

App voor analyse van middeleeuwse handschriften in de maak

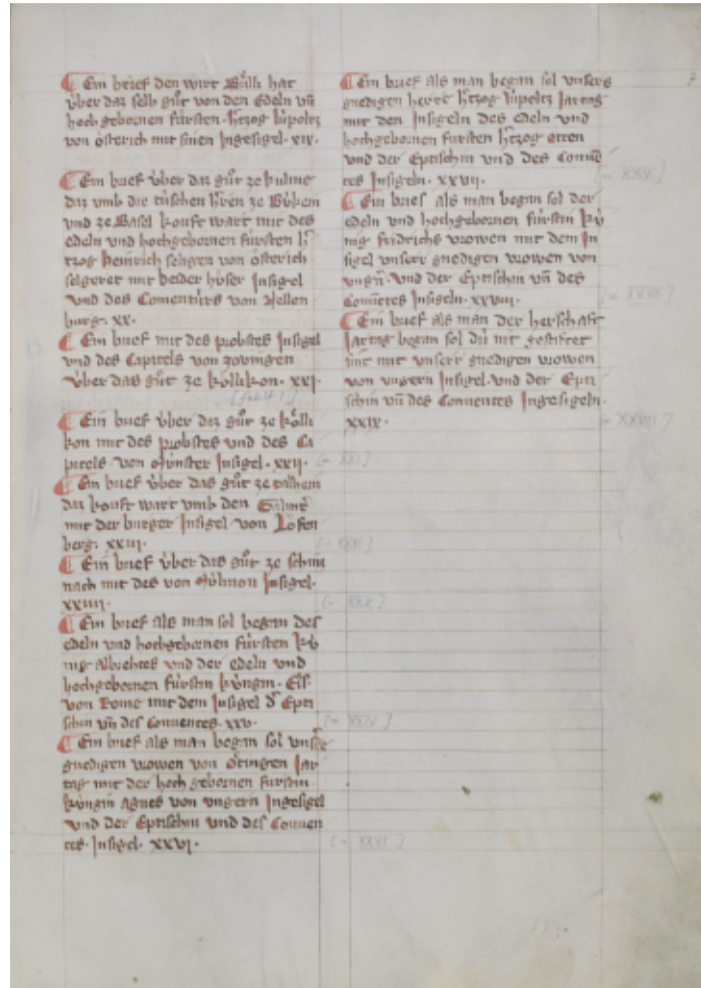
Graven in digitale data

Welke digitale tools zijn nodig voor handschriftenonderzoek? En welke vragen zou je met zo'n tool willen beantwoorden? Een onderzoeksgroep bestaande uit mediëvisten, softwareontwikkelaars, studenten, PhD's en postdocs buigt zich over deze vraag. *Mariken Teeuwen*

