

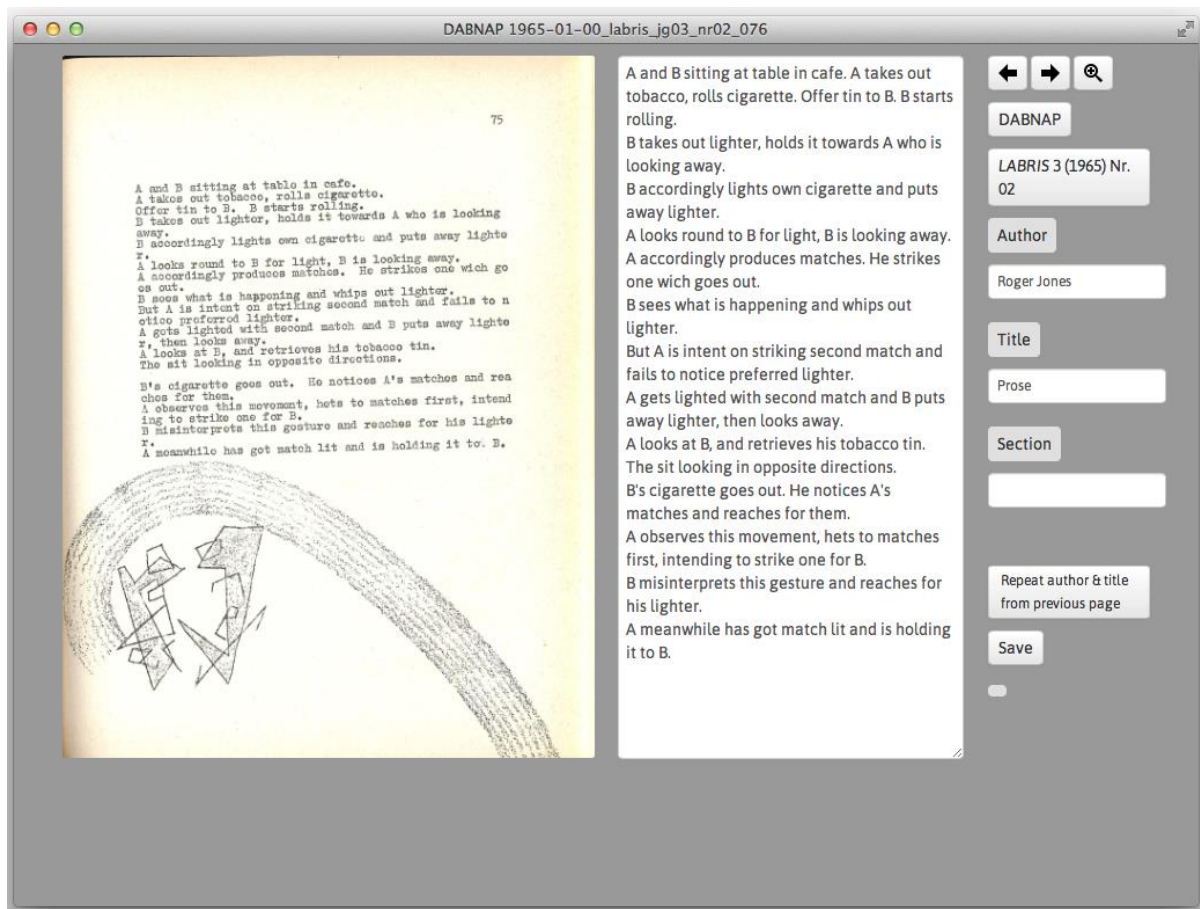
Publisher: Igitur, Utrecht Publishing & Archiving Services. Website: [www.tijdschriftstudies.nl](http://www.tijdschriftstudies.nl)

Content is licensed under a Creative Commons Attribution 3.0 License

URN:NBN:NL:UI:10-1-114147. TS > # 34, december 2013, p. 163-167.

## TS Tools

Het Digitaal Archief van Belgische Neo-Avant-garde-Periodieken (DABNAP) is een onderzoeksproject van de Koninklijke Academie voor Schone Kunsten Antwerpen in samenwerking met de bibliotheek van de Universiteit Antwerpen en de Erfgoedbibliotheek Hendrik Conscience. In oktober 2012 presenteerde Thomas Crombez dit project tijdens het ESPRit-congres *The Magazine as Medium*. In deze aflevering van TS Tools bespreekt hij samen met Sarah Bekaert de uitgangspunten en gebruiksmogelijkheden van de database. Beide auteurs zijn verbonden aan de Koninklijke Academie van Schone Kunsten, Sint Lucas Antwerpen en de Universiteit Antwerpen.



