

# Verkenning van de inzet van *topic modelling* bij het analyseren van schrijfoopdrachten

Anouschka van Leeuwen, Ayoub Bagheri, Thom Benjamin Volker en Charlotte N.C. van Brakel

**Samenvatting** In het hoger onderwijs wordt regelmatig gebruik gemaakt van tussenopdrachten die studenten zelfstandig dienen uit te voeren. In dit artikel gaan we specifiek in op de *flipped classroom*, waarbij studenten schriftelijke opdrachten voltooiën als voorbereiding op een begeleide werkgroepbijeenkomst. Deze opdrachten vormen potentieel een waardevolle bron van informatie voor de docent, die hieruit informatie kan destilleren over het begrip van de stof van de studenten. In de praktijk komen docenten er echter niet altijd aan toe om de opdrachten voorafgaand aan de werkgroepbijeenkomst te bekijken. In deze studie verkennen we daarom de perceptie van docenten omtrent de mogelijkheid om schriftelijke opdrachten te analyseren met een *text mining* techniek, namelijk *topic modelling* (TM), waarbij de meest genoemde onderwerpen uit de opdrachten gedestilleerd worden. Een team van drie docenten uit een bachelorcursus gaven middels interviews hun mening over de potentie van TM. De docenten zagen potentie voor TM op drie manieren, namelijk 1) op basis van TM met individuele studenten in gesprek gaan, 2) de uitkomsten van TM als werkvorm aan studenten voorleggen, en 3) de uitkomsten van TM met het gehele docententeam bespreken. Aan de andere kant vraagt de inzet van TM wel een serieuze tijdsinvestering. Een afweging van de voor- en nadelen moet dus in elke situatie gemaakt worden.

**Trefwoorden** Hoger onderwijs, *flipped classroom*, *topic modelling*, schrijfoopdrachten, docenten

## Probleem en context

In het hoger onderwijs wordt regelmatig gebruik gemaakt van tussenopdrachten die studenten zelfstandig dienen uit te voeren, en die niet beoordeeld worden met een cijfer maar wel dienen als verwerking van het cursusmateriaal. In dit artikel gaan we specifiek in op de context van de *flipped classroom* (Staker & Horn, 2012), waarbij studenten schriftelijke opdrachten voltooiën als voorbereiding op een begeleide werkgroepbijeenkomst. In dergelijke opdrachten beantwoorden studenten enkele vragen die betrekking hebben op de literatuur die in de desbetreffende week behandeld wordt. Studenten zetten in de opdrachten hun initiële gedachten over de cursusstof op papier, en deze opdrachten vormen daarmee potentieel een waardevolle bron van informatie voor de docent, die

## Artikelgeschiedenis

Ontvangen: 27 juni 2022  
Geaccepteerd: 4 september 2022  
Online: 18 juli 2023

## Contactpersoon

Anouschka van Leeuwen,  
A.vanLeeuwen@uu.nl

## Over de Auteur(s)

Anouschka van Leeuwen is werkzaam bij de afdeling Onderwijswetenschappen, Universiteit Utrecht; Ayoub Bagheri is werkzaam bij de afdeling Methoden & Statistiek, Universiteit Utrecht; Thom Benjamin Volker is werkzaam bij de afdeling Methoden & Statistiek, Universiteit Utrecht; Charlotte N.C. van Brakel is werkzaam bij de afdeling Onderwijswetenschappen, Universiteit Utrecht

## Copyright

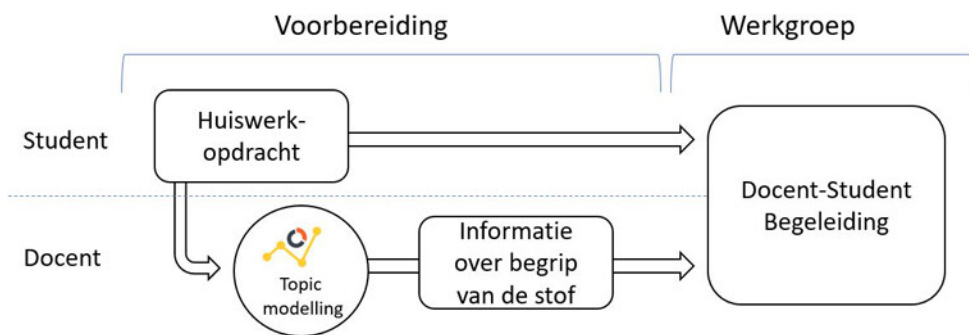
© Author(s); licensed under Creative Commons Attribution 4.0. This allows for unrestricted use, as long as the author(s) and source are credited.

hieruit informatie zou kunnen destilleren over het begrip van de stof van de studenten en daarmee adaptieve begeleiding kan bieden (Vogt & Rogalla, 2009). In de praktijk komen docenten er echter niet altijd aan toe om de opdrachten voorafgaand aan de werkgroep te bekijken, bijvoorbeeld door hoge studentaantallen of door tijdgebrek. In deze studie verkennen we daarom de mogelijkheid om schriftelijke opdrachten te analyseren met een *text mining* techniek, namelijk *topic modelling* (Blei et al., 2003), waarbij de meest genoemde onderwerpen uit de opdrachten gedestilleerd worden. We leggen de uitkomsten voor aan het docententeam van een bacheloropleiding en bestuderen de mogelijke toepassing van *topic modelling* in deze context.