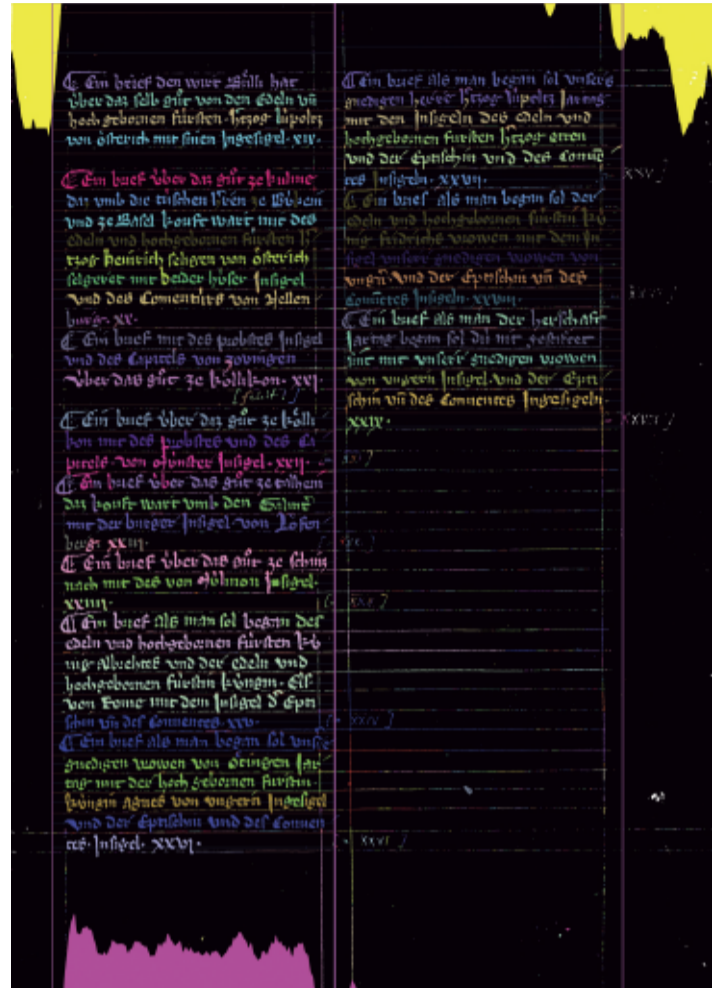
Bij Huygens ING is een denktank gestart voor ideeën en experimenten voor het ontwikkelen van digitale tools voor handschriftenonderzoek. Deze groep van onderzoekers en software engineers onderzoekt de vragen die we zouden willen (en kunnen) benaderen vanuit een digitaal perspectief, en wat we daarvoor zouden kunnen bouwen. Er is in ons vakgebied een aardverschuiving gaande. De digitalisering van handschriftencollecties was al wat langer aan de gang (sinds ca. 2000), maar steeds meer grote bibliotheken die verantwoordelijk zijn voor de opslag en ontsluiting van collecties middeleeuwse handschriften, omarmen het International Image Interoperability Framework (IIIF). Dit framework is een internationaal protocol voor het opslaan van digitale images.

Verschillende analyses

Er gaat een wereld voor ons open. Nu is het mogelijk om één viewer te bouwen waarmee handschriften uit landen van over de hele wereld naast elkaar kunnen worden getoond op één scherm (Mirador). En omdat de images vergezeld gaan van gestructureerde metadata (kenmerken zoals datum, inhoud, afmetingen, geschiedenis, etcetera), is het in principe dus ook mogelijk om met één handeling zoekopdrachten gelijktijdig uit te voeren in ver-



Links staat de originele afbeelding. Rechts staat het resultaat van de digitale toolbox: één image van waaruit handschriftenonderzoekers verschillende analyses automatisch kunnen uitvoeren. credits Aarau, Staatsarchiv Aargau, AA/0428, f. 7r – Cartulary I of Königsfelden (<http://www.e-codices.unifr.ch/en/list/one/saa/0428>)



schillende collecties. En is het mogelijk om dezelfde analyses uit te voeren op images uit verschillende collecties. Dit heeft het idee gevoed om een app te ontwikkelen waarmee de handschriftenonderzoeker vanuit één image verschillende analyses automatisch kan laten uitvoeren. De app heeft nog geen naam, maar laten we het voor nu The Digital Toolbox for the Manuscript Scholar noemen. Met deze Digital Toolbox willen we moge-

lijkheden gaan operationaliseren in één app. De onderzoeker kan de lay-out analyseren, de tekst automatisch laten transcriberen met behulp van Transkribus (een handschrift-OCR systeem dat ontwikkeld wordt met Europees geld) én de metadata zien die beschikbaar zijn in IIIF-collecties. Vervolgens heeft hij of zij mogelijkheden te over om te zoeken naar vergelijkingsmateriaal, naar lokalisering (de bepaling waar een manuscript oorspronkelijk vervaardigd is), naar datering (de bepaling

wanneer het vervaardigd is), etcetera. Kortom: het wordt een tool om snel en effectief naar materiaal te zoeken, en dat materiaal van historische, culturele en inhoudelijke context en interpretatie te voorzien. Wordt vervolgd!