Afb.1: Screenshot van de DABNAP-interface

In DABNAP staan de kunstenaarstijdschriften van de naoorlogse vernieuwers centraal, zoals de CoBRA-beweging, de visuele poëzie, de happening en het experimentele theater. Het project richt zich op periodieke kunstenaarspublicaties (*self-publishing*) van de neo-avant-garde. De initiatiefnemers waren individuele kunstenaars of kunstenaarsgroepen die tegelijk verantwoordelijk waren voor de teksten en beelden, de vormgeving, de productie én de distributie.

Als collectie bekeken, levert dat een opmerkelijke mix op van literatuur, beeldende kunst, kritiek en essayistiek. Ook op materieel vlak werden uiteenlopende keuzes gemaakt wat betreft formaat, papier en druktechniek. Sommige tijdschriften werden gedrukt, maar vaker werden ze gestencild, en in enkele gevallen werd ook van zeefdruk of fotokopie gebruik gemaakt. De periodieken bereikten in de regel een beperkt publiek. Op enkele uitzonderingen na (zoals *De Tafelronde* en *Heibel*) hadden ze bovendien een korte levensduur. De neo-avant-gardistische tijdschriften vertegenwoordigen niet alleen een tijdsgeest, maar zijn ook een verzameling kunstobjecten, en bijgevolg een unieke bron van informatie.

Hoewel de erfgoedsector en het academische milieu (internationaal gezien) een groeiende interesse toont voor neo-avant-gardetijdschriften, blijft de belangstelling in België achterlopen. Omdat de periodieken zich tussen het domein van de literatuur en dat van de beeldende kunst bevinden, dreigt deze unieke collectie in beide disciplines naar de achtergrond te verdwijnen. Het onderzoek wordt nog bemoeilijkt doordat de kunstenaarstijdschriften in eigen beheer werden uitgegeven en vaak een beperkte oplage kenden. Volledige reeksen van bepaalde titels zijn moeilijk te vinden. Sommige gestencilde tijdschriften, zoals *Anar* en *Revo*, zijn zeer kwetsbaar.

Het onderzoeksproject DABNAP wil deze lacune in het onderzoek aanpakken. Meer bepaald stelt het project zich tot doel om het kunstenaarsnetwerk dat zich heeft kunnen vormen en kristalliseren via dit uitgebreide corpus van publicaties, te visualiseren en te analyseren. Dat vraagt om een weloverwogen methodologie op twee niveaus, namelijk het tekstuele en het visuele.

Wat het visuele niveau betreft, worden de digitale scans van de originele tijdschriftpagina's op hoge resolutie (600 dpi) gemaakt, zodat de onderzoeker kan inzoomen op details die voor het blote oog moeilijk zichtbaar zijn. Zo wordt beter duidelijk via welke specifieke technieken het drukwerk tot stand kwam (zeefdruk, fotokopie, stencil, offset, e.a.) en voor welke effecten dat zorgt op gebied van typografie, kleur, compositie en lay-out.

De gebruikte resolutie is aanzienlijk hoger dan de in bibliotheken en archieven gangbare norm van 300 dpi. De keuze voor deze hoge resolutie brengt namelijk ook een betere kwaliteit van de automatische tekstherkenning met zich mee. Tegelijk zijn er ook nadelen aan verbonden. De hoge resolutie maakt het onpraktisch om met ongecomprimeerde bestandsformaten zoals TIFF te werken, maar verplicht tot gecomprimeerde formaten zoals PDF en JPG. Wat daardoor aan digitale duurzaamheid wordt ingeboet (de compressie is onomkeerbaar), compenseert de hogere kwaliteit van de herkende tekst en van de facsimiles die online kunnen worden aangeboden.

De keuze voor hoge-resolutiebeelden is een recent fenomeen. Musea als het [Louvre](#) maken gebruik van opnames op hoge resolutie (tot 15 pixels per mm<sup>2</sup>) om een beperkt aantal topwerken virtueel te exposeren, zodat de gebruiker tot op het niveau van de haarscheurtjes op de verflagen kan inzoomen. Het [Google Art Project](#) maakt in samenwerking met negentien musea van wereldniveau (waaronder de Alte Nationalgalerie, het MoMA, de National Gallery en het Rijksmuseum) meer dan duizend werken via hoge-resolutie-opnames toegankelijk.

In het geval van DABNAP worden de hoge-resolutiebeelden niet via een speciale *viewer* of zoomfunctie aan de bezoeker aangeboden. De beelden kunnen wel met de ingebouwde functies van de browser op die schaal bekeken worden. Daarnaast is de website ook aangepast aan consultering op toestellen met hoge-resolutieschermen (zogenaamde HiDPI- of Retina-schermen).

