# “LEARNING ANALYTICS” IN HET VLAAMS HOGER ONDERWIJS

Tinne De Laet e.a.



**KVAB STANDPUNTEN**
**58**
Koninklijke Vlaamse Academie van België
voor Wetenschappen en Kunsten - 2018

# “LEARNING ANALYTICS” IN HET VLAAMS HOGER ONDERWIJS

KVAB Press

**KVAB STANDPUNTEN**
**58**

Ontwerp cover: Francis Strauven
Beeld: Joris Snaet
De tekening van het Paleis der Academiën is een reproductie van het originele perspectief van Charles Vander Straeten in 1823. Jozef Cantré ontwierp het logo van de KVAB in 1947.

De KVAB Standpunten worden gepubliceerd door de Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten, Hertogsstraat 1, 1000 Brussel.
Tel. 00 32 2 550 23 23 – info@kvab.be – www.kvab.be

# “LEARNING ANALYTICS” IN HET VLAAMS HOGER ONDERWIJS


Jan Aerts
Frederik Cornillie
Tinne De Laet
Olga De Troyer
Jochen De Weerdt
Jan Elen
Anouk Gelan
Katie Goeman
Els Kindt
Dimitri Mortelmans
Wim Van den Noortgate
Joos Vandewalle
Katrien Verbert
Pieter Vos




© Copyright 2018 KVAB
D/2018/0455/08
ISBN 978 90 656 918 73

Gedeeltelijke reproductie is toegelaten mits uitdrukkelijke bronvermelding.
Partial reproduction is permitted provided the source is mentioned.
Aanbevolen citeerwijze: Tinne De Laet e.a., *"Learning Analytics" in het Vlaams hoger onderwijs*, KVAB Standpunt 58, 2018.

Drukkerij Universa



*Opgedragen aan Georges Van der Perre*
*(† 14 september 2017)*

# “Learning Analytics” in het Vlaams hoger onderwijs

INHOUDSTAFEL


Inhoud . . . . . 3

Samenvatting . . . . . 5

Executive Summary . . . . . 7

Voorwoord. . . . . 9

I. Inleiding . . . . . 11
- A. Wat is Learning Analytics? . . . . . 11
- B. Doel van dit standpunt . . . . . 12
- C. Learning Analytics in het hoger onderwijs: het kader . . . . . 12
- D. Learning Analytics als apart domein. . . . . 13
- E. Belanghebbenden . . . . . 16
- F. Kernaspecten . . . . . 16
- G. Opbouw van dit standpunt . . . . . 18

II. Waarom? . . . . . 20
- A. Onderzoek naar onderwijsleerprocessen. . . . . 20
  - A.1. Traditionele benadering . . . . . 20
  - A.2. Opportuniteiten . . . . . 21
  - A.3. Uitdagingen . . . . . 22
- B. Optimaliseren en personaliseren van de leeromgeving . . . . . 22
  - B.1. Wat, welke databron, hoe en waarom? . . . . . 23
  - B.2. Opportuniteiten . . . . . 24
  - B.3. Uitdagingen . . . . . 24
- C. Studieloopbaanbegeleiding . . . . . 26
  - C.1. Het belang van context . . . . . 26
  - C.2. Learning dashboards voor studieloopbaanbegeleiding . . . . . 27
  - C.3. Opportuniteiten . . . . . 29
  - C.4. Uitdagingen . . . . . 30

III. Data . . . . . 33
- A. Soorten data . . . . . 33
- B. *Small* of *big data* . . . . . 35
- C. Belang van context . . . . . 35

IV. Hoe? . . . . . 38
A. Dataverzameling, -opslag en -voorverwerking. . . . . 38
A.1. Dataverzameling. . . . . 38
A.2. Dataopslag . . . . . 40
A.3. Datavoorververwerking . . . . . 43
B. Data-analyse . . . . . 43
B.1. Statistische en psychometrische technieken. . . . . 43
B.2. Machinaal leren . . . . . 46
C. Visualisatie . . . . . 48
D. Descriptief, predictief of prescriptief. . . . . 50
E. Technologische oplossingen voor LA . . . . . 53

V. Randvoorwaarden . . . . . 56
A. Privacy- en wettelijke aspecten: juridische en organisatorische instrumenten . . . . . 56
A.1. Transversale principes van het gegevensbeschermingsrecht. . 59
A.2. Andere verplichtingen onder het gegevensbeschermingsrecht 61
A.3. Profilering en geautomatiseerde individuele besluitvorming . . 65
A.4. Effectbeoordeling en inschatting van de risico's . . . . . 67
A.5. Gebruik van LA voor statistiek of wetenschappelijk onderzoek 68
A.6. Besluit. . . . . 67
B. Ethische aspecten . . . . . 69
B.1. Verzameling, gebruik en interpretatie van data . . . . . 70
B.2. Actie en interventies . . . . . 71
B.3. Toestemming, beschikbaarheid en deïdentificatie van data. . . 73
B.4. Afmelding . . . . . 73
B.5. Transparantie en betrekken van belanghebbenden . . . . . 74

