

# Een door onderzoek geïnformeerde ontwikkeling van een Learning Analytics dashboard: de casus van het LAP!-project

Ellen De Bruyne, Laura Herrewijn, Amber Hoefkens, Wannes Heirman,  
Gert Vanthournout, Pieter Depessemier

**Samenvatting** Learning Analytics (LA) zijn begaan met het gebruik van data om het leerproces van studenten te begrijpen, hun leeromgeving te optimaliseren of door data-geïnformeerde innovatie mogelijk te maken. Kwalitatieve dashboards zijn cruciaal om deze doelen te bereiken. Het praktischgericht wetenschappelijk onderzoek LAP! tracht op een door onderzoek geïnformeerde manier *dashboards* voor verschillende gebruikersgroepen te ontwikkelen en gebruikt daarbij *educational design research* als raamwerk. Deze bijdrage illustreert de analyse- en exploratiefase in dit project. Hierin werd onderzoek opgezet naar de inhoud, structuur, lay-out en relevantie van deze *dashboards* aan de hand van drie acties: een literatuurverkenning; een gebruikersonderzoek bij studenten, lesgevers en opleidingshoofden; en een databankanalyse die met behulp van *machine learning* naging welke beschikbare variabelen een voorspellende kracht hebben op studiesucces. Verder werd ook de juridische en wettelijke basis van de verwerking van persoonsgegevens in de *dashboards* in kaart gebracht en geëvalueerd via een *Data Protection Impact Assessment* (DPIA). De resultaten van de drie onderzoeksacties inzake de selectie van variabelen zijn niet steeds eenduidig, maar bieden wel voldoende houvast voor de ontwikkeling van een prototype voor een generiek *dashboard*. Via cycli van ontwikkeling, implementatie en testen zal dit prototype worden verfijnd en worden aangepast aan de noden van verschillende gebruikersgroepen.

**Trefwoorden** Educational Design Research, gebruikersonderzoek, machine learning, GDPR

## Artikelgeschiedenis

Ontvangen: 30 april 2021  
Geaccepteerd: 16 februari 2022  
Online: 31 maart 2023

## Contactpersoon

Ellen De Bruyne,  
ellen.debruyne@ap.be

## Over de Auteur(s)

Ellen De Bruyne, Laura Herrewijn, Amber Hoefkens, Wannes Heirman en Gert Vanthournout zijn allen werkzaam bij AP Hogeschool Antwerpen, Innovation Lab Data gedreven Leren en Innovatie; Pieter Depessemier is werkzaam bij AP Hogeschool Antwerpen, Directie Onderwijs en Onderzoek, Cel Onderwijs en Innovatie

## 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.

## Introductie

Vandaag de dag worden grote hoeveelheden data digitaal vergaard en automatisch opgeslagen, ook over studenten en hun leerproces (Sclater et al., 2016). Dit betreft onder meer informatie die wordt verzameld bij digitale inschrijvingen, activiteiten zoals klikgedrag van studenten op de digitale leeromgeving of scores die worden ingegeven in een digitaal puntenboek. Learning Analytics (LA) trachten dergelijke data te ontsluiten met als doel het leerproces van studenten te begrijpen, hun leeromgeving te optimaliseren of door data geïnformeerde innovatie mogelijk te maken (Siemens & Gasevic, 2012).

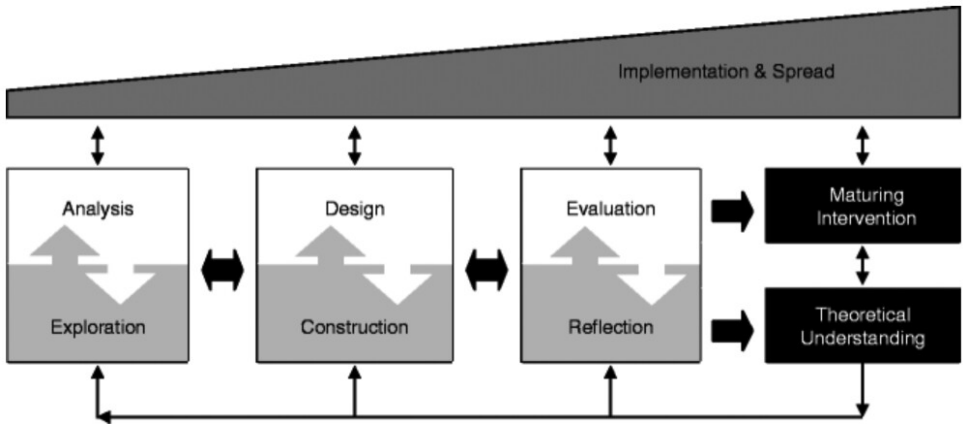
Inzichten uit LA kunnen instellingen voor hoger onderwijs helpen om in te spelen op uitdagingen zoals het identificeren en begeleiden van risico-studenten (Gasevic et al., 2017), het effectiever inrichten van (*blended*) leeromgevingen (Avella et al., 2016) of het bepalen van beleidsprioriteiten van opleidingen of departementen. Daartoe is het noodzakelijk informatie uit LA op een toegankelijke manier aan te bieden aan gebruikers (groepen). Vaak gebeurt dit via ‘dashboards’ waarin informatie multimediaal wordt weergegeven. Kwaliteitsvolle dashboards zijn een cruciale eerste stap om met LA aan de slag te gaan (Verbert, Duval, Klerkx, Govaerts, & Santos, 2013). Een dashboard verzamelt en visualiseert doorgaans gebruikerssporen van leeractiviteiten om bewustwording, reflectie, en zingeving te creëren en gebruikers in staat te stellen (nieuwe) doelen te definiëren en de voortgang naar deze doelen op te volgen (Verbert, Govaerts, Santos, Van Assche, Parra & Klerkx, 2014). Hoewel LA hoofdzakelijk gebaseerd zijn op data die worden verzameld bij studenten, zijn de inzichten die eruit voortvloeien niet enkel voor hen relevant. Ook lesgevers, studiebegeleiders en beleidsmakers zijn potentiële gebruikersgroepen van LA.

