

Workshop med DH-Lab-apper fra Nasjonalbiblioteket

2025-02-13

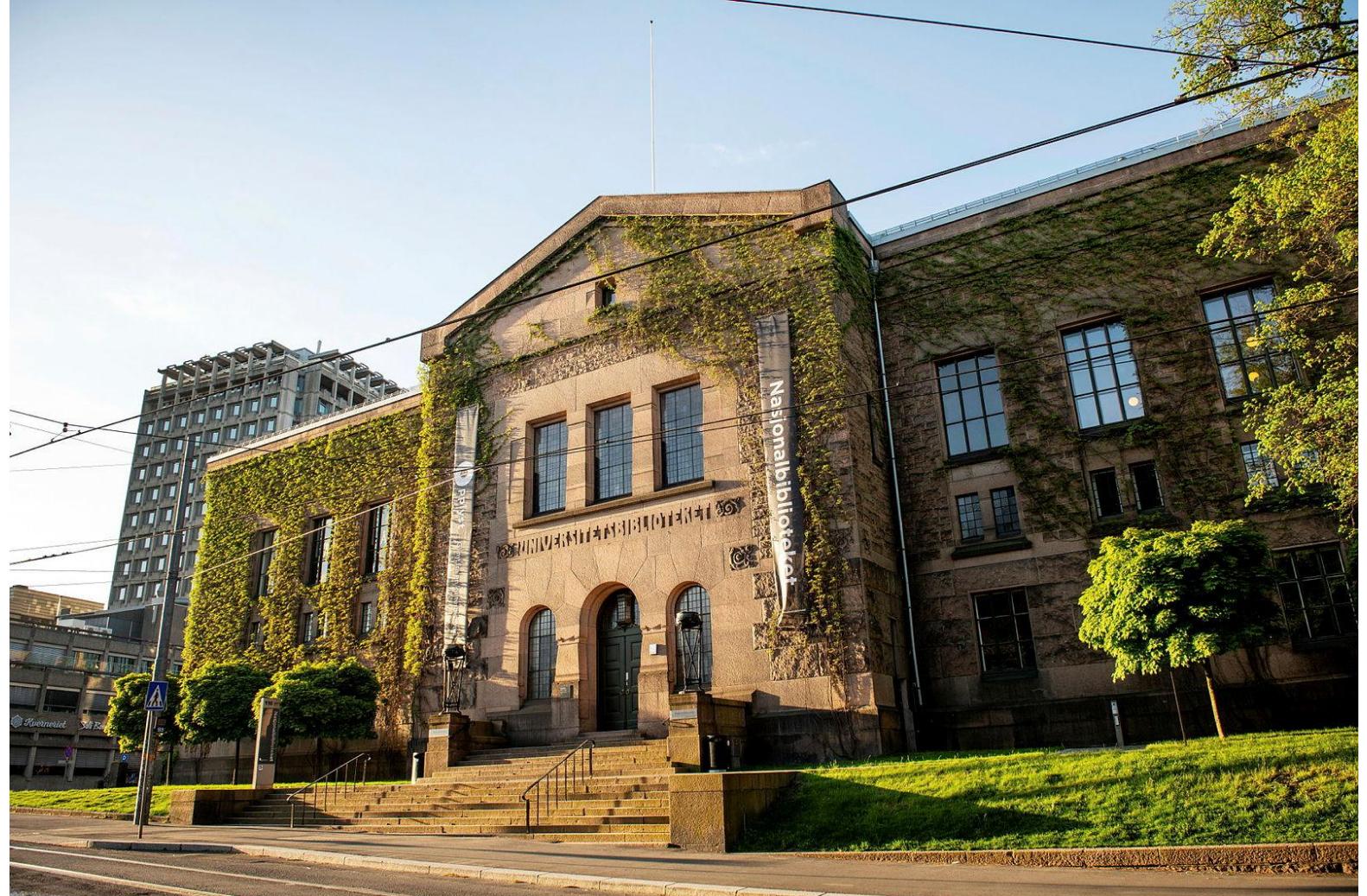
Tita Enstad

Marie Roald

Thea Tollersrud

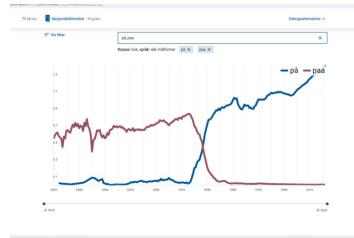


Nasjonalbiblioteket

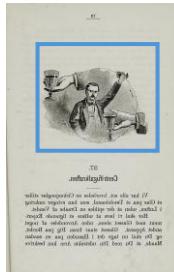




Om NB og DH-Laben



N-gram og dagsplot

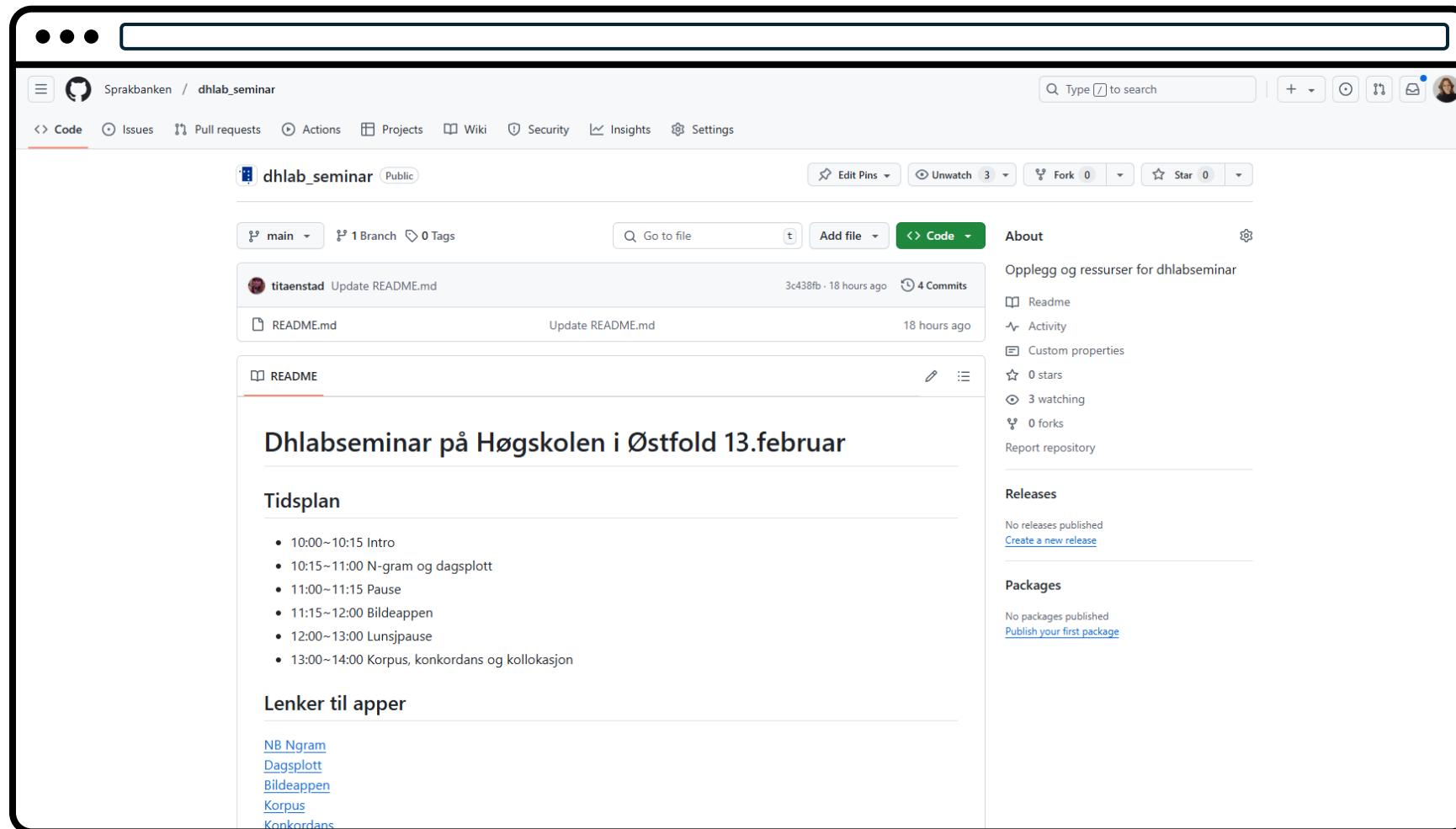


Søk etter bilder i bøker

Norges Militære
sentralbyrå Grunnskole
Uppgitt Teknisk Univ Kun
Tekniske NTH
Bjørknes
TN grunnskole
utd universitet
Videregående gående
Høyskole Universitet
Forsvarets

Korpus, konkordans og kollokasjon

Materialet fra seminaret ligger tilgjengelig på GitHub:
github.com/Sprakbanken/dhlab_seminar



Nasjonalbibliotekets samling spenner hundrevis av år



PHOTO: Statsbygg/Hans Fredrik Asbjørnsen

I 2006 startet Nasjonalbiblioteket en prosess for å digitalisere samlingen



Målet med digitaliseringsprosessen er både å bevare og tilgjengeliggjøre kulturarven



[...] bidra til at den delen av kulturarven som Nasjonalbiblioteket har som oppgave å bevare, når frem til flest mulig. [...]

Digitaliseringsprosessen har kommet langt på vei og praktisk talt alle bøker er skannet

Materialetype	Andel	Størrelse
Video og fjernsyn	22 %	2900 TB
Film	21 %	2700 TB
Aviser	19 %	2500 TB
Nettarkiv	16 %	2100 TB
Radio og lyd	12 %	1600 TB
Bøker	8 %	1000 TB
Foto	2 %	260 TB

Bokhyllaavtalen sikrer alle med norsk IP-adresse digital tilgang til alle bøker publisert før 2006

Nasjonalbiblioteket Nettbiblioteket Historier fra samlingen Hva skjer? Åpningstider Forskning English ☰ Menu

Nasjonalbiblioteket > Pressemeldinger > Ny Bokhylla-avtale gir deg digital tilgang til 400 000 bøker

Ny Bokhylla-avtale gir deg digital tilgang til 400 000 bøker