VI. Besluit en aanbevelingen . . . . . 75

VII. Terminologie. . . . . 79

VIII. Bronnen . . . . . 84
A. Wetgeving, Richtlijnen en Aanbevelingen . . . . . 84
B. Andere bronnen . . . . . 84

IX. Samenstelling van de werkgroep . . . . . 90

# Samenvatting

Learning Analytics (LA) is een nieuwe term in het onderwijslandschap. LA gaat echter verder dan de letterlijke vertaling 'analyse van het leren' doet vermoeden. LA ambieert gegevens van studenten en hun context te verzamelen, te analyseren en die vervolgens ook in te zetten om het leren te verbeteren. Het is niet verwonderlijk dat LA een *hot topic* is: het bouwt immers op gegevens over leren, die meer en meer beschikbaar komen als gevolg van de groei van onlineleren en *blended learning.* LA lijkt beloftevol om impact te hebben op het Vlaams hoger onderwijs. Maar op welke vlakken kan die impact gerealiseerd worden? En kan LA de hoge verwachtingen inlossen, want zoals Willem Elsschot al wist: "tussen droom en daad staan wetten in de weg en praktische bezwaren"?

Binnen het Vlaams hoger onderwijs zijn er drie grote toepassingsdomeinen voor LA: het onderzoek naar onderwijsleerprocessen, de optimalisatie en personalisatie van de leeromgeving, en studieloopbaanbegeleiding. Telkens zal de toepassing van LA hier raken aan veel domeinen, gaande van onderwijskunde tot *machine learning*. De effectieve realisatie van LA vergt dan ook een radicaal interdisciplinaire samenwerking van onderwijskundigen, cognitieve en motivatiepsychologen, specialisten in de concrete leerdomeinen, technologen, datawetenschappers, privacydeskundigen en ethici. Daarnaast moet LA de brug leggen tussen onderzoek en praktijk: zowel onderzoekers als *practitioners* (docenten, studieloopbaanbegeleiders, studenten) moeten van bij de eerste stap van de ontwikkeling van LA-initiatieven mee aan tafel zitten.

LA steunt op de beschikbare data over leergedragingen. Een sluimerend gevaar is dat LA leidt tot een te grote focus op feitelijke gegevens. Elke gebruiker, onderzoeker en beleidsmaker moet er daarom over waken dat LA ingezet wordt om de persoonlijke aanpak en inzichten binnen het onderwijs te versterken, en zeker niet om het menselijk contact en de ruimte voor persoonlijke omstandigheden te reduceren. Elke instelling die met LA aan de slag wil gaan kan daartoe een richtinggevende beleidstekst opstellen, met aandacht voor de privacy, de wettelijke en de ethische aspecten van LA én de kadering van LA binnen de meer algemene onderwijsvisie.

Het Vlaams hoger onderwijs kan inspiratie halen uit buitenlandse voorbeelden. De specifieke context van een Vlaamse instelling van hoger onderwijs mag hier echter niet uit het oog verloren worden. Deze heeft immers een grote invloed op veel aspecten van LA, zoals de beschikbare data (bv. *swipe cards* en meten van aanwezigheid), de onderwijscontext (bv. vrije toegang tot de opleiding of selectie), de doelstellingen van LA (bv. retentie of snelle heroriëntering), de nationale en instellingsspecifieke regelgeving en de onderwijs- en begeleidingspraktijken (bv. al dan niet een persoonlijke en professionele studieloopbaanbegeleider voor elke student).

De realisatie van LA vraagt niet enkel een investering in technische infrastructuur, maar ook de uitwerking van ethische kaders, de bewaking van de wettelijke aspecten en de ondersteuning en vorming van alle betrokkenen. LA, en bij uitbreiding educatieve technologie, belangt elke Vlaamse instelling van hoger onderwijs aan. De gemeenschappelijke context waarin deze instellingen opereren creëert gelijkaardige uitdagingen en opportuniteiten. Onder het motto "samen staan we sterk" zouden instellingen daarom de nodige structuren moeten oprichten om samenwerking, ervaringsuitwisseling en een richtinggevend kader te realiseren. Internationale voorbeelden tonen aan dat nationale samenwerkingsorganisaties rond educatieve technologie niet alleen een impuls geven aan de concrete toepassing van LA, maar ook aan het LA-onderzoek. De oprichting van een Vlaamse of Belgische samenwerkingsorganisatie rond educatieve technologie of de aansluiting bij een bestaande organisatie strekt dan ook tot aanbeveling.

Kortom, LA is een domein met potentieel, maar ook heel wat uitdagingen!

# Executive Summary

## "Learning Analytics" in the Flemish higher education system

Learning Analytics (LA) is an emerging educational technology. It aims for the collection and analysis of data from students and their context and for using this data to improve learning. It is not surprising that LA is a hot topic: it builds on educational data, which become more and more available due to the growth of online and blended learning. LA seems to have potential to impact Flemish higher education. In which areas can this impact however be realized? And, can LA meet the high expectations, as Willem Elsschot[1] already wrote in 1910 – "between dream and reality, laws stand in the way, and practical objections"[2]?