LA zijn een relatief nieuw onderzoeksveld (Papamitsiou & Economides, 2014). Op dit ogenblik zijn er nog onvoldoende ontwerprichtlijnen voor de inhoud, de structuur en de vormgeving van LA-dashboards in hoger onderwijs. Dashboards worden voornamelijk ontwikkeld door onderzoekers, commerciële organisaties en ondersteunende diensten (Buckingham Shum et al., 2019). Over het algemeen gedragen die ontwikkelaars zich als passieve gebruikers, waardoor de voorkeuren betreffende informatie en de representaties bij de toekomstige gebruikers niet volledig worden geëxploreerd. Broos, Pinxten, Delporte, Verbert en De Laet (2020) benadrukken immers het belang om dashboards te ontwikkelen met de betrokkenheid van gebruikers, zodat dashboards worden ontwikkeld op basis van hun specifieke behoeften.

Aan AP Hogeschool Antwerpen (AP) werd daarom een multidisciplinair project voor praktijkgericht wetenschappelijk onderzoek opgezet, LAP!, met als doel om op basis van onderzoek (met gebruikers) prototypes voor LA-dashboards te ontwikkelen. Als methodologisch raamwerk fungeert het model van *educational design research* (McKenney & Reeves, 2014).

### Context: Het LAP!-project of Learning Analytics op de AP Hogeschool Antwerpen

Net zoals bij andere instellingen voor hoger onderwijs is op AP informatie beschikbaar in verschillende databanken, onder meer wat betreft studentenadministratie, studieresultaten, gebruik van de digitale leeromgeving of resultaten voor de Leerstrategieën en Motivatievragenlijst (LEMO; Vanthourmout et al., 2016). Informatie uit een achttal databanken werd geanalyseerd en gerapporteerd in aparte rapporten waardoor de LA informatie voor gebruikers onderbelicht bleven. Het project voor Praktijkgericht Wetenschappelijk Onderzoek LAP!: *evidence-informed* ontwikkeling van een Learning Analytics



**Figuur 1** Model voor de aanpak van Educational Design Research (McKenney & Reeves, 2014)

*dashboard op de AP Hogeschool* geeft een aanzet tot een meer systematisch gebruik van LA op AP via het ontwerpen van kwaliteitsvolle dashboards voor drie gebruikersgroepen: studenten, lesgevers en opleidingshoofden. In lijn met suggesties uit de literatuur werd gekozen voor een multidisciplinaire aanpak: het combineert aanwezige expertise op AP inzake onderwijskundige, technologische en juridische aspecten van LA met expertise inzake datavisualisatie, *Artificial Intelligence* en gebruikerservaringen. Doorheen de vier projectjaren (2019–2023) wil het LAP!-project adaptieve *dashboards* ontwikkelen die de informatie presenteren op maat van de verschillende gebruikersgroepen. De ontwikkeling van deze *dashboards* op maat gebeurt via een cyclisch ontwikkelproces waarbij een generiek prototype geleidelijk aan wordt aangepast en verfijnd voor de verschillende gebruikersgroepen. Daarbij wordt de methodiek van educational design research gevolgd (McKenney & Reeves, 2014). De huidige bijdrage richt zich op de eerste ontwerpcyclus in het project: het ontwerpen van een generiek *dashboard*.

### Ontwerpkader: Educational Design Research

Voor het ontwerpen van de *dashboards* wordt de methode van *educational design research* gevolgd (Figuur 1; McKenney & Reeves, 2014). In onderwijskundig ontwerponderzoek is dit een generiek model. Deze gebruikers-gecentreerde methode heeft een dubbele focus op theorie en praktijk, kenmerkt een integrale onderzoek-ontwikkeling benadering en leidt tot theoretische en praktische uitkomsten die toepasbaar zijn voor andere settings. Drie iteratieve kernfasen zijn essentieel in *educational design research*: analyse en exploratie, ontwerp en ontwikkeling, en evaluatie en reflectie.

In het LAP!-project werd tussen september 2019 en januari 2020 de ‘Analyse & Exploratiefase’ afgewerkt. De voornaamste aandachtspunten in deze fase waren de inhoud,

structuur en de lay-out van de te ontwikkelen *dashboards*. Daarnaast werden vragen beantwoord naar de juridische randvoorwaarden voor het ontwerpen van een *LA-dashboard*. Tussen oktober 2020 en april 2021 werd ook een eerste cyclus in de fase 'Ontwerp & Ontwikkeling' afgerond. Daarin werd een prototype voor een generiek rapport ontwikkeld.

### Onderzoeksvragen

Het LAP!-project wil *dashboards* ontwikkelen voor drie gebruikersgroepen. Ter voorbereiding van die ontwikkeling werden verkennende onderzoeksacties opgezet. Volgende onderzoeksvragen waren daarbij richtinggevend:

- **OV1:** Welke richtlijnen en adviezen inzake relevante inhoud voor een *LA-dashboard* schuift de bestaande literatuur naar voren?
- **OV2:** Welke noden inzake inhoud en vormgeving van een *LA-dashboard* ervaren de gebruikersgroepen studenten, lesgevers en opleidingshoofden? Waar zitten gelijkenissen en verschillen?
- **OV3:** Welke variabelen in bestaande databanken op AP zijn voorspellend voor het slaagcijfer van studenten?
- **OV4:** Welke juridische en technologische uitdagingen verhinderen de ontwikkeling van een generiek *LA-dashboard*?