Kultur- og likestillingsminister Lubna Jaffery, nasjonalbibliotekar Aslak Sira Myhre og Kopinors administrerende direktør, Hege Munch Gundersen, la fram en ny Bokhylla-avtale fredag 13. oktober på Nasjonalbiblioteket. Foto: Gorm K. Gaare/Nasjonalbiblioteket

Snart vil du få digital tilgang til alle bøker som er utgitt i Norge til og med 2005 i Nettbiblioteket på nb.no. Den nye Bokhylla-avtalen, som er inngått mellom Nasjonalbiblioteket og Kopinor, omfatter 400 000 rettighetsbelagte norske bøker. Dette er 175 000 flere enn i dag.

Nettbiblioteket er en tjeneste som tilgjengeliggør den digitaliserte biten av Nasjonalbibliotekets samling



Nettbiblioteket er en tjeneste som tilgjengeliggør den digitaliserte biten av Nasjonalbibliotekets samling

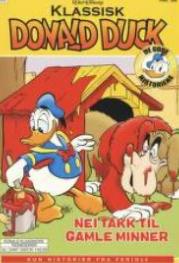


<https://www.nb.no/search>

I Nettbiblioteket kan du søke om digital tilgang til nyere bøker, aviser og magasiner for forskningsformål

[← Tilbake til trefflisten](#) [← Forrige treff | Neste treff →](#)

Klassisk Donald Duck. 2022



Klassisk Donald Duck. 20...
Egmont Kids Media
2022

Tilgjengelig etter bestemte vilkår

Tilgang
Tilgjengelig etter bestemte vilkår ([Les om tilgangsvilkår](#)).
[Logg inn](#) for å søke om tilgang.
Vitenskapelig ansatte i UH-sektoren får [direkte tilgang ved innlogging med Feide](#).

Tittel
Klassisk Donald Duck. 2022

Serietittel
Klassisk Donald Duck

Alternativ tittel
Walt Disney's klassisk Donald Duck

I Nettbiblioteket kan du søke om digital tilgang til nyere bøker, aviser og magasiner for forskningsformål

← Tilbake til trefflisten ← Forrige treff | Neste treff →

Klassisk Donald Duck. 2022

A small thumbnail image of a Donald Duck comic book cover.