Within Flemish higher education there are three major application domains for LA: the research into educational learning processes, the optimization and personalization of the learning environment, and student advising. In these three applications, LA will touch many domains ranging from educational science to machine learning. The effective realization of LA therefore requires a radically interdisciplinary collaboration between educationalists, cognitive and motivational psychologists, specialists in the specific learning domains, technologists, data scientists, privacy experts, and ethicists. In addition, LA has to bridge the gap between research and practice: both researchers and practitioners (teachers, student advisors, and students) must join the discussion from the beginning of the development of LA initiatives.

LA relies on the available data on learning behaviours. A latent danger is that LA leads to an overfocus on factual data. Every user, researcher, and policymaker must therefore ensure that LA is used to reinforce a personal approach and insights into learning, and is not used to reduce personal contact and room for individual circumstances. Every institution that wants to use LA should work out a guiding policy text, with attention for privacy, legal, and ethical aspects of LA, framed within a more general educational vision.

Flemish higher education can draw inspiration from international examples. The specific context of the own institution and Flemish higher education in general, should however not be forgotten. The local context has a great influence on many aspects of LA, such as the available data (e.g. swipe cards and measurement of attendance), the educational context (e.g. free access to higher education or selection), the goals of LA (e.g. retention or rapid reorientation), national and

[1] Flemish writer and poet, 7 May 1882 – 31 May 1960.

[2] Translated from "tussen droom en werkelijkheid staan wetten in de weg en praktische bezwaren", Willem Elsschot, Het Huwelijk, 1910.

institution-specific regulations and teaching and advising practices (e.g. whether or not each student has a personal and professional student advisor).

The realization of LA requires not only an investment in technical infrastructure, but also the elaboration of ethical frameworks, the monitoring of the legal aspects, and the support and training of all those involved. LA, and by extension educational technology, is of interest to every Flemish higher education institution. The common context in which these institutions operate creates similar challenges and opportunities. Under the motto "together we are strong", institutions should therefore set up the necessary structures to realize cooperation, exchange of experiences, and a guiding framework. International examples show that national cooperation organizations on educational technology give an impulse to both the concrete application of and research on LA. The establishment of a Flemish or Belgian cooperation organization on educational technology or the connection with an existing organization is therefore recommended.

In short, LA is a domain with great potential but also many challenges!

# Voorwoord

De reeks Standpunten van de Academie is een bijdrage tot het wetenschappelijk onderbouwd debat over actuele maatschappelijke en artistieke thema's. De auteurs, leden van KVAB en werkgroepen van de Academie schrijven in eigen naam, onafhankelijk en met volledige intellectuele vrijheid. De goedkeuring voor publicatie door één of meerdere Klassen van de Academie waarborgt de kwaliteit van de gepubliceerde studies. Dit Standpunt werd goedgekeurd voor publicatie door de Klasse Technische Wetenschappen op 30 september 2018.

Het thema "*Learning Analytics" in het Vlaams hoger onderwijs* kadert in de ruimere wereldwijde evolutie van de digitalisering, de *big data*, het toenemend gebruik van internet, en de toenemende kracht van algoritmiek en in het bijzonder van de artificiële intelligentie. De impact hiervan in de samenleving zal zich ook concreet meer en meer laten voelen op het onderwijsgebeuren. Studenten laten immers allerhande sporen na van hun onderwijsactiviteiten bij de onderwijsinstellingen. Die informatie kunnen de onderwijsinstellingen nuttig gebruiken om de studenten, docenten en studieloopbaanbegeleiders beter te ondersteunen, en dat is precies het thema van dit Standpunt. Maar deze nieuwe instrumenten moeten kaderen in een onderwijsvisie en houden ook diverse gevaren in voor de privacy, gebrek aan transparantie, ongeoorloofd maar ook commercieel gebruik. Omdat de begrippen en de mogelijkheden van Learning Analytics nog niet breed gekend zijn, en om eerder een convergerende dan een concurrerende aanpak in het Vlaams hoger onderwijs te verkrijgen, leek het voor de KVAB gewenst om een werkgroep met experten en geïnteresseerde betrokkenen bijeen te brengen om daar samen over te reflecteren. Om een voldoende concrete beschrijving en aanpak te kunnen formuleren, is er van bij het begin gekozen om het Standpunt te beperken tot het Vlaams hoger onderwijs. Het doel was immers om een werkbaar en breed gedragen voorstel uit te werken, waarmee de Vlaamse hoger onderwijsinstellingen hun systemen kunnen opzetten met als doel een betere begeleiding van de studenten te verkrijgen.

