

Verslag van de Datasprint Annoteren kun je leren, donderdag 17 oktober 2024, in Amsterdam

In deze datasprint gingen we in kleine groepjes aan de slag met verschillende soorten archiefmateriaal uit de kaartcollectie van het Allard Pierson en met een van de dagboeken uit de Amsterdam Diaries Time Machine.

We bekeken welke soorten annotaties we konden maken op dit materiaal (tags, transcripties, identificatie, etc.), hoe we die informatie het beste konden modelleren en hoe ze dan aansloten bij andere Linked Open Databronnen (bijvoorbeeld inventarissen in RiC-O, thesauri in SKOS, of andere entiteiten in Wikidata) en onderzoek.

Het hogere doel van de middag was om te komen tot een *best practice* van een gedeeld 'IIIF Canvas' voor archiefdocumenten. Daarnaast was het natuurlijk belangrijk om gewoon weer eens bij te praten over nieuwe mogelijkheden voor het werken met Linked Open Data, IIIF en Allmaps met alle data-geïnteresseerde collega's uit den lande.

Leon van Wissen van het Create lab trapte de middag af met een korte inleiding over annoteren. Wat is dat eigenlijk, annoteren? Wat zijn manifesten en welke tools zijn er nu al beschikbaar om gedigitaliseerd materiaal te verrijken? Hoe kan een erfgoedinstelling ervoor zorgen dat digitale bronnen geannoteerd kunnen worden door anderen en als Linked Open Data voor iedereen beschikbaar komen om verder verrijkt te worden? En hoe zorg je ervoor dat al die informatie ook door anderen gevonden en gebruikt kan worden? Vol inspiratie ging iedereen vervolgens in drie verschillende groepen aan de slag.

Onder de pakkende titel **Op de kaart gezet: Atlassen en ander kaartmateriaal** ging een groep samen met Jules Schoonman van de TU Delft Library/Allmaps aan het werk met – je raadt het al - historisch kaartmateriaal en atlassen. Er werd druk geannoteerd, geselecteerd, virtueel geknipt en geplakt aan de hand van willekeurige kaarten in [Allmaps](#). Hierdoor konden bijvoorbeeld verdwenen gebouwen weer zichtbaar gemaakt worden op een actuele kaart. Er werd gekeken hoe verschillende soorten annotaties elkaar konden aanvullen. Immers, als je toponiemen annoteert en linkt, dan wordt het eenvoudiger om een kaart te georefereren, en vice versa. Ook zagen we hoe digitale kaarten niet alleen geografische informatie verstrekken, maar ook als waardevolle bronnen voor historisch onderzoek kunnen dienen. Zie voor meer over deze tafel de link naar het notebook van Jules over de datasprint: [Datasprint: Annoteren kun je leren! / Allmaps | Observable](#)

Een tweede groep werd gevormd rond een tafel met het thema **Amsterdam Diaries Time Machine: dagboeken van vrouwen uit de Tweede Wereldoorlog**. Onder de bezielende leiding van Janna Aerts ging de groep aan de slag. Janna was tot november 2024 als post-doc-onderzoeker verbonden aan Amsterdam Time Machine. In de voorbereidende fase van de [Diaries Time Machine](#) heeft zij veel voorwerk in het begeleiden van een aantal studenten bij het selecteren, transcriberen en taggen van de oorlogsdagboeken. Het dagboek dat ter tafel lag was het digitale [dagboek van T.C. Avis-Hulstijn" \(Avis-Heelsteyn\)](#) dat het NIOD ons voor deze middag heel vriendelijk ter beschikking had gesteld. Er werd geoefend met metadata-verrijking, het maken van transcripties en met het toevoegen van contextuele

informatie. Zie bijvoorbeeld onderstaand voorbeeld van hoe de transcripties en entiteitsannotaties zichtbaar gemaakt werden in een IIIF-viewer:



Of dit voorbeeld van geannoteerde van entiteiten (personen, plaatsen, organisaties en datums) in Transkribus op de herkende tekst in een dagboekscan:



De derde groep, van de tafel **Terug naar de basis: Van traditionele archiefbeschrijving, naar digitale instantiaties** ging niet hands-on aan het werk, maar verloor zich in een inhoudelijke beschouwing over de transformatie van traditionele archiefbeschrijvingen naar Linked Open Data (in het bijzonder de publicatie als IIIF Manifest).

Onder tafelvoorzitterschap van Bob Coret van de KB/NDE/De Gouda Time Machine werd onder meer gesproken over hoe archiefinstellingen de overgang kunnen maken van fysieke beschrijvingen naar gedigitaliseerde en geannoteerde versies van hun materiaal. Ook kwam aan de orde wat het belang is van het creëren van consistente metadata die kunnen aansluiten bij internationale standaarden zoals RiC-O. Zo werd inzichtelijk hoe traditionele archiefdata door middel van digitalisering toegankelijker en rijker kunnen worden, waardoor deze bronnen nieuwe kansen bieden voor onderzoek en hergebruik.

Uit de geanimeerde discussie kwam een aantal to-do-punten of wish-list- onderwerpen naar voren. Die kun je hier onder lezen:

- Vindbaarheid van IIIF Manifesten moet verbeterd worden. Erfgoedorganisaties kunnen hierbij helpen door bij hun beelden een IIIF logo te plaatsen die een link

vormt naar het manifest van het betreffende beeld. Ook is er behoefte aan een soort zoekfunctie: waar kan ik manifesten vinden op basis van geografische of tekstuele zoekopdracht.

- Zou het niet fijn zijn als platformen die beelden annoteren, zoals *VeleHanden* en *HetVolk* maar ook Transkribus en Loghi, hun "ding" kunnen doen op basis van IIF images - waardoor je als erfgoedorganisaties niet steeds massa's aan afbeeldingen uit je DAM moet selecteren en transporteren naar het betreffende platform.
- Zou het niet fijn zijn als genoemde platforms het werk van de vrijwilligers of HTR gaan zien als annotaties en aan hun opdrachtgevers bieden via een Web Annotations Server.

Voor de community en de experts leverde deze tafel dus mooie input om mee naar huis te nemen en het gesprek ook in de eigen instelling verder op te pakken.

Rond half vijf werd na een plenaire terugkoppeling onder leiding van dagvoorzitter Ingeborg Verheul van de Amsterdam Time Machine nog gezellig nageborreld. En uiteraard werden er alweer nieuwe plannen gesmeed voor weer nieuwe datasprints in het volgende jaar.