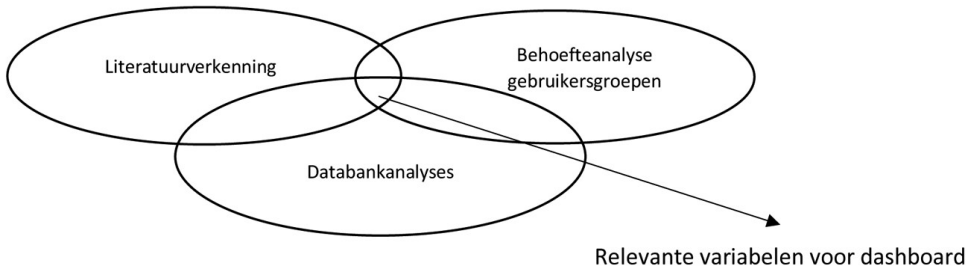
### De analyse en exploratiefase voor het ontwerpen van *LA-dashboards*

#### Aanpak

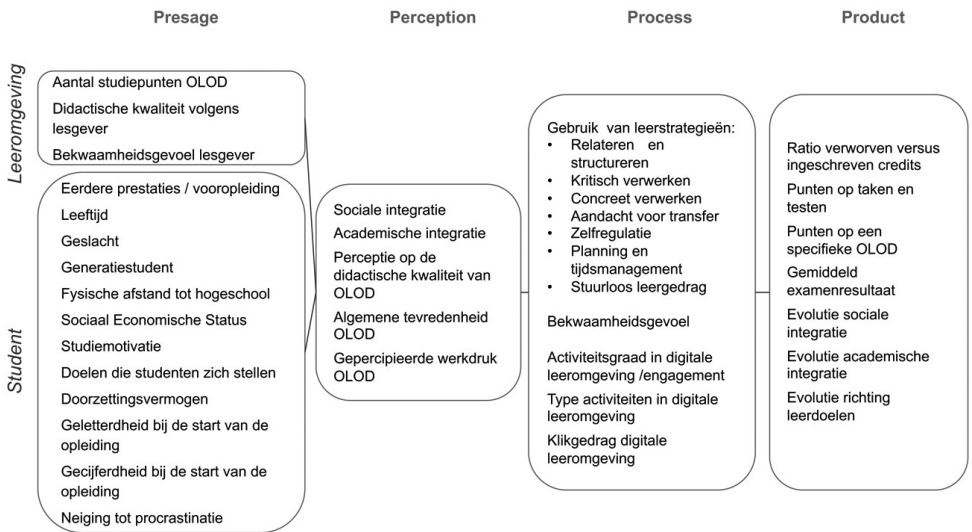
Om de onderzoeksvragen uit de analyse en exploratiefase te beantwoorden werden drie onderzoeksacties opgezet: (1) een literatuurverkenning (OV1), (2) een behoefteanalyse bij verschillende gebruikersgroepen via focusgroepgesprekken en *survey's* (OV2) en (3) een databankanalyse naar voorspellers voor studiesucces in de beschikbare data op AP (OV3). De resultaten van deze drie onderzoeksacties werden naast elkaar gelegd met als doel een doordachte en relevante selectie van variabelen te maken die in het prototype van het generieke *dashboard* kunnen worden opgenomen (Figuur 2). De resultaten van de focusgroepgesprekken werden niet enkel gebruikt om de inhoud van de *dashboards* te bepalen, maar inspireerden ook mee de lay-out van het prototype.

#### Literatuurverkenning

Via een literatuurverkenning werd een inventaris gemaakt van potentiële variabelen die deel uit kunnen maken van een *LA-dashboard* (OV1). Daarbij werden op de eerste plaats reviewstudies en meta-analyse studies geconsulteerd die voorspellers van studiesucces of andere gewenste uitkomstmaten in het hoger onderwijs in kaart brachten (Baeten et



**Figuur 2** Onderzoeksacties i.f.v. selectie van relevante variabelen



**Figuur 3** Overzicht van door de literatuur relevant geachte variabelen die beschikbaar zijn in de databanken van AP

*Noot:* OLOD = opleidingsonderdeel

al., 2010; Hattie, 2015; Price, 2013; Richardson et al., 2012; van Rooij et al., 2018). De variabelen werden geïnventariseerd en gestructureerd met behulp van het 4P-model (Gijbels, Donche, Richardson & Vermunt, 2013; Price, 2013; Price & Richardson, 2004; Biggs, 2003). Vervolgens werd voor deze inventaris nagegaan welke variabelen in de databanken op AP aanwezig waren. Het resultaat van deze oefening is weergegeven in Figuur 3.

**Behoefteanalyse bij gebruikersgroepen**

Om de noden van gebruikersgroepen in kaart te brengen en ov2 te kunnen beantwoorden werden er in mei 2020 acht focusgroepen afgenomen. Er werden vier gesprekken georganiseerd met lesgevers (N = 16) en vier gesprekken met studenten (N = 15), waarbij

respondenten telkens werden gerekruteerd uit verschillende departementen van AP. Er werd voor de focusgroepen een semigestructureerde vragenlijst opgesteld. Daarnaast werd er ook ruimte gelaten om bepaalde antwoorden van de respondenten verder uit te diepen door middel van het stellen van bijvragen. Er werden 'mock-ups' en voorbeelden van bestaande *dashboards* en feedbackrapporten gebruikt om de input van gebruikers zo concreet mogelijk te krijgen. Omwille van de geldende COVID-maatregelen gebeurde de afname via Microsoft Teams. De gesprekken werden met toestemming van de respondenten opgenomen en ad verbatim uitgetypt. De gesprekken werden vervolgens gecodeerd en geanalyseerd in Excel.

In een tweede fase van het onderzoek werd er ook een *online survey* afgenomen bij studenten ( $N = 117$ ) en lesgevers ( $N = 97$ ) van AP in juni 2020, alsook bij opleidingshoofden ( $N = 24$ ) van AP in september 2020. Deze *online survey's* werden opgesteld en verspreid via Limesurvey (voor de studenten en lesgevers) en Qualtrics (voor de opleidingshoofden). De resulterende data werden achteraf geanalyseerd in SPSS.

In zowel de focusgroepen als de online survey werd gepeild naar hoe respondenten de relevantie van *LA-dashboards* opvatten, in welke inhoudelijke informatie uit *LA* ze interesse zouden hebben, en wat hun voorkeur zou zijn naar lay-out, hoeveelheid informatie en manier van implementatie toe.

De resultaten van de focusgroepen en online survey tonen aan dat alle gebruikersgroepen *de relevantie* zien van de implementatie van *LA-dashboards*. Uit de focusgroepen blijkt dat zowel lesgevers als studenten de *dashboards* beschouwen als een manier om zichzelf op te volgen en bij te sturen. Lesgevers willen vooral hun stijl van lesgeven monitoren en de impact die dit heeft op de resultaten van hun studenten. Studenten wensen voornamelijk om hun studieresultaten en hun vooruitgang te kunnen opvolgen. Lesgevers benadrukken ook dat ze een *dashboard* wensen waarbij het genereren en opvolgen van de data niet te ingewikkeld is, zodat ze hier zelf geen extra tijd en moeite in moeten investeren. Tevens vrezen ze dat het gebruik van *dashboards* zal worden gekoppeld aan de verwachting om meer op maat van individuele studenten te werken. Studenten geven in dit opzicht aan dat ze *LA-dashboards* ook nadelig zouden percipiëren indien lesgevers alle informatie van individuele studenten kunnen volgen.