### Theoretisch kader

Huiswerkopdrachten zijn gebruikelijk in het hoger onderwijs, vooral in *flipped classroom*-modellen. In het *flipped classroom* model zijn studenten zelf verantwoordelijk voor het bestuderen van het cursusmateriaal, waarna verdiepende discussie in begeleide werkgroepen plaatsvindt (Staker & Horn, 2012). Voor het bestuderen van het cursusmateriaal worden vaak kennisclips en huiswerkopdrachten gebruikt (Van Alten, Phielix, Janssen, & Kester, 2019). Ook in deze studie bekijken we een cursus met het flipped classroom model. Studenten bestuderen kennisclips en artikelen, en leveren vervolgens een opdracht in waarin ze reflecteren op het onderwerp van die week. Daarna doen studenten mee aan een werkgroep die door een docent begeleid wordt en waar het cursusmateriaal op een dieper niveau verwerkt wordt.

Voor docenten kan het bestuderen van de huiswerkopdrachten voorafgaand aan de werkgroep een middel zijn om het begrip van de studenten van het cursusmateriaal te monitoren (Van Leeuwen, 2019). Deze informatie kan de docent vervolgens gebruiken om de inhoud van de werkgroep aan te passen. Een docent kan bijvoorbeeld meer aandacht vestigen op concepten die (een deel van) de studenten nog niet volledig begrijpen. Door op deze manier in te spelen op het begrip van de stof, kan een docent *adaptief* onderwijs bieden (Vogt & Rogalla, 2009). Aangenomen wordt dat adaptieve begeleiding leidt tot betere kennisverwerking door de student. Het doornemen van de ingeleverde huiswerkopdrachten kan dus een mogelijke weg zijn om adaptieve begeleiding en daarmee kennisverwerking door de studenten te bevorderen. Echter, gegeven dat een gemiddelde werkgroep 25 of meer studenten bevat en docenten over het algemeen een hoge werkdruk hebben (Universiteiten van Nederland, 2022), is het voor docenten niet altijd mogelijk om huiswerkopdrachten van studenten voorafgaand aan werkgroepen door te nemen. In dit artikel verkennen we daarom of tekstanalyse van huiswerkopdrachten een informatiebron kan zijn voor docenten, zodat zij niet meer elke ingeleverde tekst afzonderlijk hoeven te lezen maar een samenvatting krijgen van de belangrijkste onderwerpen die studenten noemen. De techniek die we daarvoor in dit artikel gebruiken staat bekend onder de naam *topic modelling* (Blei, Ng, & Jordan, 2003).



**Figuur 1** Mogelijke rol van *topic modelling* in het onderwijsproces

*Topic modelling* (TM) is een techniek voor het ontdekken van de onderliggende onderwerpen die in een verzameling documenten voorkomen. Het is een instrument voor het ontdekken van verborgen semantische structuren in een verzameling tekst (Blei et al., 2003). Deze techniek analyseert dus de inhoud van de activiteiten van studenten en niet zozeer de kwantiteit van de activiteit. In een eerdere studie (Van Leeuwen, 2019) voorzagen we docenten in kwantitatieve informatie over de activiteiten van studenten, bijvoorbeeld of ze een bepaalde activiteit hadden voltooid. Deze informatie (ook wel: “*learning analytics*”) leverde slechts in beperkte mate *actionable insights* op; docenten konden op basis van deze informatie weinig vervolgcacties gericht op de cursusstof initiëren. Daarom verkennen we nu of meer inhoudelijke analyses van de opdrachten die studenten inleveren, docenten kunnen helpen in de voorbereiding op werkgroepen (Kaliisa, Kluge, & Mørch, 2020).

In voorgaande studies is de toepassing van TM vooral bekeken als een manier om summatief te evalueren. In de studie van Chen, Yu, Zhang en Yu (2016) werd TM bijvoorbeeld ingezet om de reflectieve opdrachten van docenten in opleiding te evalueren en een beoordelingssysteem te maken dat automatisch een cijfer kon toekennen aan de opdrachten. In deze verkennende studie willen we juist bekijken of TM van waarde kan zijn bij een formatieve blik op de opdrachten die studenten maken. In Figuur 1 is samengevat hoe TM zou kunnen worden gepast in een cursus met een *flipped classroom* model. De studenten maken een huiswerkopdracht ter voorbereiding op de werkgroep. TM wordt toegepast op de ingeleverde opdrachten, en de resultaten daarvan worden aan de docent doorgegeven. De docent kan hieruit mogelijk informatie krijgen over het begrip van de stof van de studenten, en deze informatie gebruiken bij de voorbereiding van de werkgroep en de begeleiding die de docent biedt tijdens de werkgroep.

In dit artikel bestuderen we of TM inderdaad bruikbare inzichten kan opleveren en hoe deze inzichten zich verhouden tot de situatie waarin de docent alle huiswerkopdrachten volledig zou lezen. In dit verkennende onderzoek voeren we TM uit op huiswerkopdrachten en leggen de resultaten voor aan een docententeam om te achterhalen of TM toegevoegde waarde kan hebben in het onderwijsproces. De onderzoeksvraag die hier-

bij is opgesteld is: *wat zijn de percepties van docenten omtrent de toepassing van Topic Modelling in het analyseren van schrijfofdrachten in bachelorcurssussen met een flipped classroom design?*

## Methode

### Design

Het doel van deze studie was het verkennen van de meerwaarde van de toepassing van *Topic Modelling* (TM) bij het analyseren van schrijfofdrachten in een bacheloropleiding in het hoger onderwijs. Omdat de precieze tijdsinvestering en opbrengst van TM nog niet bekend was, is TM eerst als prototype aan docenten voorgelegd. Hiermee werd verkend of de toepassing van TM genoeg potentieel heeft om daadwerkelijk te implementeren. Deze manier van het testen van een *prototype* is een goede manier om veilig een vernieuwing te verkennen zonder de beoogde gebruikers onnodig te belasten (Hanington & Martin, 2012). In deze studie ging het daarom om een kwalitatieve verkenning van TM in de vorm van interviews met docenten. Voor de opzet van deze studie is goedkeuring verleend door de ethische commissie van de Universiteit Utrecht, dossiernummer 21-0398.

