

## Inleiding tot de digitale tekstanalyse (A704153)

**Cursusomvang** *(nominale waarden; effectieve waarden kunnen verschillen per opleiding)*

**Studiepunten 3.0** **Studietijd 90 u**

**Aanbodsessies en werkvormen in academiejaar 2024-2025**

A (semester 1)	Nederlands	Gent	hoorcollege zelfstandig werk
B (semester 1)	Nederlands	Gent	

**Lesgevers in academiejaar 2024-2025**

Lefever, Els	LW22	Verantwoordelijk lesgever
Hoste, Veronique	LW22	Medelesgever

**Aangeboden in onderstaande opleidingen in 2024-2025**

	stptn	aanbodsessie
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Duits, Italiaans)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Duits, Russisch)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Duits, Spaans)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Duits, Turks)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Duits, taaltechnologie)	5	B
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Engels, Duits)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Engels, Frans)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Engels, Italiaans)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Engels, Russisch)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Engels, Spaans)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Engels, Turks)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Engels, taaltechnologie)	5	B
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Frans, Duits)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Frans, Italiaans)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Frans, Russisch)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Frans, Spaans)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Frans, Turks)	3	A
Bachelor of Arts in de toegepaste taalkunde: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Frans, taaltechnologie)	5	B
Micro-credential Inleiding tot de taaltechnologie	3	B
Schakelprogramma tot Master of Arts in het tolken: combinatie van ten minste twee talen, Master of Arts in het vertalen: combinatie van ten minste twee talen en Master of Arts in de meertalige communicatie: combinatie van ten minste twee talen	3	A

(afstudeerrichting Nederlands, Duits, Spaans)		A
Schakelprogramma tot Master of Arts in het tolken: combinatie van ten minste twee talen, Master of Arts in het vertalen: combinatie van ten minste twee talen en Master of Arts in de meertalige communicatie: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Engels, Duits)	3	A
Schakelprogramma tot Master of Arts in het tolken: combinatie van ten minste twee talen, Master of Arts in het vertalen: combinatie van ten minste twee talen en Master of Arts in de meertalige communicatie: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Engels, Frans)	3	A
Schakelprogramma tot Master of Arts in het tolken: combinatie van ten minste twee talen, Master of Arts in het vertalen: combinatie van ten minste twee talen en Master of Arts in de meertalige communicatie: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Engels, Spaans)	3	A
Schakelprogramma tot Master of Arts in het tolken: combinatie van ten minste twee talen, Master of Arts in het vertalen: combinatie van ten minste twee talen en Master of Arts in de meertalige communicatie: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Frans, Duits)	3	A
Schakelprogramma tot Master of Arts in het tolken: combinatie van ten minste twee talen, Master of Arts in het vertalen: combinatie van ten minste twee talen en Master of Arts in de meertalige communicatie: combinatie van ten minste twee talen(afstudeerrichting Nederlands, Frans, Spaans)	3	A
Schakelprogramma tot Master of Arts in de meertalige communicatie: combinatie van ten minste twee talen	3	A
Schakelprogramma tot Master of Arts in het vertalen: combinatie van ten minste twee talen	3	A
Vorbereidingsprogramma tot Master of Arts in de meertalige communicatie: combinatie van ten minste twee talen	5	A, B
Vorbereidingsprogramma tot Master of Arts in het tolken: combinatie van ten minste twee talen	3	A
Vorbereidingsprogramma tot Master of Arts in het vertalen: combinatie van ten minste twee talen	5	A, B
Keuzepakket Taalkunde	3	A

## Onderwijstalen

Nederlands

## Trefwoorden

digitale tekstanalyse, digital humanities, digitale geletterdheid

## Situering

Het doel van dit opleidingsonderdeel is:

- de student bewust maken van het belang van digitale geletterdheid, de competentie die hen in staat stelt om met behulp van de computer digitale bronnen te kunnen aanleggen, verrijken, verwerken en analyseren.
- de student een algemeen overzicht bieden van digitale technieken en onderzoeksmethodes binnen de onderzoeksvelden van de humanities of geesteswetenschappen, met een bijzondere focus op digitale tekstanalyse (bijv. computationele taalkunde en sociolinguïstiek, natuurlijke taalverwerking, corpustaalkunde, netwerkanalyse).
- de student kritisch laten reflecteren over de specifieke uitdagingen van datagebaseerd onderzoek, met specifieke focus op bias die geïntroduceerd wordt door de selectie van tekstcorpora voor onderzoek en het bouwen van tools.
- de student laten kennismaken met concrete onderzoeksprojecten en -technieken binnen de digital humanities, opnieuw met de focus op digitale tekstanalyse (bijv. case studies binnen taal- en vertaaltechnologie, geschiedenis, literatuurwetenschappen, kunstgeschiedenis, enz.)

## Inhoud

Dit opleidingsonderdeel biedt een overzicht van de belangrijkste digitale technieken binnen de verschillende onderzoeksvelden van de humanities, met de focus op digitale tekstanalyse. In het eerste deel wordt ingegaan op het meer algemene onderzoeksdomein van de digital humanities, waarbinnen digitale tekstanalyse een centrale plaats inneemt. Er wordt een theoretisch overzicht gegeven van de belangrijkste digitale technieken en methodes voor automatische tekstanalyse, waarbij gestart wordt met het verzamelen, digitaliseren en verrijken van bronnen, en het automatisch doorzoeken van grote corpora/tekstcollecties. Daarop volgt een overzicht van methodes om te komen tot een automatisch tekstbegrip, gaande van het

oppervlakkigere lexicale niveau tot meer complexe semantische niveaus. Hierbij wordt ook ingegaan op de problemen die inherent zijn aan het verwerken van natuurlijke taal, zoals ambiguïteit en wereldkennis.