Wat de *inhoud van de LA-dashboards* betreft, tonen de resultaten van de focusgroep-gesprekken en online survey dat lesgevers op de eerste plaats interesse hebben in de prestaties van studenten op taken en examens. Daarnaast willen ze weten hoe studenten de digitale leeromgeving van hun opleidingsonderdelen gebruiken en hoe studenten de kwaliteit van de leeromgeving en het leermateriaal percipiëren. Verder zijn ze ook geïnteresseerd in het leerstrategiegebruik en de studiemotivatie van studenten en willen ze inzicht verwerven in een aantal studentenkenmerken waaronder gecijferdheid, geletterdheid en vooropleiding. Lesgevers krijgen ook graag zicht op de algemene werkdruk van de studenten. Ten slotte vinden ze een vergelijking met andere academiejaren ook interessant, omdat het de evolutie van de studenten toont. Over vergelijkingen met andere studenten zijn de meningen meer verdeeld, waarbij een kleine meerderheid van

de lesgevers aangeeft dit interessant te vinden. De voorkeuren van de studenten liggen in dezelfde lijn, waarbij zij voornamelijk geïnteresseerd blijken in punten en feedback op hun examens en opdrachten, aantal verworven *credits* en gepercipieerde werkdruk. Ze zijn ook positief over vergelijkingen met voorgaande prestaties/academiejaren, omdat ze zo hun voortgang kunnen nagaan. De meningen van studenten zijn echter verdeeld tot negatief wat het vergelijken betreft van eigen prestaties met die van andere studenten. Sommige studenten vrezen een demotiverend effect van deze vergelijkingen. Andere studenten zien toch een meerwaarde omdat het hun eigen prestaties in perspectief kan plaatsen. Opleidingshoofden, ten slotte, zijn het grotendeels eens met de andere gebruikersgroepen, en vinden veel variabelen interessant. Zo hebben ze ook interesse in aanvullende variabelen waaronder het aantal generatiestudenten, de sociale integratie en het bekwaamheidsgevoel van studenten.

De informatie die lesgevers minder relevant vinden zijn het algemene gebruik van de digitale leeromgeving en de mate waarin er gewerkt wordt aan de leerdoelen. Ook achtergrondkenmerken zoals leeftijd, het aantal generatiestudenten, en de fysieke afstand tot de campus achten ze minder belangrijk. Studenten ervaren eveneens geen nood aan informatie over de leerdoelen, want ze vinden het de taak van de lesgevers om te zien of de leerdoelen behaald worden. Vanwege het mogelijk controlerend karakter willen ze liever niet dat lesgevers zicht krijgen op procesvariabelen zoals het gebruik van de digitale leeromgeving. Een overzicht van de voorkeuren van de verschillende gebruikersgroepen inzake de inhoud van de *LA-dashboards* kan worden teruggevonden in Tabel 1.

Met betrekking tot de *lay-out van LA-dashboards* tonen de resultaten van de focus-groepgesprekken en *online survey* dat de lesgevers en opleidingshoofden voorstander zijn van een zakelijke *lay-out*, die serieus en to-the-point is en waarbij voornamelijk gebruik wordt gemaakt van figuren, grafieken en tabellen. Studenten daarentegen staan ook open voor een speelse *look-and-feel* (of een combinatie van speels en zakelijk), waarbij minder gebruik wordt gemaakt van cijfers en grafieken, maar wel extra afbeeldingen of icoontjes en eventueel een spelelement worden geïntegreerd.

Zowel lesgevers als studenten prefereren een beperkte *hoeveelheid informatie* die op een overzichtelijke manier wordt weergegeven. Ze wensen ook controle over de informatie die ze te zien krijgen. Een standaardaanbod dat ze zelf kunnen aanpassen lijkt daarmee de voorkeursoptie voor lesgevers en studenten. Opleidingshoofden daarentegen willen veel en gedetailleerde informatie zien, met een grote autonomie in het selecteren van de te presenteren informatie. Ten slotte geven alle gebruikersgroepen aan dat ze zeggenschap willen in waar en wanneer *dashboards* verschijnen. Het automatisch verschijnen van *dashboards* in de digitale leeromgeving vinden ze geen probleem, maar ze willen deze ook op eigen initiatief kunnen activeren of uitzetten, minimaliseren of vergroten.

**Tabel 1** Voorkeuren van de gebruikersgroepen inzake de inhoud van een LA-dashboard aan de AP Hogeschool Antwerpen, op basis van de resultaten van de focusgroep-gesprekken en online survey

	Potentiële variabelen voor de <i>LA-dashboards</i> :	Voorkeuren van:		
		Studenten	Lesgevers	Opleidingshoofden
<b>Presage</b>	Didactische kwaliteit volgens lesgever		+	+
	Vooropleiding		+	+
	Leeftijd		-	
	Aantal generatiestudenten		-	+
	Planningsvermogen	+	+	+
	Fysische afstand tot de hogeschool		-	
	Studiemotivatie	-	+	+
	Bekwaamheidsgevoel			+
	Geletterdheid		+	+
	Gecijferdheid		+	
<b>Perception</b>	Sociale integratie			+
	Academische integratie	+		+
	Perceptie op de didactische kwaliteit		+	+
	Algemene tevredenheid		+	+
	Gepercipieerde werkdruk	+	+	+
<b>Process</b>	Gebruik van leerstrategieën	-	+	+
	Activiteitsgraad in digitale leeromgeving	-	+ -	+
	Type activiteiten in digitale leeromgeving	-	+ -	+
<b>Product</b>	Verworven credits	+	+	+
	Punten op taken en testen	+	+	+
	Punten op een specifieke OLOD	+		+
	Gemiddeld examenresultaat	+		
	Evolutie richting leerdoelen	-	-	
	Feedback op taken en testen	+		
	% dat taken meetellen in eindresultaat	+		
	Vergelijking met eerdere prestaties	+		+
	Vergelijking met andere studenten	+ -		+ -

*Noot:* + : de variabele wordt als interessant bevonden, - : de variabele wordt als niet interessant bevonden, + - : de meningen omtrent deze variabele zijn verdeeld, geen aanduiding: de meningen omtrent deze variabele zijn neutraal.