Dit neemt niet weg dat ook vele elementen uit het Standpunt waardevol kunnen zijn voor het secundair en basisonderwijs. In de lijn van de vorige KVAB activiteiten zoals het Denkersprogramma *E-learning en Blended Learning* in 2014 en het Standpunt *Hoger onderwijs voor de digitale eeuw* (2015) heeft confrater Georges Van der Perre dit onderwerp begin 2017 naar voor geschoven als een belangrijk thema dat verder verkend moet worden. Echter heeft hij deze werkgroep wel goed op de rails kunnen zetten, maar is daarna geconfronteerd geweest met ernstige gezondheidsproblemen, en hij is ons ontvallen op 14 september 2017. We dragen dit Standpunt dan ook graag aan hem op. Zijn collega Prof. Tinne De Laet nam de coördinatie van deze werkgroep over en leidde verder de redactie van het Standpunt. Dit Standpunt is iets uitgebreider dan de gewone Standpunten om twee redenen. Vooreerst is het onderwerp van Learning Analytics (LA) nog

niet goed gekend, en ook wordt het soms verward met andere methodieken. Ten tweede is er ook gekozen om de tekst te illustreren met een aantal concrete voorbeelden, die in een kader of figuur worden beschreven. Deze kunnen de nieuwkomers en nieuwsgierigen in het domein mee overtuigen van de concrete waarde van LA. De lezers die in het domein thuis zijn kunnen dan ook Secties I, II en III en de kaderstukjes overslaan.

De inspanningen van Tinne De Laet als trekker van deze werkgroep werden mede mogelijk gemaakt door de steun van het Erasmus+ programma van de Europese Unie in het kader van het STELA Erasmus+ project (562167-EPP-1-2015-1-BE-EPPKA3-PI-FORWARD, http://ww.stela-project.eu)

Co-funded by the
Erasmus+ Programme
of the European Union



The European Commission support for the production of this publication does not constitute an endorsement of the contents which reflects the views only of the authors, and the Commission cannot be held responsible for any use which may be made of the information contained therein.

# I. Inleiding

## A. Wat is Learning Analytics?

Learning Analytics (LA) is een nieuwe term in het onderwijslandschap. Een letterlijke vertaling is 'analyse van het leren'. Omdat er geen universeel geldende definitie beschikbaar is, citeren we enkele veel gehanteerde definities:

> *Learning Analytics is de meting, verzameling, analyse en rapportering van gegevens over studenten en hun context, met als doel het begrijpen en optimaliseren van het leren en de omgeving waarin dit gebeurt.*
>
> *vertaald van (Siemens, 2011, p. 1)*

> *Learning Analytics gaat over het verzamelen van sporen die studenten achterlaten en het gebruik van deze sporen om het leren te verbeteren.*
>
> *vertaald van (Duval, 2012, p. 1)*

> *Learning Analytics is het verzamelen en analyseren van gegevens uit leeromgevingen om het leerproces van studenten te verbeteren. Deze informatie wordt vervolgens beschikbaar gemaakt voor verschillende belanghebbenden, zoals de student zelf, de docent of het opleidingsmanagement.*
>
> *(Engelfriet, Manderveld, & Jeunink, 2017, p. 3)*

Gemeenschappelijk in al deze definities is eerst en vooral het uiteindelijke doel: de optimalisatie en verbetering van het leren is de kern van LA (Elias, 2011). Aan de bron staan gegevens of data: sporen die personen tijdens het leren, begeleiden of onderwijzen achterlaten. LA heeft als doel een brug te slaan tussen de data en de optimalisatie/verbetering van het leren, begeleiden en onderwijzen. LA is gegroeid uit het toenemend aantal beschikbare gegevens over het leren van studenten, een gevolg van de groei van onlineleren en *blended learning*.

De laatst geciteerde definitie introduceert nog een potentieel extra aspect van LA: dat de gegevens beschikbaar gesteld worden aan de belanghebbenden. Dat kan gebeuren aan de hand van *learning dashboards*, die 'rapporteren' aan de belanghebbenden. *Learning dashboards* bieden een visueel overzicht van de kerninformatie uit LA. Dat is zo vormgegeven dat het snel een volledig overzicht biedt. De *learning dashboards* willen reflectie, inzicht en zelfs gedragsverandering uitlokken, door de inzichten visueel voor te stellen.

## B. Doel van dit standpunt

Dit KVAB-Standpunt brengt de opportuniteiten en de gevaren van LA in kaart om op basis daarvan aanbevelingen te formuleren over LA in het Vlaams hoger onderwijs. De concrete ambitie is de belanghebbenden – beleidsmakers, studenten, docenten, studieloopbaanbegeleiders en onderzoekers – een houvast te bieden. Om dat doel te bereiken kadert het Standpunt het begrip LA eerst aan de hand van vier thema's: Waarom?, Data, Hoe? en Randvoorwaarden (privacy-, wettelijke en ethische aspecten). Het besluit zet de aanbevelingen op een rij.

