academically but also quite literally - forgettable television.



Data Stories: volgende presentatie

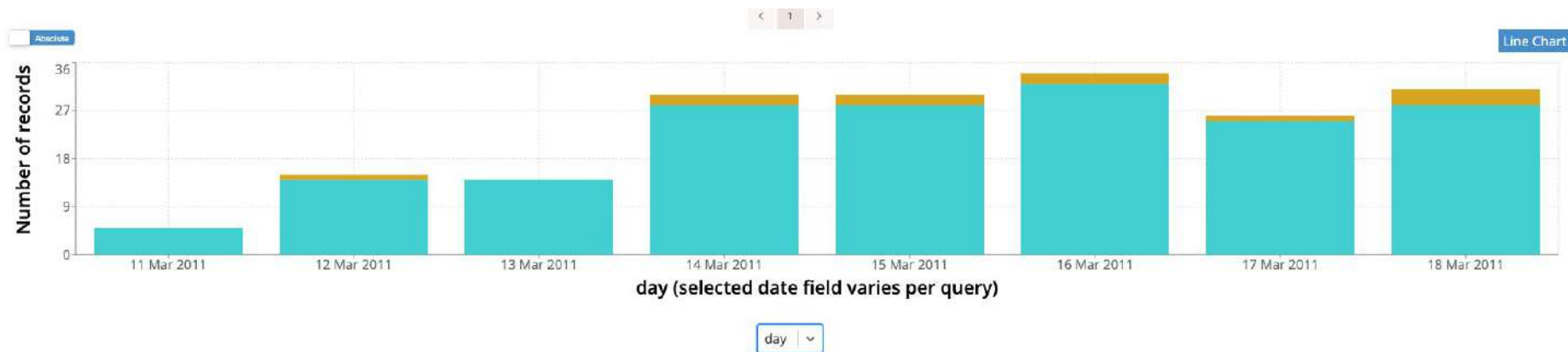
WETENSCHAP ALS BRON VAN VERHALEN

23rd June 2022

Mari Wigham, Willemien Sanders



DATA-GEDREVEN



Query #1: Fukushima

COLLECTION: Sound And Vision Archive
 QUERY TERM: Fukushima
 DATE RANGE:

Query #2: Tsjernobyl

COLLECTION: Sound And Vision Archive
 QUERY TERM: Tsjernobyl
 DATE RANGE:

WAT IS EEN DATA STORY?

- Onderzoeksvraag centraal
- Data-gedreven
- Anekdoten ter illustratie
- Context en interpretatie
- Visueel

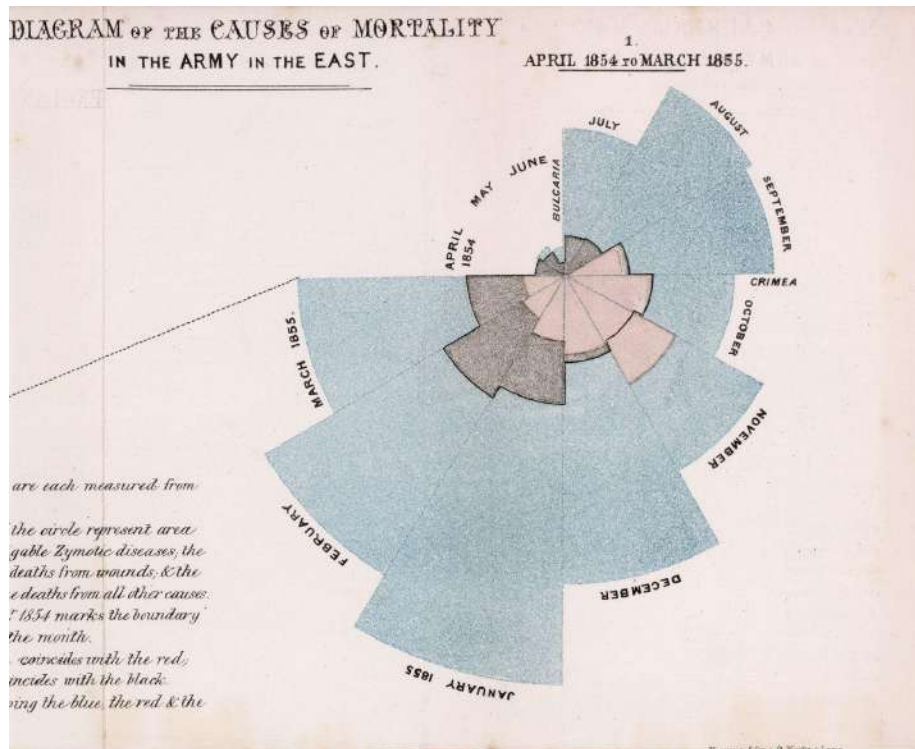
WAAROM DATA?