### Secundaire analyses op bestaande databanken

De databankanalyses om OV3 te beantwoorden, startten met een inventarisatie van beschikbare en relevante databanken. Daartoe werden zeven interviews uitgevoerd met experts uit de ICT-diensten, de cel kwaliteitszorg, de cel onderwijsadministratie en de studentgerichte diensten. Deze experts werden bevraagd over de beschikbaarheid, toegankelijkheid, structuur en inhoud van de databanken. Uit de gesprekken kon worden geconcludeerd dat twee databanken de meeste voor *dashboards* relevante informatie zouden afdekken. Enerzijds bevatte de databank achter de digitale leeromgeving niet enkel alle data betreffende het gebruik van de digitale leeromgeving. Omdat de leeromgeving ook wordt gebruikt voor de beginassessments en voor (een deel van) de kwaliteitszorgbevestigingen konden ook maten voor gecijferdheid, geletterdheid, leerstrategiegebruik, studiemotivatie, welbevinden en de percepties van studenten op de didactische kwaliteit van opleidingsonderdelen uit deze databank worden gehaald. Anderzijds bevatte de databank van het schooladministratiesysteem verschillende achtergrondvariabelen en prestatiematen.

Uit deze databanken werd als basis voor de analyses een databestand geëxtraheerd voor vier opeenvolgende academiejaren, beginnend bij het academiejaar 2016–2017. Het eerste en grootste deel van deze analyse bestond er in om de juiste gegevens te weerhouden. Om de brug te slaan tussen de beschikbare informatie in de databanken en de informatie die relevant bleek volgens de literatuurstudie of het behoefteonderzoek werden de velden van de databank geannoteerd op basis van een interactiemodel. De annotatie bracht bovendien nieuwe informatie aan het licht die achteraf een voorspellende waarde bleek te hebben zoals bijvoorbeeld de domicilieregio van studenten of de hoeveelheid communicatie tussen lesgevers en studenten.

Tijdens het weerhouden van de informatie uit de databanken werd elke variabele achtereenvolgens gecontroleerd a.d.h.v. de onderstaande criteria. Variabelen werden in de analyse niet meegenomen indien:

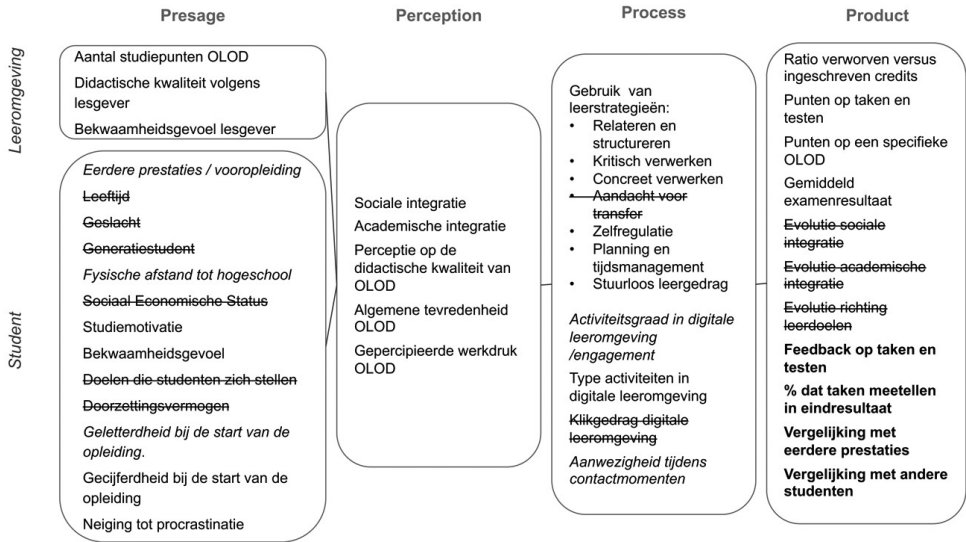
- Er (bijna) geen data beschikbaar was (*missing*);
- De betekenis van de variabele niet kon worden achterhaald (*vague*);
- De technische of inhoudelijke relevantie niet kon worden aangetoond (*irrelevant*);
- Het risico bestond dat de privacy van student, lesgever of opleidingshoofd werd geschonden volgens de opgestelde *Data Protection Impact Assessment* (Zie verder (*private*));
- Het gevoelige informatie betrof die een bepaalde doelgroep viseerde en/of mogelijke vooringenomen keuzes uit het verleden zou bevestigen (*unethical*);
- Uit verkennende analyses bleek dat de informatie geen goede weerspiegeling gaf van de realiteit (*unreliable*);
- Er wel voldoende en betrouwbare data voorhanden was maar deze een te lage informatieve meerwaarde gaf om waarschijnlijkheid in voorspellingen te bepalen (*low information*).