Omdat de focus van dit document op het hoger onderwijs ligt, hebben we het over 'student(en)', voor de bredere Engelstalige term '*learner(s)*', en over 'docent(en)', voor het Engelse '*teacher(s)*'. 'Docent' verstaan we in brede zin, van onderwijsgevende tot onderwijsverzorgende: docenten, praktijkassistenten, tutoren, studieloopbaanbegeleiders…

## C. Learning Analytics in het hoger onderwijs: het kader

LA is een jong, multidisciplinair domein dat theorie, ontwerp en datawetenschap combineert (Gašević, Kovanović, & Joksimović, 2017). Het kwam voor het eerst aan bod in het *Horizon Report*[3] van 2011. De daaropvolgende rapporten voorspelden dat LA binnen twee tot drie jaar breed ingang zou vinden (Johnson, Adams, & Cummings, 2012) en vervolgens zelfs binnen het jaar (Johnson et al., 2013). Het *Horizon Report* van 2016 vermeldt dat de verwachting is dat LA op zeer korte termijn een steeds grotere plaats zal krijgen in instellingen in het hoger onderwijs (Johnson et al., 2016).

Dit rapport kadert LA binnen de groeiende aandacht voor het meten van leren en erkent dat de toegenomen beschikbaarheid van data over het leren hierbij een faciliterende factor is. Een veelheid van gegevens is beschikbaar of kan op korte termijn beschikbaar komen: socio-demografische informatie over studenten, academische prestaties en vooruitgang, activiteiten op het leerplatform en gebruik van digitale tools tijdens of na contactmomenten... Nieuwe technologische evoluties zoals digitale en onlineleeromgevingen, of gepersonaliseerde meet-instrumenten (*quantified self*) maken het mogelijk om op grotere schaal het leer- en studeergedrag van studenten, maar ook het doceergedrag van docenten, te bestuderen. Het is dan ook niet verwonderlijk dat het domein van LA enerzijds een impuls heeft gekregen door de ontwikkeling van MOOCs en anderzijds dat LA al kort aan bod kwam in het KVAB-Standpunt *Blended Learning* (Van der Perre et al., 2015).

[3] Het NMC Horizon Project is een wereldwijd onderzoeksinitiatief dat de trends, uitdagingen en technologische ontwikkelingen verkent die waarschijnlijk een impact zullen hebben op onderwijs en leren (https://www.nmc.org/nmc-horizon/).

Er zijn heel wat ontwikkelingen en concrete experimenten gaande, maar eigenlijk begint de sector nu pas een eerste inzicht te krijgen in hoe deze gegevens zinvol analyseerbaar en inzetbaar zijn in het leerproces. Enerzijds argumenteren onderzoekers dat, ondanks de enorme groei aan onderzoek, LA nog in zijn kinderschoenen staat (Leitner, Khalil, & Ebner, 2017). Anderzijds zijn er stemmen die stellen dat LA al een fase van volwassenheid is ingegaan waarin het onderzoek de praktijk en het beleid beïnvloedt (Gašević et al., 2017). Ondertussen zijn er al commerciële pakketten om LA in te zetten binnen het hoger onderwijs, die hoge verwachtingen creëren over de impact van LA. LA omvat heel wat aspecten, gaande van dataverzameling, -opslag, -beveiliging, -analyse en -interpretatie tot ethiek. Die maken dat het hier enerzijds gaat om een beloftevolle ontwikkeling, maar dat die anderzijds ook heel wat uitdagingen met zich meebrengt die de technologie overstijgen. Een recent literatuuronderzoek naar het gebruik van LA in het hoger onderwijs wereldwijd (Leitner et al., 2017) onderzocht welke struikelblokken er voor het gebruik van LA zijn. De meest voorkomende zijn de tijdsinvestering om de gegevens voor te bereiden en resultaten te verkrijgen, de omvang van de beschikbare gegevens en onderzochte populatie, en ethische aspecten. Daarnaast blijkt dat veel studies er maar in beperkte mate in slagen om de verschillende belanghebbenden erbij te betrekken. Experts halen aan dat er naast ethische richtlijnen ook nood is aan de integratie van expertise vanuit de pedagogische wetenschappen (Ferguson, 2012). Daarnaast is de ontwikkeling van methodologieën vereist die kunnen omgaan met uiteenlopende datasets en met een sterkere focus op het perspectief van de studenten (Ferguson, 2012).

### D. Learning Analytics als apart domein

LA is een onderdeel van het bredere domein Educational Analytics, dat vanaf 2010 uitgroeide tot een eigen deeldomein. Hoewel het niet mogelijk is om sterke schotten op te trekken tussen de verschillende domeinen binnen Educational Analytics, zijn er toch pogingen om aan de hand van duidelijkere definities de deeldomeinen beter af te bakenen (Ferguson, 2012):

– ***Educational Data Mining*** (Mohamad & Tasir, 2013; Cristóbal Romero, Romero, & Ventura, 2014; Cristobal Romero & Ventura, 2013) focust op de ontdekking van nieuwe patronen, verbanden en trends door de analyse van grote hoeveelheden data over leren. Deze patronen, verbanden en trends zouden met andere technieken moeilijk of onmogelijk te analyseren zijn.

**Voorbeeld: *Educational Data Mining*** (Merceron & Yacef, 2005)

Sydney University beschikt over een dataset verzameld binnen een webgebaseerd tutoring-instrument dat studenten binnen een opleidingsonderdeel ondersteunt bij het inoefenen van logische formele bewijzen. Het instrument identificeert of de stappen in een bewijs juist of fout zijn. De onderzoekers pasten datamining-algoritmes toe op de dataset om exploratief nieuwe kennis te ontdekken. Hierbij beantwoordden ze vragen als: vinden clusteringalgoritmes groepen van gelijkaardige studenten en oefeningen? Zijn er terugkerende oefenpatronen bij studenten die slecht presteren voor het vak?