A small logo for 'GAMLE MINNER' with the text 'KJEN HISTORIER FOR VEDVOLGEN' below it.

Tilgjengelig etter bestemte vilkår

Tilgang
Tilgjengelig etter bestemte vilkår ([Les om tilgangsvilkår](#)).
[Logg inn for å søke om tilgang](#)

Vitenskapelig ansatte i UH-sektoren får [direkte tilgang ved innlogging med Feide](#).

Tittel
Klassisk Donald Duck. 2022

Serietittel
Klassisk Donald Duck

Alternativ tittel
Walt Disney's klassisk Donald Duck

Seksjonen for Språkbank og DH-Lab fasiliterer forskning og utvikling basert på NBs samling



Nasjonalbiblioteket Nasjonalbiblioteket Nettbiblioteket Historier fra samlingen Hva skjer? Åpningstider Forskning English Meny

Nasjonalbiblioteket > DH-LAB > Apper fra DH-LAB

Apper fra DH-LAB

DH-LAB Apper fra DH-LAB Forskningsprosjekter Kontakt

Appene bruker [vårt API](#) og er [tilgjengelige som åpen kildekode](#).

NB N-gram

NB N-gram er en søketjeneste som gir deg muligheten til å finne og sammenligne ordfrekvenser, dvs. når og hvor ofte ord forekommer i et historisk perspektiv.

Bildesøk

Søk etter bilder i Nasjonalbibliotekets bøker.

Lignende bilder

Lærerib / naturhistorie for folkeskolen Side 70 Larsson, P. A. 1893 Blomkvist, Olof Den første bog for barnet Side 35 Vellesen, Jonas

Naturkundebog - Nr. 11 Større udgave Om dyrene, planterne og steinerne Side 84 Beer, Johannes

Læren om mennesket og dyrene : til skolebrug Side 62 Beer, Johannes

Norsk ABC-alder Lærebok sat omasbørn Side 24 Vellesen, Jonas

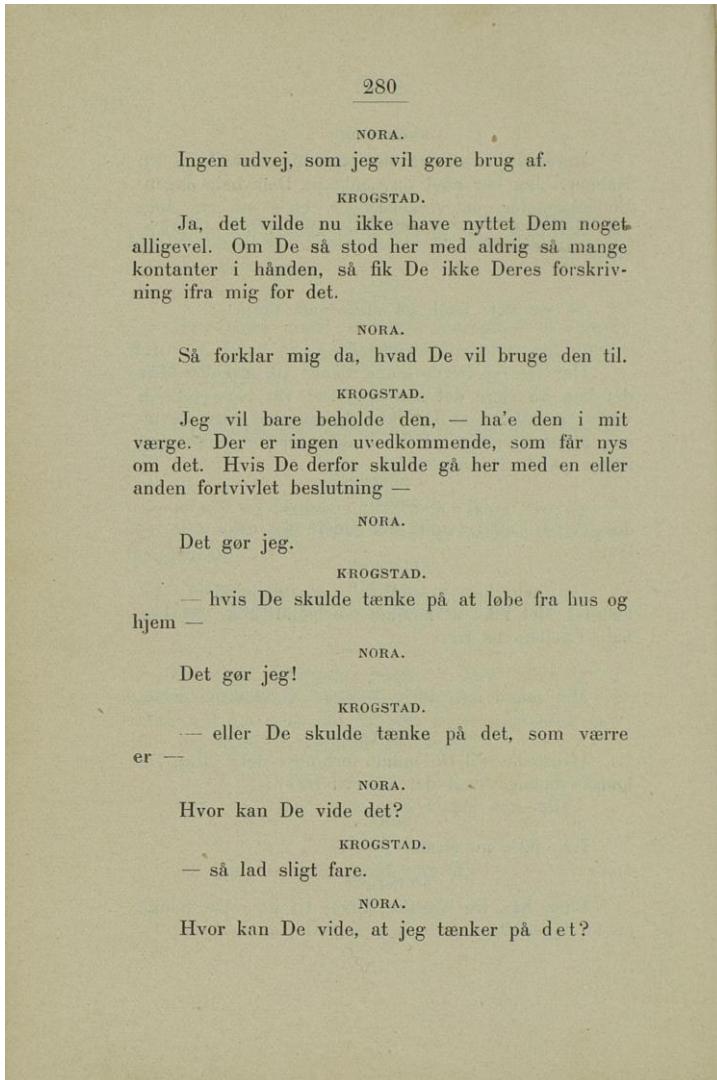
Søk i 1800-tallets bilder

Søk i bilder i bøker fra 1800-tallet, både med tekst og oppplastet bilde.

title

/ Zern , Leif / Bredal , Bjørn / Grieg , Edvard / Festspillbo

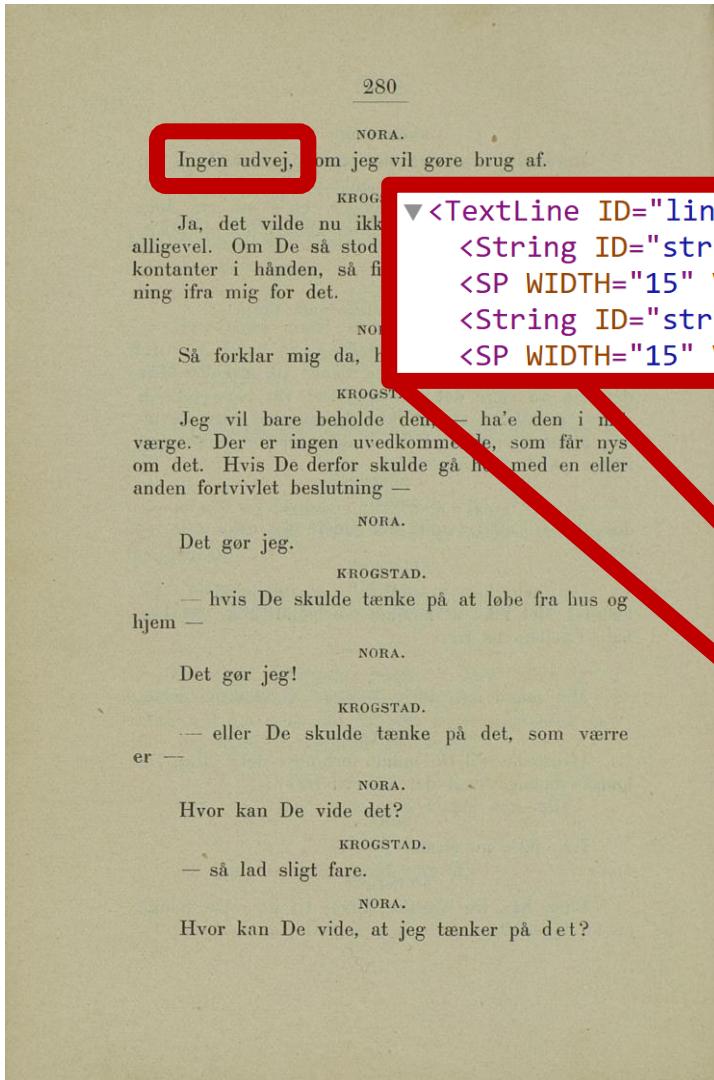
Nasjonalbiblioteket har tekstinnholdet til dokumenter lagret som ALTO-XML filer