- Grote schaal
- Nieuwe bron
 - kan bevestigen wat we al vermoedden - of het juist ontkrachten
- Onderliggende patronen ontdekken
- De onverwachte vinden



DATA(-VIZ) SPREEKT BOEKDELEN

“In 2022, **29%** of corporate board seats at North American and European companies were held by women” - CNBC



WAAROM EEN 'STORY'?

- Verhalen vertellen zit mensen in het bloed
- Wetenschappelijke publicatie is ook een 'verhaal'
- Context bepaalt het verhaal
- Interpretatie van data is essentieel





VOORBEELD: VERKIEZINGEN

1213

UITZENDINGEN MET TRANSCRIPTEN

58

PROGRAMMA'S

1742

POLITICI
MET GESLACHT EN PARTIJ

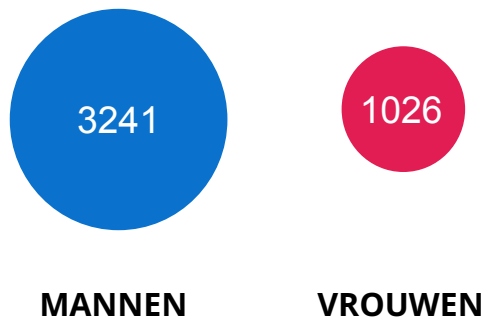
220

POLITICI
MET GEZICHT/STEMMODEL



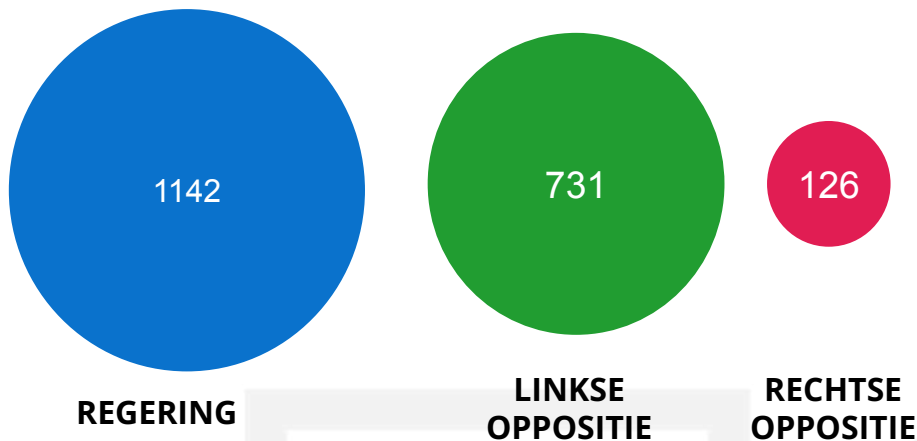
RESULTATEN

Mannen werden vaker
genoemd dan vrouwen



AANTAL KEER GENOEMD

Politici van de regering spraken het
meeste



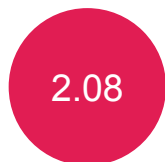
TOTALE SPREEKTijd (MIN)

The chart is a complex bubble chart visualization of the Dutch political landscape in 2017. The chart consists of numerous overlapping circles of various colors and sizes, each representing a political party or coalition. The size of each circle corresponds to the party's share of the vote. The colors are grouped by political affiliation: purple for the center-right (Lilianne Ploumen, Sigrid Kaag), green for the center-left (Lilianne Ploumen, Farid Azarkan), red for the right-wing (Mark Rutte), orange for the left-wing (Gert-Jan Segers), and blue for the far-left (Geert Wilders). The chart also includes smaller circles for various regional and local parties, as well as a small circle for the 'Other' category. The overall layout is a dense cluster of circles, with the largest circles (Mark Rutte, Gert-Jan Segers, Lilianne Ploumen, Sigrid Kaag) dominating the center. The chart is set against a white background with a faint grid.