– ***Academic Analytics*** focust op beleidsuitdagingen en/of politieke en economische uitdagingen: hoe kunnen leeropportuniteiten en onderwijs-resultaten verbeterd worden op opleidings- of instellingsniveau, of zelfs op regionaal, nationaal en internationaal niveau? Het betreft vooral cijfers over studiesucces en uitval die opleidingen, instellingen en overheden gebruiken om hun effectiviteit en efficiëntie te meten en te verbeteren.

**Voorbeeld: *Academic Analytics* op opleidingsniveau**

Opleidingen brengen de studievoortgang van de studenten binnen de opleiding systematisch in kaart. Deze analyses gebeuren typisch op het groepsniveau, eerder dan met individuele studenten. Op basis van de analyse kan de opleiding curriculumhervormingen implementeren of evalueren.

**Voorbeeld: *Academic Analytics* op Vlaams niveau**

Het Vlaamse ministerie van Onderwijs en Vorming voerde vanaf 2008-2009 het leerkrediet in om de kosten die samengaan met het te lang of zelfs onbeperkt studeren tegen te gaan. De analyses van de studietrajecten van studenten voor en na de invoering van het leerkrediet maken het mogelijk om de impact en het effect van de maatregel na te gaan. Deze analyses gebeuren typisch op Vlaams niveau (alle studenten in het hoger onderwijs), eerder dan op het niveau van de instellingen, de opleidingen of de individuele studenten.

– ***Learning Analytics*** focust op de onderwijskundige uitdaging en op data-ondersteunde beslissingen: hoe kunnen gegevens gebruikt worden om het leren te verbeteren? Het gaat over het inzetten van de data, die niet '*big*' hoeven te zijn, om implementeerbare inzichten (*actionable insights*) aan te reiken voor mensen die leren, onderwijzen en begeleiden. Daarnaast worden in dit domein expliciet de technische, ethische en pedagogische dimensies geïntegreerd.

**Voorbeeld: Learning Analytics op het niveau van een opleidings-onderdeel (Gelan et al., 2018)**

Een exploratief onderzoeksproject van de UHasselt verzamelde en onderzocht data van vier opleidingsonderdelen *blended* en *distance learning* aan drie universiteiten. Op basis van berekende LA-metrieken werd voor elk opleidingsonderdeel het onlineleergedrag vergeleken van succesvolle (geslaagd op het eindexamen) en niet-geslaagde studenten. Specifieke kenmerken van het onlineleergedrag correleerden positief met studiesucces (bv. het aantal onlineleersessies, het aantal bekeken en actief uitgevoerde leeractiviteiten, doorklikken naar de theorie, aantal herhalingen van activiteiten). Werken op de dag voor de *assessment deadline* (*cramming*) had een negatief verband met studiesucces. De doelstelling van het exploratieve onderzoek was niet om causale relaties te vinden tussen onlineleergedrag en studiesucces. Onlineleren is immers maar een onderdeel van het volledige leerproces. Bovendien spelen andere factoren een rol in studiesucces. De correlaties helpen wel om het leerproces beter te begrijpen en gedurende het semester de leervoortgang te monitoren en mogelijke problemen te detecteren. *Processmining*-analyse van de leerdata maakte het verder mogelijk om onlineleerpatronen te identificeren en ze te verbinden met bepaalde studentenprofielen. Aan de hand van leerdashboards kregen studenten feedback over hun leergedrag en tips van succesvolle studenten over leerstrategieën.

In het kader van dit Standpunt is vooral het verschil tussen *Academic* en Learning Analytics van belang. Hoewel er geen strikte scheiding is tussen deze twee, zijn er toch aspecten die helpen om ze te typeren, met name het niveau waarop ze werken. Onderwijs speelt zich af op verschillende niveaus: het niveau van de interactie tussen een student, begeleider en docent (micro), het niveau van de organisatorische setting waarin het leren en begeleiden van studenten (en de ondersteuning ervan door een docent of begeleider) plaatsvinden (meso) en het niveau van de overkoepelende regelgeving en het beleid (macro).

Het is de ambitie van *Academic Analytics* om het **beleid, management en de administratie** van instellingen in het hoger onderwijs en de overheid de nodige data te geven, om strategische beleidsbeslissingen te nemen op het vlak van onderwijs. *Academic Analytics* is al het ware de *business intelligence* van onderwijsinstellingen. *Academic Analytics* richt zich dus eerder op het macroniveau van het onderwijs.

LA daarentegen ambieert in de eerste plaats om meer over het leerproces te weten te komen en dit te optimaliseren. LA richt zich dan ook tot de docenten, studieloopbaanbegeleiders en studenten zelf, en heeft de ambitie om het concrete leerproces of de begeleiding te beïnvloeden. LA spitst zich dus eerder toe op het microniveau van het onderwijs.