Prof. Dr. Mariken Teeuwen is senior wetenschapper afdeling Wetenschapsgeschiedenis, Huygens ING.
huygens.knaw.nl

Kea Tijdens, coördinator WP8-SERISS-project:

‘Zelf beroepstitels kiezen’

Dankzij het SERISS-project is er een online database met duizenden gecodeerde beroepstitels beschikbaar, inclusief software en handleiding. De coördinator van deze taak, Kea Tijdens, vertelt.

Maurice Martens

“Bijna alle sociaal-economische surveys vragen respondenten naar beroep en opleiding. Het zijn belangrijke indicatoren voor de identiteit van een werkende, voor zijn sociaal-economische status. Ieder land heeft tienduizenden beroepstitels. Om titels tussen landen te kunnen vergelijken, moeten ze geclassificeerd zijn. Het achteraf omzetten naar classificaties kost tijd en geld. Het is efficiënter om respondenten zelf te laten kiezen uit een lijst met beroepstitels. Deze keuzelijst moet

groot genoeg zijn om een beroepstitel te vinden of ten minste één die ‘dicht’ genoeg bij het eigen beroep ligt om tot dezelfde classificatie te worden gecodeerd”, steekt Kea Tijdens van wal.

Wageindicator

Het begon eind jaren negentig. “Paulien Osse, directeur van WageIndicator, maar destijds journaliste, kwam bij me langs. Weekblad *Intermediair* had een salarischek op z'n website geplaatst, maar deze was volledig gericht op academische mannenberoepen. Paulien vroeg of het mogelijk was ook voor vrouwen een salarischek te maken. We hebben toen voor een heel aantal beroepen uitgezocht wat vrouwen verdienen. We hebben verschillende enquêtes opgenomen in de grootste vrouwenbladen en ook een online

versie. 15.000 werkende vrouwen gaven ons hun loon en hun beroep, een groot succes. Later hebben we mannenberoepen toegevoegd en een loonwijzer voor iedereen gemaakt, deze draait inmiddels in 92 landen en heeft 40 miljoen bezoekers per jaar.”

Sterker onderzoek

SERISS staat voor Synergies for Europe's Research Infrastructures in the Social Sciences. Dit vierjarig project richt zich op het versterken van sociaalwetenschappelijk onderzoek in Europa. Eén van de werkpakketten binnen SERISS gaat over het online beschikbaar maken van codeerbare items en de gereedschappen om deze items te integreren in vragenlijsten. Ook de beroepenlijst die Tijdens heeft samengesteld, hoort hierbij. Op surveycodings.org kunnen de sets met

duizenden gecodeerde beroepstitels, industrienaamen, opleidingen en vraagteksten worden opgevraagd. Er wordt uitgelegd wat de beste manier is om deze vragen te stellen en hoe deze te integreren in webvragenlijsten.

surveycodings.org

Over Kea Tijdens:

Kea Tijdens is socioloog en promoveerde in 1989. Zij werkt sinds 1982 bij de Universiteit van Amsterdam en is sinds 2000 onderzoekscoördinator bij AIAS. Ze was gedurende 10 jaar (2004-2014) hoogleraar Vrouw en Arbeid aan de faculteit Sociologie, Erasmus Universiteit Rotterdam. Sinds 2001 is ze ook wetenschappelijk coördinator van de WageIndicator/Loonwijzer web survey.

Linked Data en Syrische bronnen

Reeds in de oudste getuigen van het Syrische christendom zien we een complexe interactie tussen het bijbelse literaire erfgoed en de hellenistische cultuur. De bestudering hiervan vereist een optimaal gebruik van encyclopedische, geografische en lexicografische bronnen. Er komen steeds meer van dergelijke bronnen online beschikbaar, maar ze zijn vaak moeilijk toegankelijk en missen de verbinding met elkaar en de teksten waarop zij betrekking hebben. In de CLARIAH research-pilot LinkSyri: Linking Syriac Data, uitgevoerd door het Eep Talstra Centre for Bible and Computer (VU) en DANS worden deze bronnen met Linked Data toegankelijk gemaakt en met elkaar verbonden. (WvP) clariah.nl/projecten/research-pilots/linksyri