Voor de eigenlijke predictieve analyses werd de data opgesplitst in een training en test set. Op de eerste set werd een *machine learning* algoritme toegepast, een *elastic net*-geregulariseerde lineaire regressie (Zou & Hastie, 2005). Uit het resulterende lineaire model valt af te leiden welke variabelen een significante bijdrage leveren aan de voorspellingen en welke niet. De eerste analyses toonden aan dat met 148 van de 400 weerhouden variabelen het slaagcijfer van studenten kan worden voorspeld en dat het toevoegen van nog meer variabelen de voorspelling niet aanzienlijk verbetert. Op dit ogenblik is er nog geen definitieve uitspraak rond de accuraatheid en de precisie van de voorspellingen, maar er kunnen reeds een aantal vermoedelijke voorspellers van het slaagcijfer aangeduid worden (niet in volgorde van sterkte):

- Het departement, de opleiding en het opleidingsonderdeel
- De tussentijdse cijfers op opdrachten
- De lesgever van een opleidingsonderdeel
- De vooropleiding van de student
- De activiteit van studenten op de digitale leeromgeving
- De taalvaardigheid van de student
- De aanwezigheid van de student tijdens de contactmomenten
- De complexiteit van de cursus (= aantal verschillende modules)

#### **Selectie van variabelen voor het prototype van het generiek LA-dashboard**

Op basis van het gebruikersonderzoek en de databankanalyses werd de initiële selectie aan variabelen uit de literatuurverkenning verder verfijnd. Er kon in deze fase van het project niet eenduidig worden gekozen om enkel de variabelen in de doorsnede van de drie onderzoeksacties te selecteren omwille van verschillende redenen. Ten eerste gaven verschillende gebruikersgroepen soms tegenstrijdige voorkeuren inzake het opnemen van variabelen in het *dashboard*. Zo waren lesgevers en studenten het oneens over het opnemen van procesvariabelen zoals de activiteitsgraad op de digitale leeromgeving. Ten tweede spraken de resultaten van verschillende onderzoeksacties elkaar soms tegen. Zo bleek uit de databankanalyses dat o.m. de activiteitsgraad van studenten op de digitale leeromgeving of de fysieke afstand tot de hogeschool wel degelijk een voorspellende waarde hebben voor studiesucces, hoewel (sommige) gebruikersgroepen de relevantie van de informatie niet inzagen. Ten derde kon het effect van sommige relevante variabelen nog niet adequaat worden ingeschat in de databankanalyses omdat ze slechts recent werden geïmplementeerd. Zo werd het beginassessment op AP pas in 2019 ingevoerd en werden de eerste pilots met bevragingen over de didactische kwaliteit van opleidingsonderdelen via de digitale leeromgeving pas in 2020 uitgevoerd. De daadwerkelijke meerwaarde van deze potentieel relevante variabelen kon dus nog onvoldoende worden ingeschat. Ten vierde werden in het gebruikersonderzoek ook een aantal variabelen vermeld die niet werden geïnventariseerd in de literatuurverkenning. Om voldoende rekening te houden met de behoeften van de gebruikers werden deze ook meegenomen.



**Figuur 4** aanpassingen aan de selectie van potentiële variabelen voor de *LA-dashboards* op basis van gebruikersonderzoek en databankanalyses

*Noot:* doorstreept: variabele niet opgenomen, *italic:* sterke voorspeller uit databankanalyses, **bold:** extra variabele uit gebruikersonderzoek

Daartoe werd besloten om de selectie van variabelen voldoende ruim te nemen, deze te verwerken in de eerste prototypes en vervolgens (opnieuw) af te toetsen bij gebruikers. Figuur 4 geeft een overzicht van de aangepaste selectie van variabelen.

### Onderzoek naar juridische drempels bij de ontwikkeling van LA

Ten slotte werd de juridische en wettelijk basis voor het verwerken van persoonsgegevens bij het gebruik van *LA-dashboards* onderzocht (ov4). Zo werd verkend in welke mate het publiceren van gegevens in *LA-dashboard* in lijn is met de algemene verordening gegevensbescherming (GDPR) en het privacy-beleid van de AP Hogeschool. Gesprekken met experts op AP en met de *Data Protection Officer* van AP toonden aan dat het noodzakelijk was een *data protection impact assessment* uit te voeren (DPIA). Deze analyse moet worden uitgevoerd bij dataverwerkingsactiviteiten die hoge risico's stellen aan de rechten en vrijheden van individuen, bijvoorbeeld als er aan *profiling* wordt gedaan of wanneer er met gevoelige informatie wordt gewerkt. De DPIA brengt de grootste risico's (vb. datalekken) van deze verwerking in kaart en belicht de manier waarop de gegevensverwerkers met deze risico's zullen omgaan. In overleg met de DPO, vertegenwoordigers van de ICT-diensten en juridische experts werd deze DPIA gemaakt. Vervolgens werd deze voorgelegd op het hoogste beleidsorgaan op AP, het directieteam, en daar formeel goedgekeurd. Op basis van het DPIA werd o.m. de privacyverklaring van AP aangepast waardoor het gebruik van data voor *LA-dashboards* mogelijk wordt gemaakt.

Inhoudelijk stellen de DPIA en het privacy-beleid aan dat studenten of lesgevers eigenaar blijven van hun eigen data. Ze hebben het recht om hun data in te kijken, te laten corrigeren of verwijderen, voor zover de taken van AP mogelijk blijven. De *dashboards* in het LAP!-project verwerken wel degelijk een aantal persoonsgegevens, vooral dan van studenten, maar ook hier werden een aantal doordachte keuzes gemaakt. Zo worden geen persoonsgegevens opgevraagd die behoren tot de bijzondere categorieën van persoonsgegevens zoals ras of geloof en worden enkel die gegevens opgevraagd waarvan het onderzoek in het LAP!-project aantoont dat ze relevant zijn. De data van studenten worden zes jaar gebruikt. Zolang studenten ingeschreven zijn op de AP Hogeschool wordt daarbij studentnummer meegenomen om gegevens te koppelen. Nadien worden de gegevens geanonimiseerd, maar wel nog gebruikt voor vergelijkingen en historische analyses. Studenten en lesgevers zien hun eigen gegevens. Gebruikers zien gegevens van anderen enkel op een geaggregeerd niveau.