### Deelnemers

Er is een cursus bij de opleiding onderwijswetenschappen geselecteerd die was opgezet volgens het *flipped classroom* model. Deze cursus bestond uit wekelijkse digitale lessen en literatuur die zelfstandig bestudeerd dienden te worden, een huiswerkopdracht en een werkgroep waarin het cursusmateriaal en de huiswerkopdracht werden besproken. In de huiswerkopdrachten werden studenten gevraagd te reflecteren op de literatuur van de desbetreffende cursusweek, bijvoorbeeld door te reageren op een stelling. De cursus werd begeleid door een team van drie docenten (één vrouw en twee mannen), elk met een werkgroep bestaande uit ongeveer 20 studenten. De docenten hadden 2, 6 en 10 jaar ervaring met lesgeven. Eén van de docenten gaf dit specifieke vak voor de eerste keer, de andere twee docenten hadden het vak vijf keer eerder gegeven. Na afloop van het vak werden de huiswerkopdrachten uit één van de cursusweken gebruikt om TM op toe te passen. De docenten werden kort na afloop van de cursus uitgenodigd voor het interview. De docenten kregen een cadeaubon als waardering voor hun deelname.

### Instrumenten

De dataverzameling vond plaats middels interviews, waarbij een gestructureerde interviewleidraad werd gebruikt (Baarda & Van der Hulst, 2017). Dat betekent dat de onderwerpen en de te stellen vragen van te voren opgesteld waren. De interviewleidraad bestond uit vier onderdelen. Eerst werden docenten gevraagd naar hun indruk van de rol van huiswerkopdrachten in de cursus; welke rol hadden de huiswerkopdrachten volgens

hen voor de student en voor de docent. In onderdeel 2 werd uitleg gegeven over wat TM is en werden de uitkomsten van de TM getoond op drie manieren (zie onder). Na elk van de drie resultaten werd aan de docenten gevraagd om hun reactie te geven op de uitkomsten en of ze dachten dat deze uitkomsten nuttig zouden kunnen zijn voor hun onderwijs. Naast de uitkomsten van de TM werden docenten in onderdeel 3 ook gevraagd twee volledige huiswerkopdrachten te lezen die we willekeurig hadden gekozen uit alle 60 ingeleverde opdrachten. Daarna werden de docenten gevraagd te reageren op wat ze uit deze opdrachten konden opmaken en wat ze hiermee zouden kunnen in hun onderwijs. Tenslotte werd de docenten in onderdeel 4 gevraagd hun eindoordeel te geven over TM en of zij dachten dat dit een waardevolle toepassing zou kunnen zijn in het hoger onderwijs.

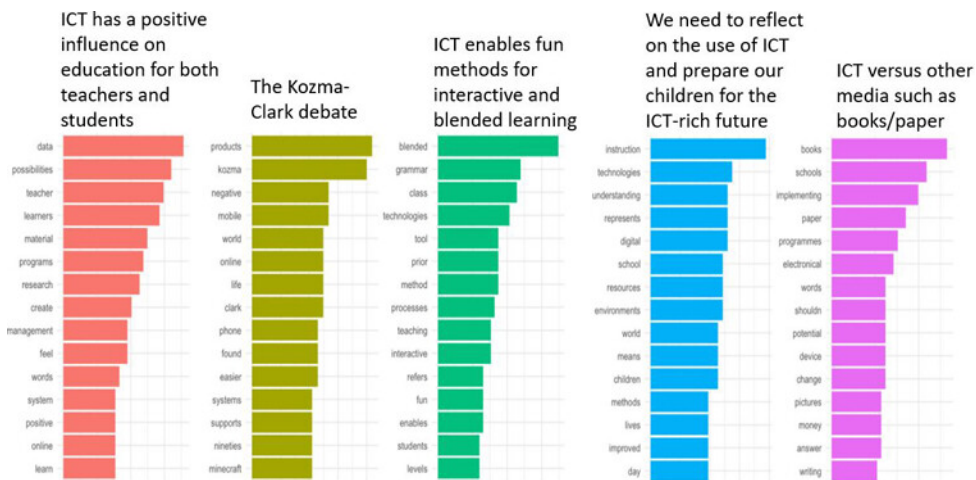
Om validiteit van de resultaten te vergroten werd tijdens de interviews na elk onderdeel steeds gecheckt of de interviewer het goed begrepen had door een samenvatting van het gezegde te geven en de geïnterviewde te vragen of dit klopte. Ook hadden de docenten aan het einde van het interview de ruimte om eventuele zaken te benoemen die zij belangrijk vonden, maar die nog niet aan de orde waren gekomen.

De interviews werden afgenomen door de eerste auteur. De interviews duurden een uur en vonden online plaats via Teams. Er werd een opname gemaakt van het interview, die werd verwijderd nadat het interview getranscribeerd was.