Recent maakt ook een nieuwe term opgang: *Student Analytics*. Dat wordt gezien als de brug tussen *Academic* en Learning Analytics. Net als LA ambieert *Student Analytics* afstand te nemen van een *'one-size-fits-all'*-oplossing en streeft het naar een persoonlijke benadering. Door beschikbare data te analyseren gaat *Student Analytics* na wat daarin voorspellend is voor studiegedrag of -succes. Als eenmaal



© Joris Snaet

de belangrijke factoren geïdentificeerd zijn, kunnen ze aan de basis liggen van een gedifferentieerde, meer persoonlijk en data-gebaseerde studentenbegeleiding. Het doelpubliek is hier in de eerste plaats de **studieloopbaanbegeleider en de student zelf.**

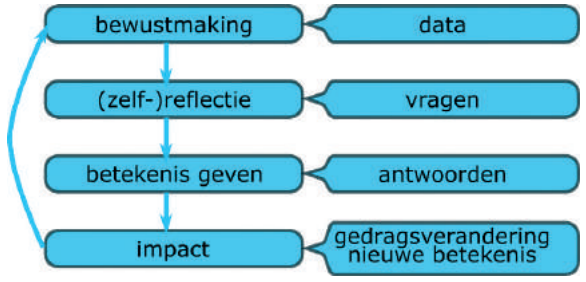
### E. Belanghebbenden

LA heeft heel wat belanghebbenden. Hieronder volgt een opsomming, met enkele voorbeelden van wat LA voor hen kan betekenen:

- **Individuele student:** reflectie over het eigen leer- en studeerproces, met inbegrip van de behaalde resultaten, in vergelijking met andere studenten (uit hetzelfde opleidingsonderdeel, dezelfde cursus of dezelfde opleiding: de zogenaamde *peers*) of vergelijkbare (al dan niet succesvolle) studenten uit het verleden.
- **Docenten en onderwijsondersteuners:** reflectie over het aangeboden leermateriaal en de gehanteerde leervormen, met bijvoorbeeld als doel de verbetering of optimalisatie hiervan, of de opvolging van de individuele student.
- **Studieadviseurs en studieloopbaanbegeleiders:** data-gebaseerde feedback en opvolging van studenten om de student zo goed mogelijk te adviseren en te coachen in het studieproces.
- **Docententeams, opleidingsverantwoordelijken:** verbetering en optimalisatie van de opleiding en het curriculum.
- **Onderzoekers:** LA biedt een waaier aan multidisciplinaire onderzoeks-mogelijkheden, zoals onderzoek naar het leer- en studeergedrag van studenten, de impact van de gebruikte leervormen, de effecten van data-gebaseerde feedback en data-analyse- of visualisatietechnieken.

### F. Kernaspecten

Figuur 1 toont het proces van LA: bewustmaking door het gebruik, de analyse en/of de visualisatie van data; gevolgd door zelfreflectie die vragen oproept; gevolgd door bewustmaking, waarbij antwoorden geformuleerd worden op de vragen; en ten slotte impact in de vorm van gedragsverandering of bijvoorbeeld nieuwe betekenissen die ontdekt worden (Verbert, Duval, Klerkx, Govaerts, & Santos, 2013).




*Figuur 1: Procesmodel voor Learning Analytics* (Verbert et al., 2013).

Figuur 2 toont de zes kritische dimensies van LA die houvast bieden binnen het multidisciplinaire LA-domein. Binnen elke dimensie worden ook enkele deelaspecten opgesomd. De zes dimensies zijn kritisch in de zin dat in elke toepassing van LA aandacht voor elk van de dimensies vereist is.




*Figuur 2: Zes kritische dimensies van Learning Analytics volgens Greller & Drachsler* (Greller & Drachsler, 2012).

De zes dimensies zijn: belanghebbenden, doelstellingen, data, instrumenten, externe en interne beperkingen. Bij de **belanghebbenden** horen zowel de data-betrokkenen (*data subjects*) als de data-klanten (*data clients*). De data-betrokkenen zijn de personen over wie data worden verzameld. Dikwijls zijn dit studenten, maar ook docenten of bijvoorbeeld studieadviseurs kunnen het onderwerp zijn van de data. De data-klanten zijn de begunstigden van de data, diegenen op wie de concrete LA-interventie is gericht en van wie er gehoopt wordt dat er reflectie, inzicht of gedragsverandering wordt uitgelokt.

Het is belangrijk om de **doelstellingen** van LA expliciet te maken. Het doel kan onderzoeksmatig zijn, bijvoorbeeld beter begrijpen hoe onderwijsleerprocessen verlopen (zie II.A). Ook de optimalisatie en personalisatie van de leeromgeving (zie II.B) of de data-gebaseerde ondersteuning van de studieloopbaanbegeleiding (zie II.C) kunnen het doel zijn. Binnen deze drie toepassingsdomeinen is er dikwijls ook sprake van een doelstelling die is gerelateerd aan reflectie of predictie. Reflectie is de kritische zelfevaluatie van een data-klant op basis van de ter beschikking gestelde data. Het kan zowel gaan over eigen data (bv. studenten

die nadenken over hun leergedrag op basis van hun waargenomen activiteiten op het leerplatform) als over data van andere belanghebbenden (bv. docenten die nadenken over het online aangeboden leermateriaal op basis van de activiteit van studenten op het leerplatform).