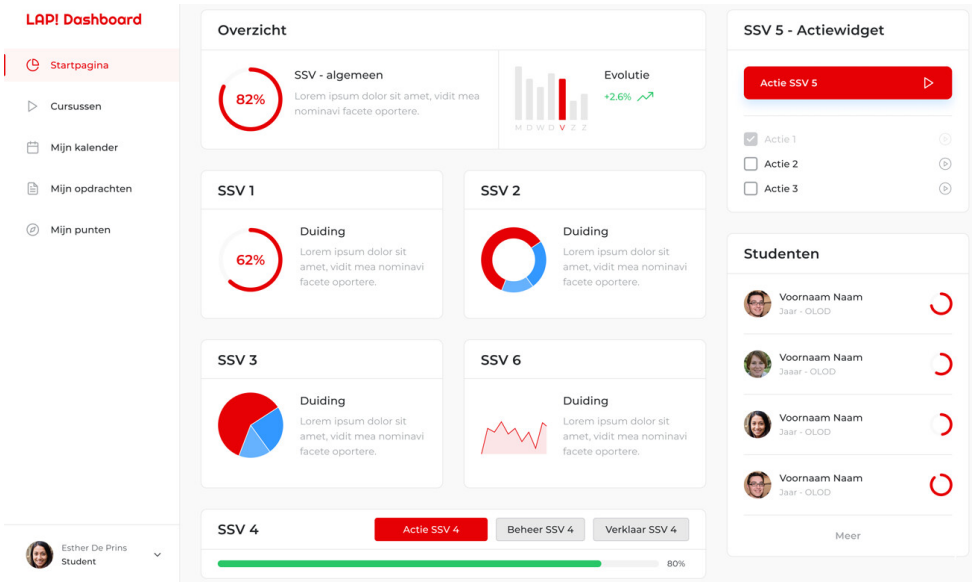
### Ontwerp en ontwikkelingsfase

In oktober 2020 werd de volgende fase uit het model voor educational design research opgestart: de ontwerp en ontwikkelfase. Deze fase zal drie academiejaren duren, waarbij in elk academiejaar een ontwerp- en een testfase elkaar afwisselen. Via deze cyclische manier van ontwerpen kunnen *dashboards* geleidelijk aan worden verfijnd op maat van de verschillende gebruikersgroepen en kan ook de adaptieve en interactieve component worden uitgebreid.

Tussen oktober 2020 en maart 2021 werden zo verschillende varianten van een prototype voor het generieke *dashboard* ontwikkeld (zie ook Figuur 5) gebaseerd op de eerder beschreven onderzoeksactiviteiten.

Vanaf mei 2021 wordt via onderzoek bij de verschillende gebruikersgroepen nagegaan (1) welke informatie de gebruikersgroepen raadplegen in de *dashboards*; (2) hoe de gebruikersgroepen de *dashboards* gebruiken en welke interpretaties ze maken; (3) welke additionele (ondersteunende) informatie ze wensen om de *dashboards* adequaat te gebruiken; (4) waar de inhoud en *lay-out* van de *dashboards* kunnen worden verbeterd en (5) hoe de *dashboards* verder kunnen worden verfijnd naar de verschillende gebruikersgroepen toe. Inzichten uit deze fase zullen dan opnieuw een volgende ontwerpcyclus aansturen.

Deze cyclus van ontwerpen, implementeren en testen herhaalt zich nogmaals in de volgende twee academiejaren, waarbij in het laatste academiejaar de nadruk wordt gelegd op het adaptief maken van de *dashboards* op basis van AI. Het doel daarbij is dat het *dashboard* mee evolueert met de voorkeuren van de gebruikers. Ook zal het algoritme dat gebruikt werd in de databankanalyses verder worden ontwikkeld voor het beantwoorden van complexere vragen. Zo moet het leeralgoritme een onderscheid kunnen maken tussen dynamische variabelen waarvoor een actie van de student vereist is, en stabiele



**Figuur 5** Impressie van een prototype van het generieke *LA-dashboard*

variabelen, zoals de vooropleiding. Met deze uitbreiding van het algoritme willen we een meer complexe vraag beantwoorden zoals: gegeven een bepaalde vooropleiding, wat is de meest effectieve actie die de student kan ondernemen om studiesucces te verhogen? Zo kan AI ook interveniëren wanneer blijkt dat gebruikers belangrijk geachte informatie onvoldoende aandacht geven.

## Reflecties

Aangezien het LAP!-project op dit ogenblik enkel de analyse en exploratiefase en een deel van de ontwerp en ontwikkelfase doorliep, leek het ons te voorbarig om conclusies te trekken. Daarom eindigen we deze bijdrage met een aantal reflecties, waarvan we hopen dat ze anderen kunnen inspireren.

Zoals eerder aangegeven bleek een (definitieve) selectie van op te nemen variabelen niet vanzelfsprekend. Het lijkt erop dat een geleidelijke verfijning op basis van cycli van ontwerp, implementatie en evaluatie meer aangewezen is. Tevens is het belangrijk om bij de selectie van variabelen en de uitwerking van de *dashboards* het uiteindelijke doel van de *dashboards* voor ogen te houden. Die doelstelling was bij de aanvraag op de eerste plaats 'ondersteunend' voor het leerproces en de didactische informatie. Het feit dat er voorspellende analyses worden gebruikt als onderzoeksactie werkt daarbij soms vertroebelend. Enerzijds wijzen deze analyses op relevante variabelen. Anderzijds zijn de geïdentificeerde variabelen niet steeds variabelen waarnaar je op zoek bent van-

uit het oogpunt van ondersteuning. Het is belangrijk om een kritische reflectie op de onderzoeksdata te blijven maken vanuit de doelstellingen. Deze reflectie kan ook door worden getrokken naar (toekomstige) visualisaties. Ook daar kan de samenhang tussen variabelen op basis van correlaties of voorspellingen informatief zijn. Toch moeten ook hier afwegingen gemaakt worden tussen informatieve meerwaarde enerzijds en 'over-informeren' of 'over-compliceren' langs de andere kant. Op basis van toekomstige ontwerpcycli zal deze balans per gebruikersgroep moeten worden geïdentificeerd.

Uit gebruikersonderzoek blijkt dat verschillende gebruikersgroepen wel een gelijkaardige relevantie zien in het gebruik van *LA-dashboards*, namelijk het monitoren van processen. Ze percipiëren wel gedeeltelijk andere variabelen om dit mogelijk te maken. Positief daarbij is dat gebruikersgroepen op de eerste plaats variabelen selecteren die veranderbaar zijn en waarop ze een invloed kunnen hebben. Lesgevers en opleidingshoofden prefereren tevens niet alleen studentvariabelen, maar ook variabelen waarmee ze hun eigen rol kritisch in vraag kunnen stellen, zoals percepties op de didactische kwaliteit. Het is opvallend dat gebruikersgroepen (en vooral studenten) uitkomst- of prestatievariabelen prefereren en minder de procesvariabelen die de uitkomstvariabelen (mee) kunnen helpen verklaren. Dit stelt de intrigerende vraag of uitkomstvariabelen voor gebruikers voldoende handvatten bieden om hun processen te monitoren en bij te sturen. Voor studenten wordt dit hoofdzakelijk gelinkt aan een gevoel van onveiligheid en controle. Echter, een meerwaarde van *LA-dashboards* is net dat ze gegevens uit verschillende bronnen aan elkaar linken om zo processen inzichtelijk te maken. Het spanningsveld tussen een veilig klimaat en de informatieve waarde van variabelen is iets waar in de verdere ontwikkeling rekening mee zal worden gehouden, bijvoorbeeld via het anonimiseren of aggregeren van data.