Ook op het tekstuele niveau worden een aantal innovatieve technieken ingezet, om de digitale collectie voor onderzoekers toegankelijk te maken. Dat gebeurt door de automatisch herkende tekst van de gedigitaliseerde tijdschriften in vier stappen te verwerken:

*(a) Controle van automatische tekstherkenning*

Alle scans worden omgezet in tekst met behulp van courant beschikbare OCR-software (*Optical Character Recognition*). Vanwege de grote diversiteit in gebruikte lettertypes en andere vormen van belettering (handschrift, kleeftletters, zeefdruk) is de kwaliteit van de herkende tekst erg wisselend. Bij gedrukte tijdschriften, zoals *Cobra* of *De Tafelronde*, is de OCR-output nagenoeg foutloos. Bij gestencilde tijdschriften ligt de foutenmarge heel wat hoger. Daarom werd een interface ontwikkeld die de projectmedewerkers in staat stelt om de automatisch herkende tekst na te kijken en te corrigeren (zie afbeelding 1).

*(b) Toevoeging van metadata op basisniveau*

Op dit niveau wordt de documenttekst manueel van de basismetadata voorzien (op niveau van het nummer: tijdschrift, nummer, datum; op niveau van de bijdragen: auteur, titel, paginering) en daarna geïndexeerd. Dat gebeurt via dezelfde interface als hierboven vermeld. De combinatie van deze elementaire data en metadata maakt het mogelijk om op de DABNAP-website niet alleen facsimiles van alle tijdschriftpagina's te presenteren, maar ook *full-text search* aan te bieden. De resultaten van de zoekacties worden gevisualiseerd via eenvoudige grafieken, die de wisselende frequentie waarmee een woord over het hele corpus vermeld wordt, in beeld brengen. Afbeelding 2 toont een screenshot van deze zoekinterface. Het gezochte woord, 'happening', blijkt vooral frequent in de late jaren zestig.