### Achtergrond Topic Modelling

De *Topic Modelling* (TM) analyse werd uitgevoerd door het toepassen van *Latent Dirichlet Allocation* (LDA; Blei, Ng & Jordan, 2003) op de opgeschoonde versie van één van de in het Engels geschreven huiswerkopdrachten uit de cursus van 48 studenten. Hiervoor zijn de opdrachten ontdaan van getallen en andere niet-alfabetische karakters en werden hoofdletters vervangen door kleine letters. Daarnaast zijn stopwoorden die voorkomen in de “stop words” data uit het R-pakket “tidytext” verwijderd, evenals een aanvullende set van acht termen<sup>1</sup> (de lijst met stopwoorden bevat termen die relatief vaak voorkomen maar weinig inhoudelijke waarde hebben, zoals *i*, *we*, *is*, *any* en *very*). De overgebleven tekst werd gewogen middels *term frequency-inverse document frequency* (*tf-idf*), wat inhoudt dat termen gewogen worden gebaseerd op hoe vaak ze voor komen per opdracht per student, maar gecorrigeerd voor hoe vaak ze bij alle studenten samen voorkomen.

*LDA* is een methode om observaties (documenten) te clusteren in latente klassen (*topics*), gebaseerd op de gelijksoortigheid van de documenten. *LDA* beschouwt elk document als een samenstelling van meerdere *topics*, en modelleert de inhoudelijke overlap van de documenten aan de hand van het voorkomen van de *topics*. Elk *topic* wordt gemodelleerd als een samenstelling van verschillende termen die relatief vaak samen komen binnen de individuele documenten. De output van een *LDA*-model kan gebruikt worden om inzicht te genereren in drie verschillende componenten: 1) de verschillende, overkoeplende *topics* die in de opdrachten voorkomen, 2) de inhoudelijke gelijksoortigheid van



**Figuur 2a** De weergave van de uitkomsten van het identificeren van clusters van *topics*

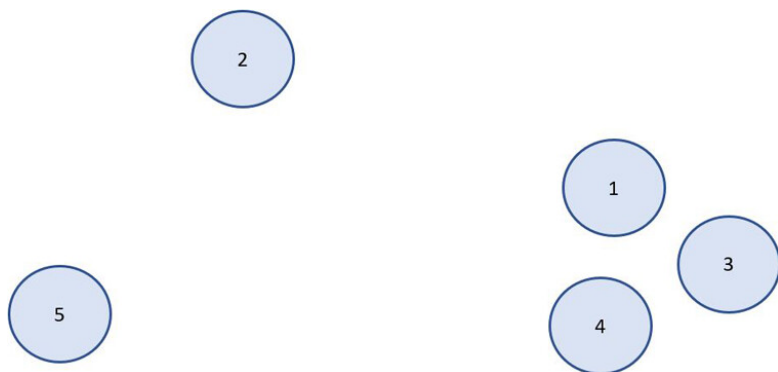
de geïdentificeerde *topics* en 3) het identificeren van de *topics* in individuele opdrachten. Deze drie componenten werden aan de docenten getoond, zie de figuren 2a t/m 2c.

**Identificeren van clusters van *topics***

Een *LDA*-model schat gelijktijdig hoe de *topics* zijn opgebouwd aan de hand van de woorden die dit *topic* kenmerken, alsmede welke *topics* kenmerkend zijn voor ieder document. De *topics* worden opgebouwd aan de hand van de woorden die (relatief) vaak samen voorkomen binnen documenten. De opdrachten zijn geanalyseerd met *LDA* zoals geïmplementeerd in het R-pakket “textmineR” (Jones, 2021), waarbij de resultaten van modellen met 2, 5, 10, 15, 20, 25, 30, 35 en 40 *topics* werden vergeleken.<sup>2</sup> Gebaseerd op inhoudelijke en empirische gronden (de zogenoemde *perplexity*-maten van de geëvalueerde modellen, zie Blei et al., 2003) werd de oplossing met vijf *topics* gekozen als beste representatie van de data, waarna de tien woorden uitgelicht werden die het meest kenmerkend voor dit *topic* zijn. Op deze 5 clusters zijn labels geplakt door de kernwoorden te bekijken, de kernwoorden tussen clusters te vergelijken, en de huiswerkopdrachten te herlezen. Twee onderzoekers (de eerste en vierde auteur) hebben eerst afzonderlijk labels gemaakt, en daarna hun keuze met elkaar besproken. Uit deze discussie kwamen de uiteindelijke labels voort per cluster. In de visualisatie die aan docenten is getoond, werden de rijen kernwoorden per cluster weergegeven met daarboven de labels (Figuur 2a)

**Identificeren van semantische relatie tussen *topics***

De output van het *LDA*-model kan daarnaast gebruikt worden om de semantische afstand tussen *topics* uit te drukken. Iedere term heeft een bepaalde kans om tot elk *topic* te behoren. Wanneer termen met een relatief hoge waarschijnlijkheid in dezelfde *topics*



**Figuur 2b** De weergave van de uitkomsten van het identificeren van semantische relaties tussen *topics*

voorkomen, is de afstand tussen de *topics* kleiner. De afstand tussen *topics* is berekend middels de *Jensen-Shannon divergence* (Zhang et al., 2018), waarna de afstanden tussen *topics* worden terug geschaald naar twee dimensies middels multidimensionaal schalen, en gevisualiseerd. Dit hele proces werd uitgevoerd met het R-pakket LDavis (Sievert & Shirley, 2014). Het uiteindelijke resultaat van deze analyse was een visualisatie van de 5 *topics* als 5 bollen, waarbij de afstand tussen de bollen aangeeft hoeveel de *topics* semantisch gezien van elkaar verschillen (Figuur 2b).

#### Identificeren van *topics* per individuele opdracht

Ten slotte beschouwt *LDA* ieder document als een mix van verschillende *topics*, gebaseerd op de termen in dit document die met een bepaalde waarschijnlijkheid tot ieder *topic* behoren. Gebaseerd op de termen in ieder document kan uitgedrukt worden in welke mate een document tot een *topic* behoort. Het resultaat van deze analyse was een tabel met in de kolommen de 5 *topics*, in de rijen de individuele opdrachten, en in de cellen de verdeling van de *topics* over de opdracht in percentages (Figuur 2c).