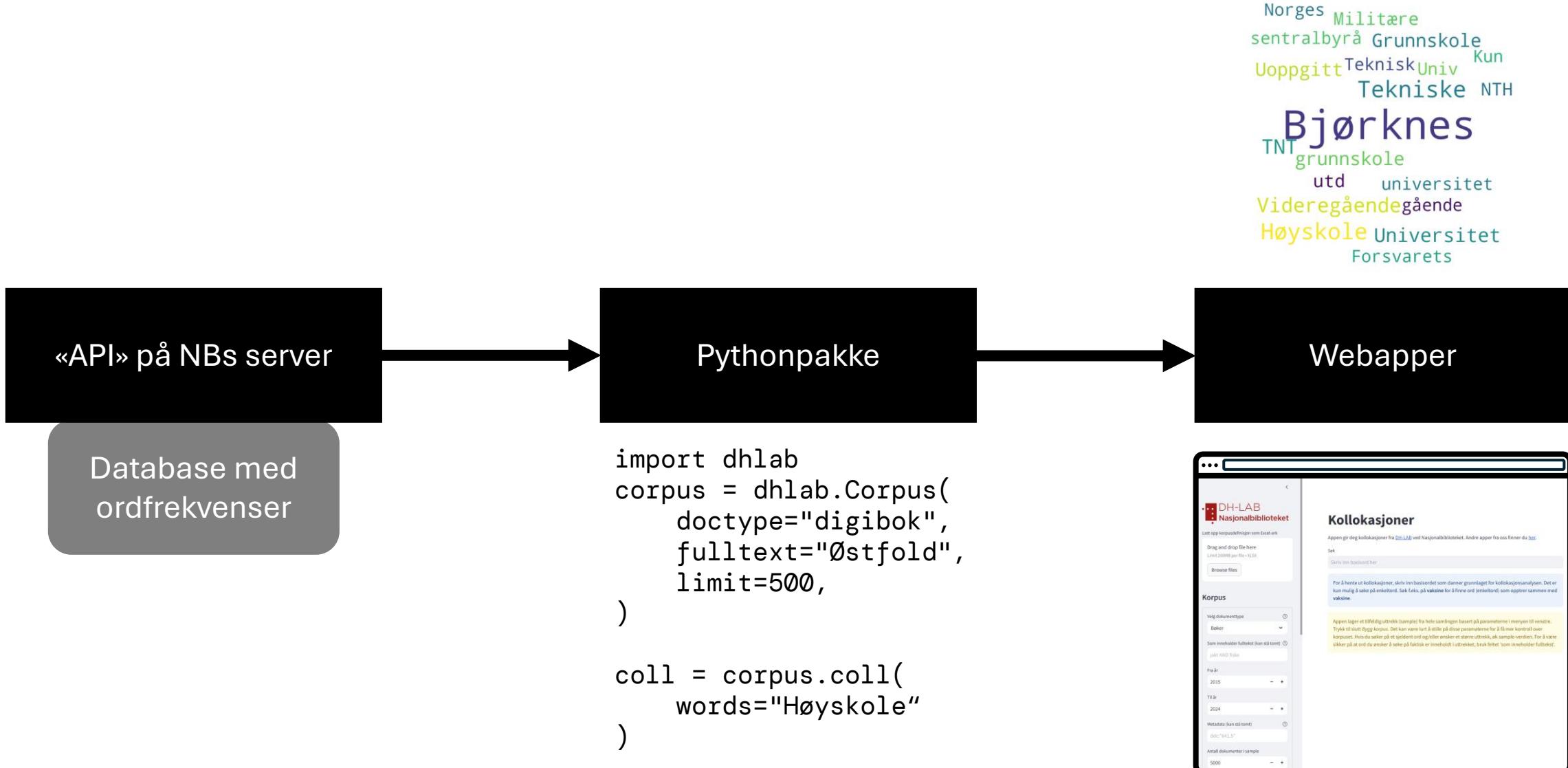
This XML file does not appear to have any style information associated with it. The document tree is shown below.

```
<alto xmlns="http://www.loc.gov/standards/alto/ns-v3#" xmlns:xlink="http://www.w3.org/1999/xlink" xmlns:xsi="http://www.w3.org/2001/XMLSchema-instance" xsi:schemaLocation="https://www.loc.gov/standards/alto/ns-v3.xsd">
  <Description>
    <MeasurementUnit>mm10</MeasurementUnit>
    <SourceImageInformation>
      <fileName>https://www.nb.no/services/image/resolver/URN:NBN:no-nb_digibok_2009040303005_0324/full/pct:50/0/native.jpg</fileName>
      <sourceImageInformation></sourceImageInformation>
    </SourceImageInformation>
    <OCRProcessing ID="OCR_0">
      <ocrProcessingStep>
        <processingSoftware>tesseract 5.2.0</processingSoftware>
        </processingSoftware>
      </ocrProcessingStep>
    </OCRProcessing>
  </Description>
  <Styles>
    <TextStyle ID="TXT_0" FONTSIZE="32" FONTFAMILY="Times Roman"/>
    <ParagraphStyle ID="PAR_BLOCK" ALIGN="Block"/>
  </Styles>
  <Layout>
    <Page WIDTH="1192" HEIGHT="1790" PHYSICAL_IMG_NR="0" ID="page_0">
      <PrintSpace HPOS="0" VPOS="0" WIDTH="1192" HEIGHT="1790">
        <ComposedBlock ID="cblock_0" HPOS="210" VPOS="208" WIDTH="824" HEIGHT="394">
          <TextBlock ID="block_0" HPOS="582" VPOS="208" WIDTH="230" HEIGHT="20" STYLEREFS="TXT_0 PAR_BLOCK">
            <TextLine ID="line_0" HPOS="582" VPOS="208" WIDTH="230" HEIGHT="20">
              <String ID="string_0" HPOS="582" VPOS="208" WIDTH="164" HEIGHT="16" WC="0.75" CONTENT="NoRA." STYLEREFS="TXT_0"/>
              <SP WIDTH="32" VPOS="208" HPOS="746"/>
              <String ID="string_1" HPOS="779" VPOS="216" WIDTH="32" HEIGHT="10" WC="0.10" CONTENT=")" STYLEREFS="TXT_0"/>
            </TextLine>
          </TextBlock>
          <TextBlock ID="block_1" HPOS="288" VPOS="242" WIDTH="606" HEIGHT="34" STYLEREFS="TXT_0 PAR_BLOCK">
            <TextLine ID="line_1" HPOS="288" VPOS="242" WIDTH="606" HEIGHT="34">
              <String ID="string_2" HPOS="288" VPOS="242" WIDTH="79" HEIGHT="34" WC="0.92" CONTENT="Ingen" STYLEREFS="TXT_0"/>
              <SP WIDTH="15" VPOS="242" HPOS="367"/>
              <String ID="string_3" HPOS="383" VPOS="244" WIDTH="95" HEIGHT="34" WC="0.92" CONTENT="udvej," STYLEREFS="TXT_0"/>
              <SP WIDTH="15" VPOS="244" HPOS="479"/>
              <String ID="string_4" HPOS="495" VPOS="254" WIDTH="47" HEIGHT="16" WC="0.92" CONTENT="som" STYLEREFS="TXT_0"/>
              <SP WIDTH="15" VPOS="254" HPOS="543"/>
              <String ID="string_5" HPOS="559" VPOS="242" WIDTH="47" HEIGHT="34" WC="0.92" CONTENT="jeg" STYLEREFS="TXT_0"/>
              <SP WIDTH="15" VPOS="242" HPOS="606"/>
              <String ID="string_6" HPOS="622" VPOS="244" WIDTH="47" HEIGHT="24" WC="0.92" CONTENT="vil" STYLEREFS="TXT_0"/>
              <SP WIDTH="15" VPOS="244" HPOS="670"/>
              <String ID="string_7" HPOS="686" VPOS="254" WIDTH="63" HEIGHT="24" WC="0.92" CONTENT="gøre" STYLEREFS="TXT_0"/>
              <SP WIDTH="15" VPOS="254" HPOS="750"/>
              <String ID="string_8" HPOS="766" VPOS="242" WIDTH="63" HEIGHT="34" WC="0.92" CONTENT="brug" STYLEREFS="TXT_0"/>
              <SP WIDTH="15" VPOS="242" HPOS="830"/>
              <String ID="string_9" HPOS="846" VPOS="244" WIDTH="47" HEIGHT="24" WC="0.91" CONTENT="af." STYLEREFS="TXT_0"/>
            </TextLine>
          </TextBlock>
        </ComposedBlock>
      </PrintSpace>
    </Page>
  </Layout>
</alto>
```