Vrouwen spraken gemiddeld
langer dan mannen



MANNEN



VROUWEN

GEMIDDELDE SPREEKTijd (MIN)

Rechtse politici werden het vaakst
genoemd



LINKS



CENTRUM



RECHTS

AANTAL KEER GENOEMD



NPO brengt Wopke Hoekstra het meest in beeld tijdens verkiezingscampagne

16 maart 2021 08:11

41 NUjj-reacties v



NOS

Nieuws v

Sport v

Uitzendingen

NOS NIEUWS • BINNENLAND • POLITIEK • 16-03-2021, 06:00

Onderzoek: Rechts domineert de verkiezingscampagne op radio en tv



Joop

Nieuws — 16 maart 2021 • Laatste update 13:50

Q 142 <

Publieke omroep heeft uitgesproken voorkeur voor rechtse mannen



ONZE DATA STORIES

- *15 Jaar DWDD* - analyse van De Wereld Draait Door
- *Factual divergences in Dutch television news* -
misinformatie rondom corona
- *Ladies and gentlemen...* - onderzoek van gender in media
- *Who's speaking* - verkiezingsverhaal

Komende attracties:

- *20 jaar Tegenlicht* - analyse van Tegenlicht
- *Mediaoorlog* - onderzoek in oorlogskranten

DE INSTINKERS

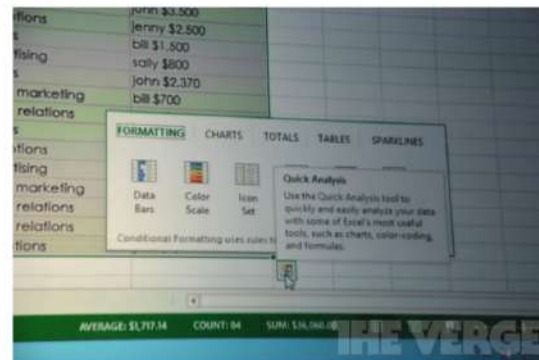
- Eigenaardigheden in de (meta)data, bijv.:
 - ontbrekende (meta)data
 - heterogeniteit
- Problemen in de analyse, bijv.:
 - fouten
 - verkeerd gebruik van analyses
- Fouten in interpretatie
- Fouten in communicatie

A simple Excel calculation error puts a famous economic study under scrutiny

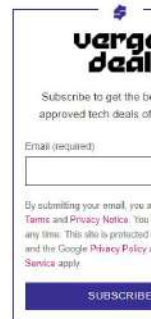
By Nathan Ingraham | Apr 17, 2013, 11:06am EDT

Source: Financial Times, University of Massachusetts, and Next New Deal | Via Business Insider and Bloomberg Businessweek

f t e SHARE



Excel 2013 stock



Een fout antwoord ziet er vaak hetzelfde uit als een juist antwoord...

HET TEAM

Multidisciplinair

- Inhoudelijke expert
- Archief expert
- Data expert
- Eindredacteur



Photo by Vlad Hilitanu on Unsplash

HET PROCES

Zorgvuldigheid is nodig - je wilt geen onzin verkopen!

- Exploratiefase - data exploreren en de vraag formuleren
- Onderzoeksfase - data analyseren en de vraag beantwoorden in een verhaal
- Reviewfase - alles goed controleren
- Publicatiefase - publiceren, monitoren, reageren

HET ARCHIEF CENTRAAL!

Uitdaging/kans

Aandacht

Kritiek



HET ARCHIEF CENTRAAL!

Archief expert

- Metadata goed documenteren (metadata archeologie, provenance)
- Ondersteuning bieden aan onderzoekers
- Uitdaging/kans: Metadata verbeteren

Data expert

- Data digitaal beschikbaar maken
- Analysemethodes beschikbaar maken



By Silar - Own work, CC BY-SA 4.0, <https://commons.wikimedia.org/w/index.php?curid=106541058>

BEEELD EN GELUID: STATUS

- Data beschikbaar gemaakt
- Proces ontwikkeld
- Verhalen gepubliceerd

Toekomst:

- Metadata documenteren
- Tools beschikbaar maken
- Onderzoekers ondersteunen



JIJ OOK?

- Zin om ook mee te doen aan data stories?
- Vragen of opmerkingen?