



Opschrift

Vergadering van 15 april 2026

Nummer: 2026_MV_00238

Onderwerp:

Mondelinge vraag van raadslid Mathieu Cockhuyt: Inzet van artificiële intelligentie voor het ontsluiten van historische archieven

Raadslid(-leden):

Mathieu Cockhuyt - N-VA

Bevoegd: Christophe Peeters

Omschrijving van de vraag

Toelichting:

Oude archiefdocumenten zoals schepenregisters bevatten een enorme hoeveelheid waardevolle informatie over het stedelijke verleden. Tegelijk vormen ze door hun omvang en moeilijk leesbaar handschrift vaak een hoge drempel voor onderzoekers, historici, genealogen en geïnteresseerde burgers.

In Mechelen werd onlangs een interessant project opgezet waarbij het Stadsarchief samen met historici en archiefvrijwilligers artificiële intelligentie inzet om laatmiddeleeuwse schepenregisters automatisch te transcriberen. Via het platform Transkribus worden handgeschreven documenten uit het ancien régime omgezet naar doorzoekbare tekst. Daardoor worden deze bronnen plots veel toegankelijker voor een breed publiek. Inmiddels staan daar al tientallen registers volledig doorzoekbaar online.

Ook in Gent bevinden zich in het stadsarchief grote hoeveelheden historische documenten die nog niet geïndexeerd of getranscribeerd zijn en daardoor slechts beperkt toegankelijk blijven.

In 2024 werd door de stad nog een persbericht uitgestuurd over het vernieuwde, state-of-the-art scanatelier in De Zwarte Doos. Het lijkt me dat we dus minstens de hardware hebben om in Gent een gelijkaardig traject op te zetten om deze waardevolle teksten te ontsluiten naar onze burgers.

Vraag:

Graag had ik hierover volgende vragen gesteld:

- Wordt binnen het Gentse stadsarchief al onderzocht hoe artificiële intelligentie kan worden ingezet om historische documenten sneller te ontsluiten?
- Zijn er vandaag projecten of pilootinitiatieven rond automatische transcriptie of digitalisering van handgeschreven bronnen?

- Indien dit niet het geval is, is de schepen bereid te bekijken of Gent hier, eventueel samen met vrijwilligers of onderzoeksinstituten, een gelijkaardig traject kan opstarten om deze waardevolle archieven beter toegankelijk te maken?

Antwoord

Wordt binnen het Gentse stadsarchief al onderzocht hoe artificiële intelligentie kan worden ingezet om historische documenten sneller te ontsluiten?

- Archief Gent beschikt inderdaad over de vereiste hardware om (historische) documenten te digitaliseren en zet de verschillende scanapparatuur ook bijna dagelijks in om stukken te digitaliseren op vraag (van internen of externen) of om projectmatig collecties te digitaliseren. Dit zowel door eigen personeel als vrijwilligers.
- Ook het toepassen van AI gebeurt al: de collectie Gentse Wegwijzers zijn gedigitaliseerd en via OCR (optical character recognition) doorzoekbaar gemaakt. We wachten momenteel de komst van de nieuwe beeldbank af (project in samenwerking met District09) om te zien of het ontsluiten van de digitale bestanden mét OCR mogelijk zal zijn (niet elke toepassing kan dit zomaar). Deze nieuwe beeldbank zal eind 2026 opgeleverd worden en Archief Gent zit mee in het projectteam.
- Het archiefbeheersysteem dat Archief Gent gebruikt, kan een koppeling maken met de diensten van Transkribus voor de transcriptie van handgeschreven bronnen. Die koppeling is (nog) niet aangekocht door Archief Gent omdat (1) eerst onderzocht moet worden hoe efficiënt de AI-modellen werken op de eigen collectie maar vooral omdat (2) er geen e-depot beschikbaar is om de gedigitaliseerde beelden in op te slaan en via API's door te sturen vanuit het archiefbeheersysteem naar Transkribus en (3) omdat elke transcriptie geld kost, ook met een API-koppeling. De efficiëntie uittesten zonder dat we beelden effectief kunnen doorsturen (vanwege gebrek aan opslag en middelen) zou ook niet nuttig zijn aangezien de AI-modellen razendsnel evolueren en verbeteren.
- AI wordt ook gebruikt voor het transcriberen van onze audiovisuele collectie: zowel de opties binnen het Adobe-pakket als (zeer recent) een opensource toepassing waarvoor District09 onlangs (13/03) toestemming gaf om dit te mogen downloaden en testen. Die collectie wordt ontsloten op Hetarchief.be van Meemoo.

Zijn er vandaag projecten of pilootinitiatieven rond automatische transcriptie of digitalisering van handgeschreven bronnen?

- Archief Gent neemt momenteel niet deel aan een project; (zie ook volgende vraag)

Indien dit niet het geval is, is de schepen bereid te bekijken of Gent hier, eventueel samen met vrijwilligers of onderzoeksinstituten, een gelijkaardig traject kan opstarten om deze waardevolle archieven beter toegankelijk te maken?

- In de erfgoedsector worden verschillende projecten opgezet, zoals recent het MODAL-project met ook een Gentse speler of een project van de KU Leuven. De

- medewerkers van Archief Gent houden deze projecten en de resultaten in de gaten en nemen ook deel aan studiedagen of voorstellingen van de resultaten. Bij veel van deze projecten is ook Meemoo betrokken, waarvan Archief Gent een Contentpartner is en waarvan medewerkers verschillende studiedagen volgen.
- Het zélf opzetten van een AI-project vraagt (IT/technische) kennis die Archief Gent niet in huis heeft en vaak veel middelen waardoor samenwerkingen nodig (en wenselijk) zijn. Het zou efficiënter zijn om als archiefinstelling in een samenwerking te stappen (maar zonder technische kennis geen voortrekkersrol aan te nemen) Onze collecties aanbieden om AI-modellen te trainen, is zeker mogelijk. Maar de vraag gaat hier specifiek over het ontsluiten en toegankelijk maken van de stukken, nadat er AI op toegepast is. En daar knelt het schoentje: enerzijds door de overgang naar de nieuwe beeldbank, maar zeker ook wegens gebrek aan e-depot om als opslaglocatie te dienen.

Archief Gent volgt de ontwikkelingen dus op, neemt deel aan studiedagen en overleggroepen rond het onderwerp, past AI toe wanneer het betaalbaar is en duidelijke resultaten oplevert (projecten hebben niet altijd de gewenste resultaten, zeker niet als het gaat over pilootprojecten die een nieuwe technologie uittesten) maar heeft nog niet ingezet op grote AI-projecten. Binnen de sector zijn er ook stemmen die net met opzet niet mee willen springen op de nieuwste hype maar net kwalitatief verder werken aan de basiswerking en ontsluiting om geen middelen te spenderen aan onzekere technologieën (pilotprojecten) of dure commerciële tools (transkribus), zonder de kwaliteit en mogelijkheden van die tools te ontkennen natuurlijk.