Nasjonalbiblioteket har tekstinnholdet til dokumenter lagret som ALTO-XML filer



DH-Laben har bygd infrastruktur for å analysere samlingen basert på ordfrekvensdata



DH-Laben gjør samlingen mer tilgjengelig ved å tilby begrenset tilgang til tekstinneholdet



DH-Labens ressurser tilgjengeliggjør samlingen for forskning

ImagiNation 1814-1905

INNSTILLINGER
Forfatter fornavn:
Forfatter etternavn:
Dato fra og til: 1814 - 1905
Velg kategorier: BYGGE KORPUS

KART AV GEOGRAFISKE OBJEKTER
 A

VELG ELEMENTER I KORPUS
1326 elementer i korpuset
 4 geografiske objekter
1818 Lehmann, Jacob
Reglement for Infanteriet af den Kongelige norske Armee. 1818
 19 geografiske objekter
1820 Winther, Hans Thøger / Hansen, S.
Den Lette Regnekunst, eller Arithmetiske Tabeller, hvorved Enhver, om han end ikke har lært at regne, paa en hurtig Maade kan finde den Summa, som flere Stykker, ALEN, PUND, POTTER &c. koste, naar han alene veed Prisen paa et enkell Stykke, ALEN, PUND, POT &c.; beregnet til norske Specier: Hosfuet en Forklaring over Vægt og Maal, samt en Tabel over Guld- og Sølvmynter, saaledes,
 TRONDHEIM OSM Carto

© Prosjekt ImagiNation 1814-1905 · Norges teknisk-naturvitenskapelige universitet NTNU · 2024

Rediscovering the 1890s: A Norwegian Poetry Corpus*

Ranveig Kvinnland^{1,*}, Ingerid Leyning Dale^{2,†} and Lars Magne Tungland^{2,†}

¹Centre for Ibsen Studies, University of Oslo
²The National Library of Norway

Abstract
What language was used for written poetry in the 1890s in Norway? What was the gender distribution of published poets in that decade? To what degree does the lyrical subject "reveal" itself explicitly or implicitly, in the poems? This article presents answers to these questions in relation to a corpus of poetry written by Norwegian and Danish poets, published between 1890 and 1899. The corpus contains 3,440 poems from 81 books, encoded in TEI XML, and annotated with author and publication metadata, as well as lyrical features such as rhyme schemes and presence of the lyrical subject. The objective of constructing this corpus is to fill a resource gap for research on Norwegian poetry. It provides empirical data for investigating historical claims about the literary period, and comparing findings from both close and distant readings of the data. It is structured in a well-known format (TEI) that enables testing and development of automatic poetry analysis tools, as well as the addition of further annotations. We have published the data in a public repository, making it easily findable, accessible, interoperable and reusable (FAIR).