Het samenvoegen van gegevens uit de verschillende databanken was een grotere uitdaging dan oorspronkelijk gedacht. Enerzijds was er de uitdaging om alle mogelijke privacy-risico's te inventariseren en om strategieën te bedenken om deze te voorkomen of remediëren. Vervolgens vergde het nodige overleg om onderzoekers toestemming te geven om data uit deze databanken te exporteren. Ten slotte bleek het ook moeilijk om de databanken correct te koppelen omdat dit niet enkel diende te gebeuren op basis van de identiteit van studenten, maar ook op basis van opleidingsonderdelen en lesgevers. De tijd die nodig was om dit proces te doorlopen werd onderschat, waardoor de projectplanning onder druk kwam. In sommige gevallen kan het werken met een kleinere set aan gesimuleerde data een tussenoplossing bieden.

Kwalitatieve *LA-dashboards* zijn cruciaal in door data-geïnformeerde begeleiding en innovatie. Met het LAP!-project werden de eerste stappen gezet in de door onderzoek geïnformeerde ontwikkeling hiervan. Het eerste onderzoekjaar richtte zich op basis van *educational design research* op onderzoek inzake de inhoud, structuur, relevantie, *layout* en juridische grondslagen van deze *dashboards*. Het leverde relevante resultaten op die niet altijd uitsluitend gaven, maar wel voldoende richting voor de ontwikkeling van een eerste prototype dat als basis kan dienen voor verdere verfijning.

## Literatuur

- Avella, J., Kebritchi, M., Nunn, G., & Kanai, T. (2016). Learning analytics methods, benefits, and challenges in higher education: A systematic literature review. *Online Learning*, 20 (2), 13–29.
- Baeten, M., Kyndt, E., Struyven, K., & Dochy, F. (2010). Using student-centred learning environments to stimulate deep approaches to learning: Factors encouraging or discouraging their effectiveness. *Educational Research Review*, 5, 243–260.
- Biggs, J.B. (2003). *Teaching for quality learning at university*. Buckingham: Open University Press.
- Broos, T., Pinxten, M., Delporte, M., Verbert, K., & De Laet, T. (2020). Learning dashboards at scale: early warning and overall first year experience. *Assessment & Evaluation in Higher Education*, 45(6), 855–874.
- Buckingham Shum, S., Ferguson, R., & Martinez-Maldonado, R. (2019). Human-Centred Learning Analytics. *Journal of Learning Analytics*, 6(2), 1–9.
- Gasevic, D., Jovanovic, J., Pardo, A., & Dawson, S. (2017). Detecting learning strategies with analytics: Links with self-reported measures and academic performance. *Journal of Learning Analytics*, 4(1), 113–128.
- Gijbels, D., Donche, V., Richardson, J.T., & Vermunt, J.D. (Eds.) (2013). *Learning patterns in higher education: Dimensions and research perspectives*. Routledge.
- Hattie, J. (2015). The applicability of Visible Learning to higher education. *Scholarship of teaching and Learning in Psychology*, 1(1), 79.
- McKenney, S. & Reeves, T. (2014). *Conducting Educational Design Research*. London: Routledge
- Papamitsiou, Z. & Economides, A. (2014). Learning Analytics and educational data mining in practice: A systematic literature review of empirical evidence. *Educational Technology & Society*, 17(4), 49–64.
- Price, L. (2013). *Modelling factors for predicting student learning outcomes in higher education*. In D. Gijbels, V. Donche, J. Richardson, & J. Vermunt (Eds.), *Learning Patterns in Higher Education: Dimensions and Research Perspectives* (pp. 56–77). Routledge.
- Richardson, M., Abraham, C., & Bond, R. (2012). Psychological correlates of university students' academic performance: a systematic review and meta-analysis. *Psychological Bulletin*, 138, 353–387.
- Sclater, N., Peasgood, A., & Mullan, J. (2016). *Learning Analytics in Higher Education: A review of UK and international practice*. Bristol: JISC.
- Siemens, G. & Gasevic, D. (2012). Learning and knowledge analytics. *Journal of Educational Technology & Society*, 15(3), 1–2.
- van Rooij, E., Brouwer, J., Fokkens-Bruisma, M., Jansen, E., Donche, V., & Noyens, D. (2018). A systematic review of factors related to first-year students' success in Dutch and Flemish higher education. *Pedagogische Studiën*, 94, 360–406.
- Vanhoumout, G., Van de Mosselaer, H., Donche, V., & Vansteenkiste, M. (2016). Discovering and strengthening learning strategies and motivation using the LEMO-instrument. In P.J. Bonne & D. Nutt (Eds.), *Ten Times the First Year: Reflections on ten years of the European First Year Experience Conference* (pp. 53–70). Leuven: LannooCampus.

- Verbert, K., Duval, E., Klerkx, J., Govaerts, S., & Santos, J.L. (2013). Learning Analytics Dashboard Applications. *American Behavioral Scientist*, 57(10), 1500–1509. <https://doi.org/10.1177/0002764213479363>
- Verbert, K., Govaerts, S., Duval, E., Santos, J.L., Van Assche, F., Parra, G., & Klerkx, J. (2014). Learning dashboards: an overview and future research opportunities. *Personal and Ubiquitous Computing*, 18(6), 1499–1514. doi:10.1007/s00779-013-0751-2.
- Zou, H., & Hastie, T. (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 67, 301–320. <https://doi.org/10.1111/j.1467-9868.2005.00503.x>