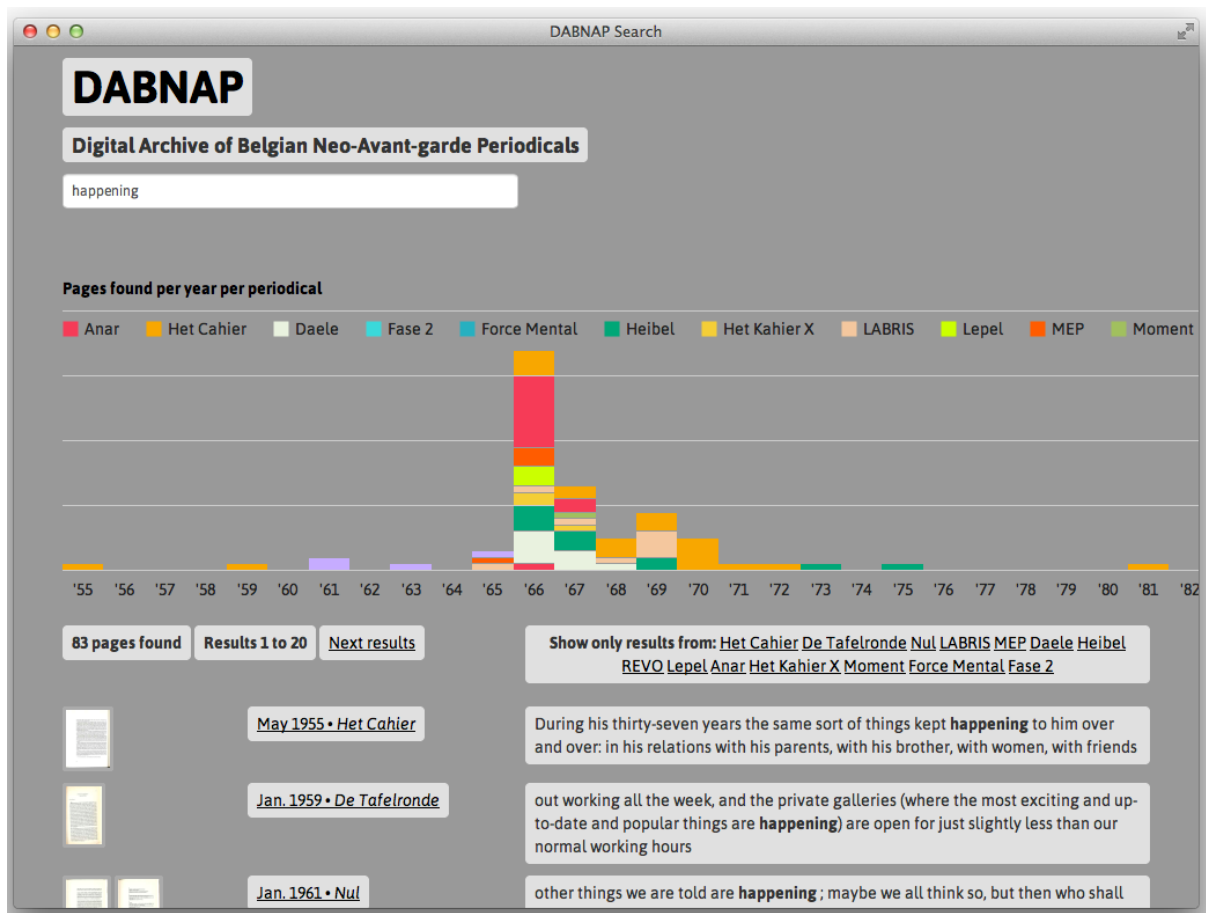
*(c) Elementaire semantische metadata*

Op het derde niveau wordt de documenttekst door computerlinguïstische software automatisch geanalyseerd. Zo worden de documenten van een tweede semantische laag voorzien. Linguïstische software, meer bepaald het Frog-pakket, ontwikkeld door de ILK-groep (Induction of Linguistic Knowledge) van de Universiteit van Tilburg en door

het CLiPS-centrum (Computational Linguistics & Psycholinguistics) van de Universiteit Antwerpen, maakt het mogelijk om automatisch eigennamen te detecteren in Nederlandse tekst. Ook deze informatie wordt als metadata toegevoegd aan de interface van het te ontwikkelen webarchief.

(d) *Relationele semantische metadata*

De elementaire metadata die centraal staan op het derde niveau, maken elk tijdschriftnummer en elke pagina van een dergelijk nummer tot een semantisch rijk document. Deze documenten worden in het project geconcipeerd als een mentale ‘ontmoetingsruimte’ of een ‘evenement’ waar die namen, organisaties en locaties met elkaar in contact komen. Op het vierde niveau worden deze gegevens automatisch geanalyseerd als een semantisch netwerk van betekenisvolle relaties. De software die aan de basis ligt van het webarchief kan eenvoudige relaties automatisch detecteren (zoals ‘is-auteur-van’, ‘is-lid-van’, ‘heeft-geëxposeerd-in’). Ook deze meta-informatie wordt op geautomatiseerde wijze toegevoegd aan de webinterface zodat clusters van vaak samen vermelde of met elkaar gerelateerde namen ook samen aan de gebruiker worden gepresenteerd.



Afb. 2: De zoekinterface van DABNAP

DABNAP ging in januari 2012 van start. Waar staat het project vandaag, na de eerste twee werkjaren? Op dit ogenblik zijn al 27 periodieken gescand en naar tekst omgezet. In totaal betrof het meer dan 33.000 tijdschriftpagina's. Het werkplan voorziet dat er begin 2014 ca. 50.000 pagina's van 40 tijdschriften raadpleegbaar zullen zijn.

Van ongeveer de helft van de nu beschikbare pagina's werd de automatisch herkende tekst al nagekeken. Momenteel wordt de aandacht vooral gericht op de automatische herkenning van eigennamen, en de mogelijke procedures en interfaces om deze metadata manueel na te kijken.

Voorlopig is DABNAP enkel toegankelijk voor onderzoekers (zowel studenten, promovendi als senior-onderzoekers) die een bijzondere belangstelling voor het materiaal hebben en een motivatie kunnen voorleggen. Tegelijk doet zich een groot probleem voor wat betreft de publieke toegankelijkheid van het gedigitaliseerde materiaal. Hoe kan de levensvatbaarheid van het project worden verzekerd, nadat de voorziene onderzoeksmiddelen zijn uitgeput? Gezien de culturele waarde en de moeilijke beschikbaarheid van de periodieken, ligt een publieke ontsluiting voor de hand. Maar vrijwel alle visuele en tekstuele bijdragen vallen onder het auteursrecht. Daarom werd bij de Vlaamse overheid een erfgoedproject aangevraagd om het enorme corpus aan kunsterfgoed dat in het kader van het onderzoek werd gedigitaliseerd, voor een breed publiek te kunnen ontsluiten. Tegelijk wensen we het traject van deze niet voor de hand liggende ontsluiting (aangezien er op vrijwel alle documenten nog auteursrechten rusten) uitgebreid te documenteren, en als 'logboek' van een *casestudy* in digitale ontsluiting voor de erfgoedsector beschikbaar te maken.