Keywords
poetry, corpus construction, Norwegian literary history, computational literary analysis, rhyme

1. Introduction

The purpose of this paper is to present a new corpus of digitized poetry published between 1890 and 1899 by Norwegian and Danish poets for a Norwegian audience. The construction of this specific dataset is part of the research project Norwegian Romantic Nationalisms (NORN) at the University of Oslo, aiming to model nineteenth century literature.¹ NORN Poems is a corpus that encapsulates many different linguistic and genre-specific features from a decade with a large literary production in Norway. While previous literary research on the period has mainly focused on a few canonized authors, this corpus enables research on lyrical features both as a cultural trend and as a marker of its time of production.

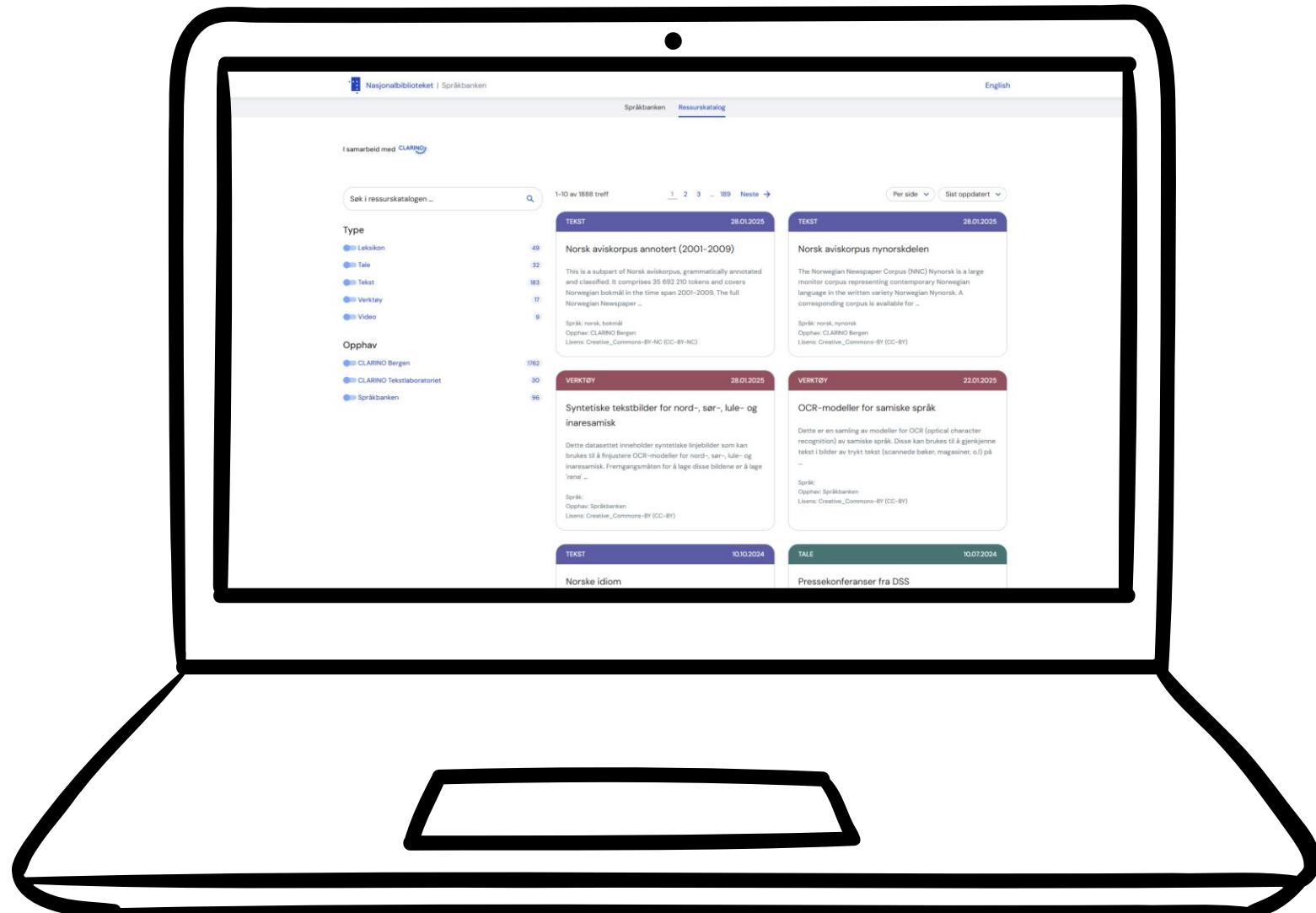
There are two goals connected to this material. The first and main goal is to use computational approaches to examine genre-specific trends in a wide variation of poems, produced

CHR 2024: Computational Humanities Research Conference, December 4–6, 2024, Aarhus, Denmark