#### Analyse

De interviews werden getranscribeerd in tabelvorm, waarbij elk van de onderwerpen in het interview een aparte rij voorstelde en de drie docenten de kolommen vormden. Vervolgens werd een gestructureerde analyse toegepast om de overkoepelende thema's en ideeën in de interviews te identificeren (Miles & Huberman, 1994). Aangezien de interviews een gestructureerd karakter hadden, was het in de analyse mogelijk om dezelfde structuur aan te houden. Voor elke docent werd voor elk onderwerp uit het interview een korte samenvatting gemaakt in termen van een kernzin of kernthema (de zogenaamde *themacodes*, zie Wester & Peters, 2004). Deze samenvattingen werden op dezelfde manier als de transcripties in tabelvorm geplaatst, wat leidde tot een *case-ordered matrix* (Miles & Huberman, 1994). Om de betrouwbaarheid te garanderen werd dit proces door twee



Student	Topic 1	Topic 2	Topic 3	Topic 4	Topic 5
1	0%	<b>80%</b>	<b>20%</b>	0%	0%
2	0%	0%	0%	<b>100%</b>	0%
3	0%	<b>100%</b>	0%	0%	0%
...					
25	0%	0%	<b>15%</b>	0%	<b>85%</b>
26	<b>100%</b>	0%	0%	0%	0%
27	0%	<b>10%</b>	<b>90%</b>	0%	0%

**Figuur 2c** De weergave van de uitkomsten van het identificeren van *topics* per individuele opdracht

onderzoekers (de eerste en vierde auteur) afzonderlijk doorlopen. Daarna werden de samenvattingen vergeleken en waar nodig bediscussieerd om tot een gedeelde visie te komen. Omdat het een verkennende studie betrof met een kleine sample van 3 docenten, was het samen doorlopen van de volledige *dataset* haalbaar.

## Resultaten

Tabel 1 geeft de resultaten weer van de analyse van de interviews. Het betreft een case-ordered matrix waarbij de kolommen de docenten representeren, en de rijen de onderwerpen uit het interview. De cellen geven de samenvatting weer van de reactie van de desbetreffende docent op het desbetreffende thema. Hieronder gaan we verder in op elk van de onderwerpen.

*Reflectie op inzet huiswerkopdrachten.* Alle docenten gaven aan dat de huiswerkopdrachten hoofdzakelijk bedoeld waren als voorbereiding voor de student op de werkgroep. De docenten keken wel of de opdrachten ingeleverd werden, maar – zoals verwacht – niet altijd naar de inhoud. De reden daarvoor was niet alleen dat het tijdsintensief is om alle opdrachten te bekijken, maar ook dat ze op hun ervaring van voorgaande jaren konden leunen. Twee docenten wisten op basis van eerdere jaargangen van dit vak welke onderwerpen studenten zouden aansnijden en welke concepten studenten moeilijk vinden. Het lezen van de opdrachten heeft dan minder relevantie. Alle docenten gaven wel aan dat ze de huiswerkopdrachten in de werkgroep bespreken. Dat deden ze bijvoorbeeld door studenten tijdens groepswork vragen te stellen, of door de vragen uit de huiswerkopdrachten plenair te bespreken.

*Reactie op analyse 1; de 5 topics.* Bij de eerste analyse die de docenten te zien kregen, hadden zij alle drie vragen over de oorsprong en de validiteit van de labels die op de 5 clusters geplakt waren. Na uitleg over hoe de onderzoekers dit hadden aangepakt,



**Tabel 1** Resultaten van de interviews (rijen) per geïnterviewde docent (kolommen)

Onderwerp	Docent 1	Docent 2	Docent 3
<b>Ervaring als docent</b>	2 jaar ervaring	6 jaar ervaring	10 jaar ervaring
<b>Aantal keer dit vak gegeven</b>	1 <sup>e</sup> keer dit vak	6 <sup>e</sup> keer dit vak	6 <sup>e</sup> keer dit vak
<b>Reflectie op inzet huiswerkopdrachten</b>	Het huiswerk diende als voorbereiding voor de student, ik checkte of het ingeleverd was, niet de inhoud.	Het huiswerk diende als voorbereiding voor de student, ik checkte of het ingeleverd was, niet de inhoud. Tijdens de werkgroep vroeg ik studenten naar hun antwoorden op de huiswerkopgaven.	Het huiswerk diende als voorbereiding voor de student, ik keek er soms naar. Tijdens de werkgroep bespraken we soms plenair de huiswerkopgaven. Ik wist uit ervaring al wat ik ongeveer kon verwachten.
<b>Reactie op analyse 1: de 5 topics</b>	Het is niet duidelijk hoe de labels voor de <i>topics</i> gemaakt zijn en of het tijd zou besparen. De <i>topics</i> lijken te passen bij de opdracht. Inhoudelijk weet ik niet of studenten het snappen, maar het geeft een overzicht van de onderwerpen. De analyse laat zien dat studenten de literatuur hebben bekeken.	Het is niet duidelijk hoe de labels voor de <i>topics</i> gemaakt zijn, of het tijd bespaart, en of ze objectief kloppen. Ik zou de onderwerpen kunnen voorleggen aan studenten als discussieonderwerpen. De analyse zegt niets over individuele studenten, ik zou het alleen gebruiken in het plenaire deel.	Het is niet duidelijk hoe de labels voor de <i>topics</i> gemaakt zijn. Ik herken de onderwerpen van de ingeleverde stukken van vorige jaren. Je kunt de analyses gebruiken om de huiswerkopdracht specifiek te maken.
<b>Reactie op analyse 2: samenhang topics</b>	De uitkomst lijkt me correct. Ik weet niet hoe ik het zou kunnen gebruiken.	De uitkomst lijkt me correct. Het zegt me niets over het begrip van de studenten. Ik zou het als werkvorm kunnen inzetten door deze analyse aan de studenten te tonen.	De uitkomst lijkt me niet helemaal correct. Het zegt me niets over het begrip van de studenten. Ik zou het als werkvorm kunnen inzetten door deze analyse aan de studenten te tonen.