Naast deze theoretische onderbouwing wordt ook uitgebreid aandacht besteed aan kritische reflectie over de selectie en het gebruik van bronnen voor datagebaseerd onderzoek, het bouwen en toepassen van technologie, de impact van bijvoorbeeld zelflerende systemen en AI-toepassingen op de maatschappij en het belang van een goed begrip van de onderliggende algoritmes voor het correct toepassen en interpreteren van de uitvoer van deze tools.

In het tweede deel komen verschillende use cases aan bod uit het bredere onderzoeksdomein van de digital humanities, opnieuw met focus op digitale tekstanalyse. Daarbij wordt het hedendaags onderzoek en de belangrijkste tendensen in kaart gebracht voor enkele specifieke onderzoeksdomeinen (bijv. automatische tekstanalyse voor klassieke talen, het gebruik van computermethodes voor de exploratie en analyse van grote tekstcollecties of van hedendaagse digitale media) en er worden een of meerdere representatieve onderzoeksprojecten voorgesteld.

Tijdens de begeleide zelfstudie lezen de studenten wetenschappelijke artikels ter voorbereiding van specifieke cursusonderdelen, doorlopen ze modules van een digitaal leerpad en maken ze oefeningen.

### **Begincompetenties**

Algemene vorming die kan verwacht worden vanuit het secundair onderwijs.

### **Eindcompetenties**

- 1 Basiskennis en inzicht hebben in de verschillende digitale technieken en onderzoeksmethodes binnen de geesteswetenschappen, in het bijzonder methodes voor digitale tekstanalyse. [B.1.7.; met toetsing], [LWDC.6.1.; met toetsing]
- 2 Denk- en redeneervaardigheid en het vermogen tot kritische reflectie betreffende het bouwen en toepassen van tools voor digitale tekstanalyse [B.3.3.; met toetsing], [B.5.1.; met toetsing]
- 3 Onderzoeksvaardigheden: wetenschappelijke literatuur. [B.2.3.; met toetsing]
- 4 Reflectie over de wetenschappelijke en maatschappelijke rol van Digital Humanities-onderzoek binnen de socioculturele context waarin het gevoerd wordt. [B.5.1.; met toetsing], [LWDC.6.2.; met toetsing]

### **Creditcontractvoorwaarde**

De toegang tot dit opleidingsonderdeel via creditcontract is open: de student houdt zelf rekening met voorkennis uitgedrukt in begincompetenties

### **Examencontractvoorwaarde**

De toegang tot dit opleidingsonderdeel via examencontract is open

### **Didactische werkvormen**

Hoorcollege, Zelfstandig werk

### **Toelichtingen bij de didactische werkvormen**

Theoretisch deel: hoorcolleges, discussie van wetenschappelijk onderzoek

Praktisch deel: begeleide zelfstudie via digitaal leerpad

### **Studiemateriaal**

Type: Handboek

Naam: Taaltechnologie ontrafeld  
Richtprijs: € 27  
Optioneel: nee  
Taal : Nederlands  
Auteur : Véronique Hoste, Cynthia Van Hee

Type: Slides

Naam: slides  
Richtprijs: Gratis of betaald door opleiding  
Optioneel: nee  
Beschikbaar op Ufora : Ja

### **Referenties**

Rens Bod, De Vergeten wetenschappen, Prometheus, 2010.

Daniel Jurafsky & James Martin, "Speech and Language Processing. An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition", 2008. Updated

### **Vakinhoudelijke studiebegeleiding**

Ondersteunend discussieforum

### **Evaluatiemomenten**

periodegebonden en niet-periodegebonden evaluatie

### **Evaluatievormen bij periodegebonden evaluatie in de eerste examenperiode**

Schriftelijke evaluatie

### **Evaluatievormen bij periodegebonden evaluatie in de tweede examenperiode**

Schriftelijke evaluatie

### **Evaluatievormen bij niet-periodegebonden evaluatie**

Participatie

### **Tweede examenkans in geval van niet-periodegebonden evaluatie**

Examen in de tweede examenperiode is mogelijk

### **Toelichtingen bij de evaluatievormen**

Traject met twee vreemde talen:

- niet-periodegebonden evaluatie: De studenten worden beoordeeld op het afwerken van digitale leerpaden en het indienen van kleine opdrachten
- periodegebonden evaluatie: Schriftelijke examens

Traject met taaltechnologie: idem als traject met twee vreemde talen. Bij het traject taaltechnologie wordt een bijkomende verdiepende taak uitgewerkt (ter waarde van 2 studiepunten).

### **Eindscoreberekening**

- permanente (niet-periodegebonden) evaluatie: 10%
- periode-gebonden evaluatie: 90%

### **Faciliteiten voor werkstudenten**

Mogelijkheid tot vrijstelling van aanwezigheid

Geen mogelijkheid tot examen op ander tijdstip in het academiejaar

Mogelijkheid tot feedback via e-mail of na afspraak tijdens de kantooruren, beperkt tot het beantwoorden van specifieke vragen.