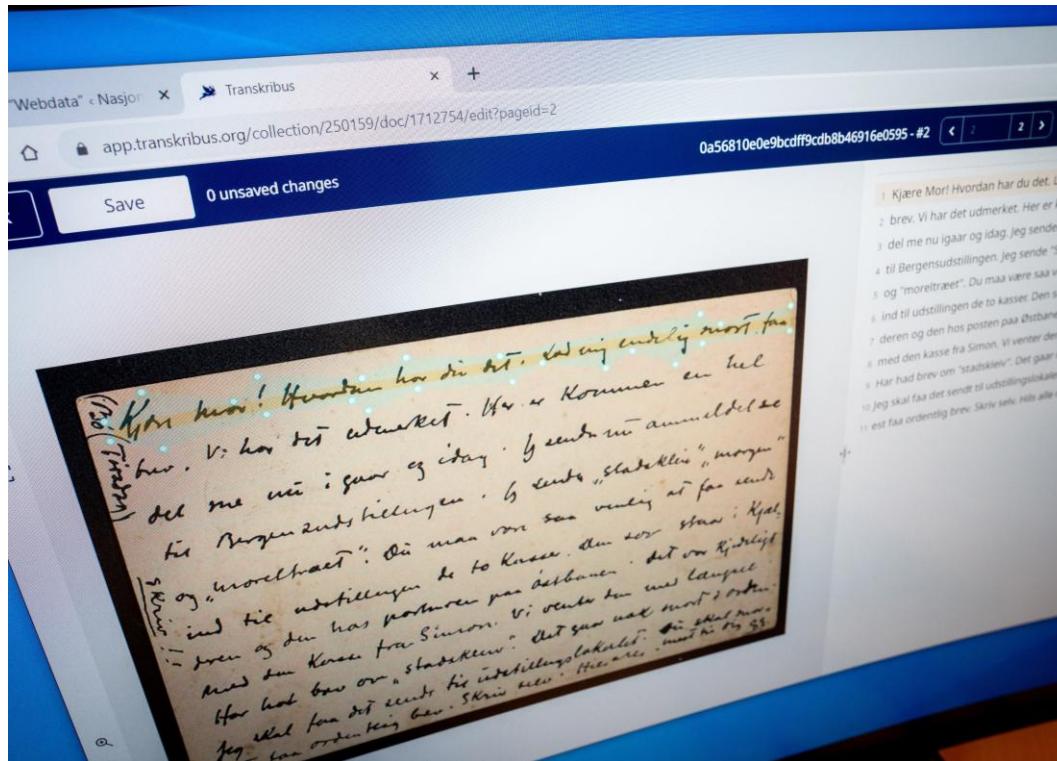
*Corresponding author.
*These authors contributed equally.
✉ ranveig.kvinnland@oslo.no (R. Kvinnland); ingerid.dale@nh.no (I.L. Dale); lars.tungland@nh.no (L.M. Tungland)
○ 0099-0091-0597-5136 (R. Kvinnland); 0099-0091-4134-6169 (I.L. Dale)
○ 0099-0091-0597-5136 (R. Kvinnland); 0099-0091-4134-6169 (L.M. Tungland)
○ 2024 Copyright for this page by its author. Use permitted under Creative Commons License Attribution 4.0 International (CC BY 4.0).
https://www.kf.uio.no/dhi/english/research/projects/norn-norwegian-romantic-nationalisms/index.html



Språkbanken lager og tilgjengeliggjør ressurser for norsk språkteknologi



Hugin-Munin prosjektet transkriberte historiske brev for håndskriftsgjenjening



N Ms.8° 3015 Dagbøker 1824-1830 +

← Tilbake til trefflisten Forrige treff Neste treff →

November 24^{de}

Kjære Mor! Hvordan har du det. Det er endelig mottatt fra ^(det) din brev. Vi har fått videnheten. Her er Kommer en tel del me inn i galleri og etasjer. Og sentet inn annedelene til Bergen utstillingen. Og sentet inn "stadekken" morgon ^{og} mørkhaet. Da man var saa veldig at far sendt inn til utstillingen de to kasser. Den ene med den kasse fra Simon. Vi ventet den. Har hadt brev om "stadekken". Det gaar ikke jeg skal faa der sendt til utstillingssalen. Det er ikke ordentlig brev. Skriv selv. Hils alle m

November 25^{de}

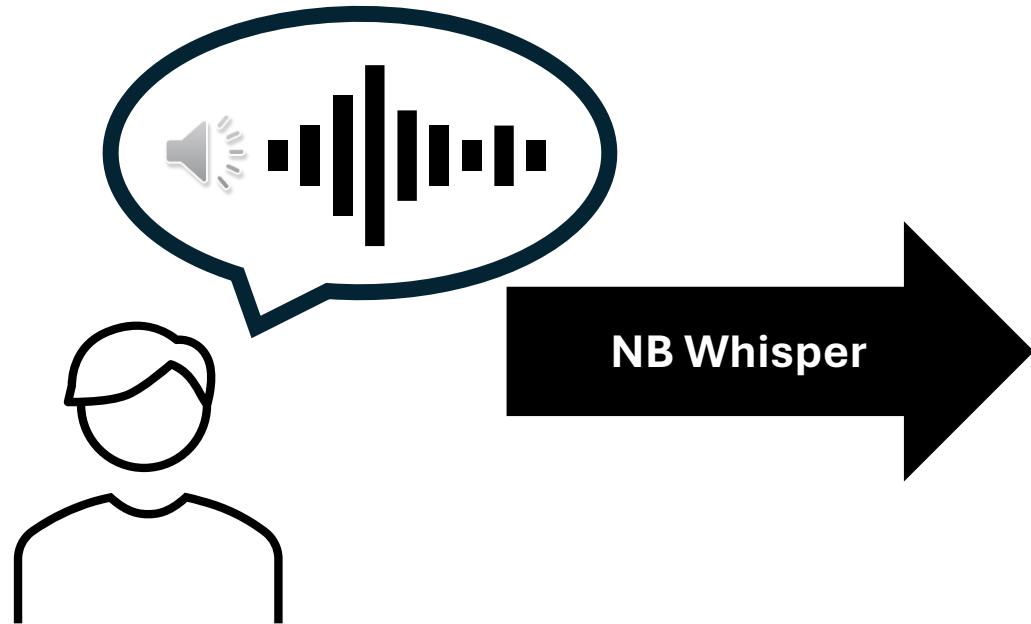
grader. Tidlig morgengang har den været i Dag at seje i jeg er ret Glad at jeg er fremkommet nogenlunde til Freds. — November 24de Slagtnings Syster har været min Dags Forretning som nu ogsaa er forbie. Ingen videre Fremmed, Taalelig Humeur, men Saalelig Post. November 25de Veltlfredes ender jeg denne Dag, et trofrie Sind har været min Glæde, og Son mit Berlids Aftenstunden har jeg dog fornøjet mig med at tylgning noede ^{med} venner til minne ene Elfrida Henriksen, B. og S. En beffjordende Dronne har jeg i Nat, haft om disse Njere. S. Lar min far elskværdig, at mit Hjerte endnu liges i Flame, Jon en Elfrid. — Bertha! derimod synes at være i en meget sonnende Tidstand, dog saa dybt sjukke. Og aldrig minne vere Berthas Ufylgd. O Lykkeled Livetig hellige Blomster. Wie den der føger at forståe din Sjel. —