**Tabel 1** Resultaten van de interviews (rijen) per geïnterviewde docent (kolommen) (*vervolg*)

Onderwerp	Docent 1	Docent 2	Docent 3
<b>Reactie op analyse 3: topics per student</b>	Dat een student een onderwerp bespreekt, zegt niets over inhoudelijk begrip. Bij uitschieters zou ik studenten ernaar kunnen vragen.	Ik kan dit resultaat gebruiken om te kiezen welk <i>topic</i> ik plenair bespreek. Ik kan het gebruiken in het gesprek met een specifieke student.	Ik kan dit resultaat gebruiken om te kiezen welk <i>topic</i> ik plenair bespreek en in welke volgorde ik dat doe.
<b>Reactie op een volledige huiswerkopdracht en vergelijking met uitkomsten TM</b>	Met de visualisatie kun je alleen zien dat een student met een onderwerp bezig is. Ik kan uit de tekst duidelijker opmaken hoe de student over een onderwerp denkt. Op beide manieren krijg ik input die ik in de les kan gebruiken.	Met de visualisatie krijg ik een overzicht van alle studenten. Ik kan uit de tekst duidelijker opmaken hoe de student over een onderwerp denkt. Bij Analyse 1 twijfel ik over de betrouwbaarheid. Analyse 2 en 3 zou ik kunnen gebruiken, eventueel samen met een paar volledige teksten van studenten.	Met de visualisatie krijg ik een overzicht van alle studenten. Informatie uit de teksten is meer gedetailleerd. Op beide manieren krijg ik input die ik in de les kan gebruiken. Onervaren docenten zouden de teksten kunnen lezen.
<b>Eindoordeel over TM</b>	Of TM bruikbaar is, is afhankelijk van het doel van de cursus / de huiswerkopdrachten. Ik zou het interessant vinden om de <i>topics</i> met andere docenten te bespreken.	Of TM bruikbaar is, is afhankelijk van het doel van de huiswerkopdrachten. Testen van toepassing van TM kan nuttig zijn.	Ik vind vooral analyse 3 erg interessant. Onderwijs-coördinatoren zouden TM kunnen gebruiken om verdiepvragen op te stellen. Als je huiswerkopdrachten specifiek zijn dan zou je meer met TM kunnen.

vroegen de docenten zich af of dit proces wel tijd zou besparen als het labelen van de clusters ook een aanzienlijke hoeveelheid tijd kost. Ook twijfelden de docenten of deze analyse waardevolle informatie geeft. De docenten gaven aan dat deze analyse een overzicht kan geven van de onderwerpen die studenten noemen, maar niet over het begrip dat de studenten van die onderwerpen hebben. Studenten kunnen een *topic* wel noemen, maar dat betekent niet dat ze het in de juiste context hebben geplaatst en alles hebben

begrepen. Docent 2 gaf aan dat je de visualisatie van de *topics* ook als werkvorm zou kunnen voorleggen aan de studenten en hen daarover zou kunnen laten discussiëren. Docent 3 gaf een geheel andere toepassing van deze analyse aan: je zou op basis van de *topics* de huiswerkopdracht voor het jaar daarop specifieker kunnen maken door in te spelen op de *topics* die studenten niet hebben genoemd maar die wel belangrijk zijn.

*Reactie op analyse 2; samenhang topics.* De tweede analyse gaf weer in hoeverre de 5 gelabelde *topics* semantisch gezien van elkaar verschillen. Twee docenten vonden de uitkomsten van de analyse passen bij hun ideeën over de inhoud van de cursus, maar één docent niet. Alle docenten waren het erover eens dat de analyse niet veel zegt over het begrip van de stof van de studenten. Twee docenten noemden, net als bij de uitkomsten van analyse 1, dat ze deze zouden kunnen inzetten als werkvorm door het aan de studenten te tonen.

*Reactie op analyse 3; topics per student.* De docenten kenden de meeste waarde toe aan de derde analyse, waarin per student wordt aangegeven welke *topics* in hun tekst zijn genoemd. De docenten gaven aan dat ze deze analyse konden gebruiken om individuele studenten aan te spreken op welke *topics* ze wel of niet hadden genoemd. Daarnaast leek het de docenten handig om informatie die aangeeft hoe vaak een *topic* genoemd is te gebruiken om in de werkgroep te bespreken, bijvoorbeeld door eerst in te zoomen op *topics* die relatief weinig genoemd worden.

*Reactie op een volledige huiswerkopdracht en vergelijking met uitkomsten TM.* Na de drie analyses werden docenten ook gevraagd twee volledige huiswerkopdrachten te lezen en hun inzichten te vergelijken met de uitkomsten van TM. De algemene uitkomst was hier dat de TM weliswaar een overzicht geeft van de onderwerpen die studenten noemen, maar dat het lezen van de geschreven opdrachten uiteindelijk meer en meer gedetailleerde informatie geeft over wat studenten precies denken over de onderwerpen.

*Eindoordeel over TM.* Alle docenten geven aan dat ze wel iets zien in het gebruiken van TM, maar niet alleen voor het doel waarvoor wij als onderzoekers het in eerste instantie hadden bedoeld. De docenten noemden bijvoorbeeld dat ze de uitkomsten van TM ook zouden willen bespreken in het docententeam van het vak, om zo met elkaar af te stemmen wat je met de uitkomsten kan en of het nuttige informatie is om mee te nemen in het ontwerp van de huiswerkopdrachten van het jaar erop.

## Discussie

### *Bespreking van bevindingen*

In dit onderzoek keken we naar de percepties van docenten omtrent de mogelijke toepassing van *Topic Modelling* (TM) in het hoger onderwijs om inzicht te geven in de ingeleverde huiswerkopdrachten van studenten. De studie werd specifiek uitgevoerd bij een cursus met een flipped classroom design, waarbij de huiswerkopdrachten een voorbereiding zijn op de verdiepende discussies in de werkgroepbijeenkomsten (Staker & Horn, 2012).

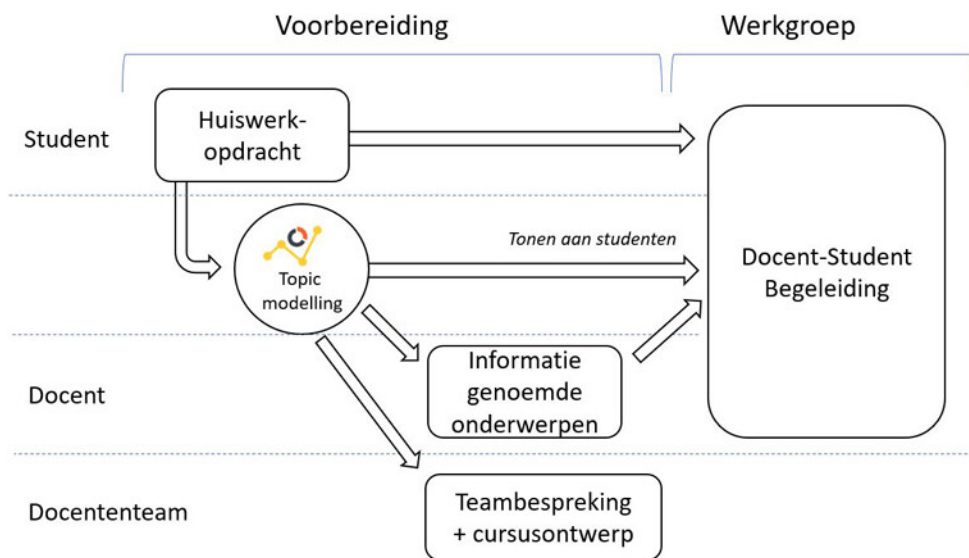
Het verschilde per docent of en hoe zij de uitkomsten van TM zouden kunnen en willen gebruiken in hun onderwijs. De perceptie van de docenten bleek samen te hangen met hoeveel ervaring zij al hadden met het specifieke vak waarin we TM hadden uitgevoerd. Docenten doen in de loop van de tijd ervaring op met welke concepten studenten moeilijk vinden. Daardoor hebben zij minder aan een overzicht van onderwerpen, en juist meer aan gegevens over hoe vaak en door wie een bepaald onderwerp genoemd wordt. Dat geeft hen *actionable* informatie omdat ze op specifieke onderwerpen in zouden kunnen spelen in de werkgroep en specifieke studenten zouden kunnen begeleiden, waardoor ze adaptief onderwijs kunnen aanbieden (Vogt & Rogalla, 2009). Voor minder ervaren docenten kan het overzicht van onderwerpen juist wél van waarde zijn om voor te bereiden op eventuele vragen van studenten, en om deze onderwerpen te bespreken met het gehele docententeam. In vergelijking met de eerdere studie waarin vooral informatie werd gegeven over het wel of niet voltooiën van activiteiten (Van Leeuwen, 2019), biedt TM zeker voor minder ervaren docenten dus meer inhoudelijke en meer actionable inzichten op.

Wel moeten we realistisch zijn in de afweging tussen de tijdsinvestering en de opbrengsten van TM, zeker gegeven de werkdruk die docenten ervaren (Universiteiten van Nederland, 2022). Het uitvoeren van TM en het labelen van de clusters vraagt kostbare tijd, en de opbrengsten van TM zoals die uit deze studie naar voren kwamen waren beperkt. Zeker bij de meer ervaren docenten was de algemene uitkomst dat TM informatief is, maar zeker niet noodzakelijk om een goede werkgroep voor te bereiden.

De docenten in dit onderzoek noemden toepassingen van TM die verder gingen dan de door ons voorziene toepassing, namelijk het tonen van de uitkomsten van TM aan docenten. De docenten in deze studie gaven aan dat er nog twee mogelijke toepassingen zijn om verder te verkennen, zie Figuur 3. Allereerst kunnen de uitkomsten van TM direct als werkvorm ingezet worden door de visualisaties tijdens de werkgroepen aan studenten te tonen en hen hierover te laten discussiëren. Daarnaast werd door de docenten genoemd dat de TM input kan zijn voor het wekelijkse overleg van het docententeam. Daarin kan besproken worden hoe en of aanpassingen aan de werkgroep wenselijk zijn en welke moeilijkheden studenten eventueel hebben ervaren met het cursusmateriaal. Dit zijn toepassingen van TM die in vervolgstudies verder onderzocht kunnen worden.

### *Beperkingen van deze studie*

Een duidelijke beperking van dit onderzoek is de kleinschaligheid van de uitgevoerde studie. TM is ingezet bij een specifieke opdracht in een specifieke cursus, met een klein docententeam. Dat betekent dat de resultaten waarschijnlijk erg contextgebonden zijn en dat toepassing van TM in een andere context mogelijk andere kansen en uitdagingen met zich meebrengt. Deze studie moet daarom, zoals eerder besproken, vooral gezien worden als een eerste verkenning van de mogelijke waarde van TM in het Hoger Onderwijs bij het analyseren van schrijfpdrachten.