November 26^{de}

Udmærket fornøjet Sind gis jeg i Stolen, Møller og Hans var hen i Stolen, Møller var Frisk og Johnsgård, som glæder mig, hun trænger dog højlig til Lægeprædelle, for at Læge jo ellers mørke

megen glæde. Kjedsommelig Føre, har der været i Dag at Rejse i, jeg er ret Glad at jeg er fremkommen Nogenlunde Tilfreds. — November 24de Slagtnings Syster har været min Dags Forretning som nu ogsaa er forbie. Ingen videre Fremmed, Taalelig Humeur, men Daarlig Bryst. November 25de Veltlfredes ender jeg denne Dag, et trostende Sind har været min Glæde, og Søn mit Arbejde. I Aftenstunden har jeg dog fornøjet mig med, at tilsende nogle Linnier med Posten til minne ene elskende Veninder, B. og S. — En besynderlig Drom har jeg i Nat, haft om disse kjere, S. var mig saa elskværdig, at mit Hjerte endnu synes i Flamme, som en Elskers. — Bertha! derimod synes at være i en meget Fornedrende Tilstand, dog, saa dybt synker dog aldrig min rene Berthas uskyld. — O Kyskhet Livetig hellige Blomster, Vee den der søger at fordærve Din Dyd. — November 26de Udmærket fornøjet Sind ejer jeg i Aften, Moster Friis, og Hans var her i Aften, Moster var Frisk og fornøjet, som glæder mig, hun trænger dog højlig til Adsprædelse, for at Læge sit ellers mørke

NB-Whisper prosjektet er ga ut en taletranskripsjonsmodell som «forstår» norske dialekter



<https://huggingface.co/spaces/NbAiLab/nb-whisper-demo>

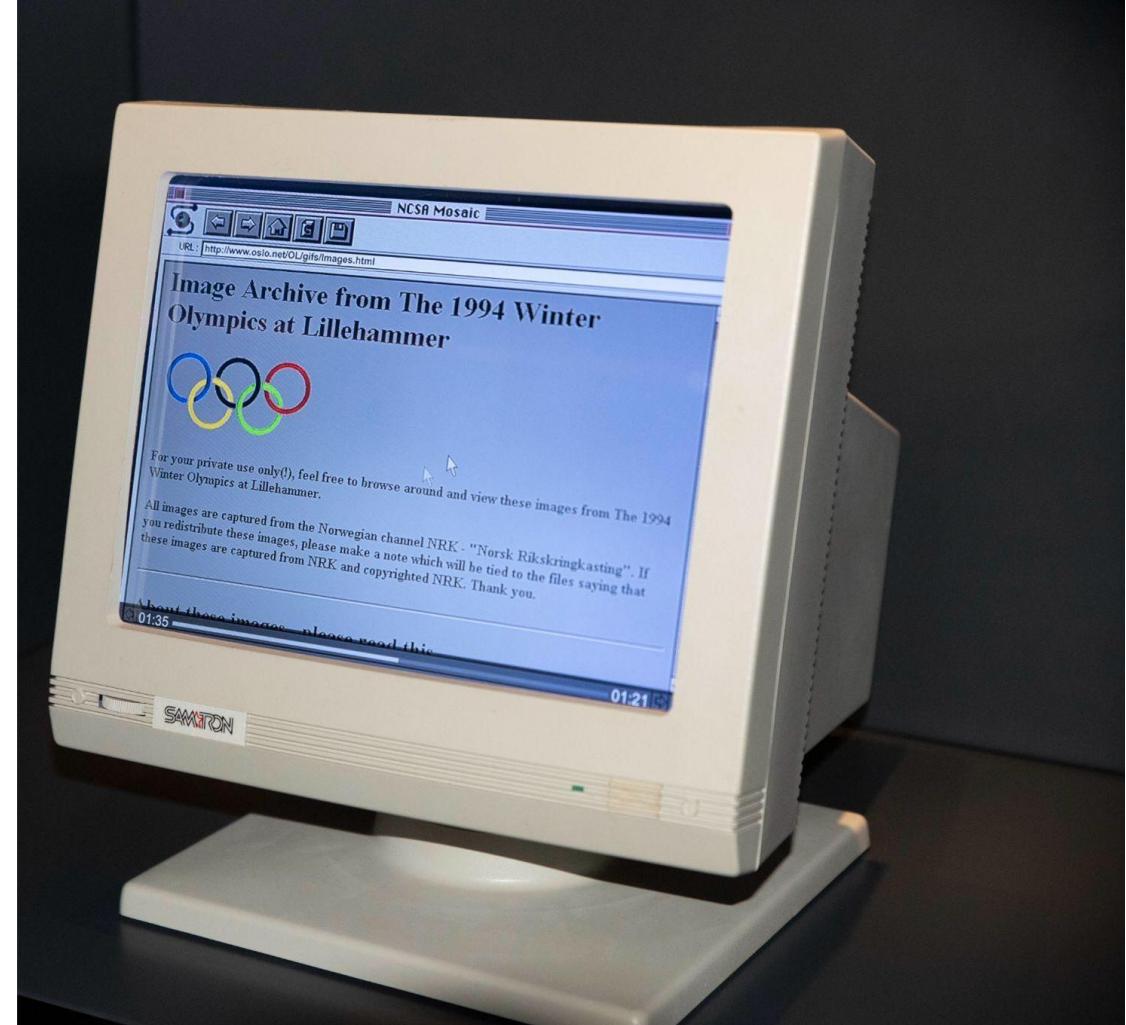
NB-Whisper prosjektet er ga ut en taletranskripsjonsmodell som «forstår» norske dialekter



[00:00:00 -> 00:00:06] Hos Jon Simonsen i Fredrikstad får de alltid kjøkkenutstyr, isenkam osv.
[00:00:06 -> 00:00:10] Gjør et forsøk med kjøkkentøy fra Jon Simonsen i Fredrikstad.

Nettarkivet inneholder arkiverte versjoner av norske nettsider

I Webdata prosjektet vil vi bygge forskningsinfrastruktur for nettarkivet



Kort fortalt gjør DH-Lab appene det mulig å forske på NBs digitaliserte samling fra et fugleperspektiv