**Figuur 3** In de interviews genoemde toepassingen van *topic modelling*

### *Implicaties en conclusie*

Opleidingen die erover denken om TM in te zetten, kunnen we een aantal tips meegeven. De inzet van TM is een analyse die nog een handmatige verwerking vereist, namelijk het bepalen van de labels die bij de onderwerpen-clusters horen. Deze tijdsinvestering zouden wij alleen maken als er binnen het docententeam ruimte en draagvlak is om de toepassing van TM te verkennen. Als besloten is om TM in te zetten, is het goed om de huiswerkopdrachten waarop TM zal worden toegepast, aan te passen zodat de uitkomsten van de TM eenduidiger en makkelijker te interpreteren zijn. Het is dan raadzaam dat de teksten over één onderwerp gaan. Een andere belangrijke voorbereidende stap is om de tijdsvolgorde van de uit te voeren stappen van TM vooraf goed in kaart te brengen. Wanneer leveren studenten hun opdracht in, wanneer wordt TM uitgevoerd, wanneer worden de labels op de clusters bepaald, en wanneer krijgen de docenten toegang tot de clusters en labels? Ook is het belangrijk dat de docenten inzicht hebben in hoe de labels tot stand zijn gekomen. Het is daarom goed om af te spreken wie de labels zal vaststellen; bij voorkeur gebeurt dit in een duo, bijvoorbeeld door de coördinatoren van het vak. De laatste tip is om te overwegen om de resultaten van TM in het docententeam van het vak te bespreken zodat docenten met elkaar kunnen uitwisselen wat ze met de uitkomsten zouden kunnen doen.

De conclusie van deze studie is dat docenten potentie zien in TM. Een vervolgstap zou kunnen zijn om van het bekijken van een prototype (zoals in deze studie) te gaan naar het uitproberen van TM in een lopende cursus. Daarmee zou meer inzicht opgedaan kunnen worden in de praktische uitvoering en meerwaarde van TM.

## Noten

- 1 Behalve de ingebouwde lijst met stopwoorden, zijn ook de termen *makes*, *question*, *shouldnt*, *whats*, *wich*, *lot*, *couldnt* en *addition* verwijderd.
- 2 Hiervoor gebruikten we de functie “FitLdaModel()”, met specificaties “iterations = 3000” en “burnin = 1000”.

## Literatuur

- Baarda, B., & van der Hulst, M. (2017). *Basisboek interviewen*. Noordhoff Uitgevers.
- Blei, D.M., Ng, A.Y., & Jordan, M.I. (2003). Latent Dirichlet allocation. *Journal of Machine Learning Research*, 3, 993–1022.
- Chen, Y., Yu, B., Zhang, X., & Yu, Y. (2016). Topic modeling for evaluating students’ reflective writing: A case study of pre-service teachers’ journals. *Proceedings of the sixth international conference on learning analytics & knowledge*, 1–5. <https://dl.acm.org/doi/10.1145/2883851.2883951>
- Hanington, B., & Martin, B. (2012). *Universal methods of design: 100 ways to research complex problems, develop innovative ideas, and design effective solutions* [E-book]. Rockport Publishers.
- Jones, T. (2021). *CRAN – Package textmineR*. textmineR: Functions for Text Mining and Topic Modeling. Geraadpleegd op 19 juni 2022, van <https://cran.r-project.org/web/packages/textmineR/index.html>
- Kaliisa, R., Kluge, A., & Mørch, A.I. (2020). Combining checkpoint and process learning analytics to support learning design decisions in blended learning environments. *Journal of Learning Analytics*, 7(3), 33–47. <https://doi.org/10.18608/jla.2020.73.4>
- Miles, M.B., & Huberman, M.A. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd edition). SAGE Publications, Inc.
- Sievert, C., & Shirley, K.E. (2014). LDAvis: A method for visualizing and interpreting topics. In Association for Computational Linguistics (Red.), *Proceedings of the Workshop on Interactive Language Learning* (pp. 63–70). Association for Computational Linguistics.
- Staker, H., & Horn, M.B. (2012). *Classifying K–12 blended learning*. Innosight institute.
- Universiteiten van Nederland (2022). Werkdruk. <https://www.universiteitenvannederland.nl/werkdruk.html>
- Van Alten, D.C.D., Phielix, C., Janssen, J., & Kester, L. (2019). Effects of flipping the classroom on learning outcomes and satisfaction: A meta-analysis. *Educational Research Review*, 28(June), 1–18. <https://doi.org/10.1016/j.edurev.2019.05.003>
- Van Leeuwen, A. (2019). Teachers’ perceptions of the usability of learning analytics reports in a flipped university course: when and how does information become actionable knowledge? *Educational Technology Research and Development*, 67(5), 1043–1064. <https://doi.org/10.1007/s11423-018-09639-y>
- Vogt, F., & Rogalla, M. (2009). Developing adaptive teaching competency through coaching. *Teaching and Teacher Education*, 25(8), 1051–1060.

- Wester, F., & Peters, V. (2004). Kwalitatieve analyse: uitgangspunten en procedures (pp. 75–103). Bussum: Coutinho.
- Zhang, C., Zheng, J., Wang, Z., Jia, Z., & Li, F. (2018). A New Evaluation Criterion with the Integration of Perplexity and Jensen-Shannon Divergence for Biterm Topic Model. *IEEE proceedings*. <https://doi.org/10.1109/IMCEC.2018.8469586>