Aan de basis van LA liggen steeds **data** (zie III). Hoewel LA dikwijls geassocieerd wordt met activiteiten op onlineleerplatformen beperkt het zich daar zeker niet toe. Andere gegevensbronnen zijn onder andere demografische data, zelf-gerapporteerde data in al dan niet gevalideerde vragenlijsten, academische resultaten, studievoortgangsresultaten, aanwezigheden bij contactmomenten, toegang tot de instellingsgebouwen of bibliotheek, of zelfs hersenactiviteit, oogbewegingen en hartslag bij bepaalde lesactiviteiten.

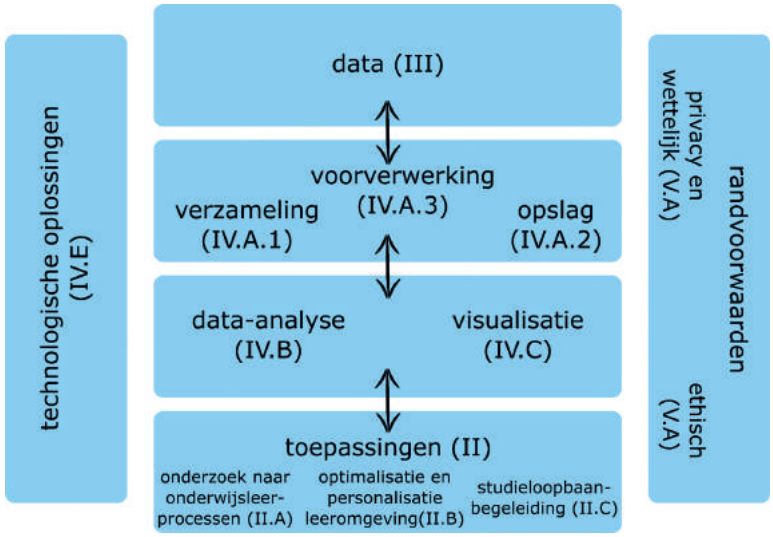
LA steunt op heel wat onderliggende **instrumenten**, niet enkel de pedagogische theorieën en inzichten inzake leren maar ook data-analysetechnieken uit domeinen als *Educational Data Mining*, psychometrie en statistiek (IV.B.1) of machinaal leren (IV.B.2). Daarnaast zijn ook instrumenten voor dataverzameling, -opslag, -voorverwerking (IV.A) en visualisatie (IV.C) van belang.

Enkele van de **externe randvoorwaarden** die LA beïnvloeden zijn: privacy- en wettelijke aspecten (V.A), ethische aspecten (V.B), organisatorische, praktische en procesmatige aangelegenheden.

Bij de **interne beperkingen** staan vooral competentie en acceptatie centraal. De data-klanten moeten de nodige competenties hebben om met de gerapporteerde gegevens aan de slag te kunnen gaan. Van data-klanten wordt dikwijls meer geëist dan een basis datageletterdheid en ethisch begrip. Nagaan welke competenties vereist zijn en de interventie afstemmen op de beschikbare competenties is dan ook noodzakelijk. Daarnaast zal LA enkel succesvol zijn als er ook acceptatie is. Dat is niet vanzelfsprekend, zelfs als de nodige competenties aanwezig zijn en alle ethische, privacy- en gegevensbeschermingsaspecten in overweging zijn genomen.

### G. Opbouw van dit standpunt

Binnen het bestek van dit Standpunt kunnen we onmogelijk een gedetailleerd beeld geven van het brede LA-domein. Dit Standpunt gaat dieper in op enkele kernaspecten, in het bijzonder op wat van belang is in de context van het Vlaamse hoger onderwijs. De overzichtsfiguur hieronder laat zien hoe het Standpunt is opgebouwd en welke onderdelen aan bod komen in de vier grote onderdelen: Waarom? (II), Data (III), Hoe? (IV) en Randvoorwaarden (V). Binnen elk thema stellen we aanbevelingen voor.


technologische oplossingen (IV.E)
data (III)
voorverwerking (IV.A.3)
verzameling (IV.A.1)
opslag (IV.A.2)
data-analyse (IV.B)
visualisatie (IV.C)
toepassingen (II)
onderzoek naar onderwijsleer-processen (II.A)
optimalisatie en personalisatie leeromgeving(II.B)
studieloopbaan-begeleiding (II.C)
randvoorwaarden
privacy en wettelijk (V.A)
ethisch (V.A)

# II. Waarom?

Dit onderdeel behandelt drie mogelijke globale toepassingen van LA op microniveau: onderzoek naar onderwijsleerprocessen; optimalisatie en personalisering van de leeromgeving; studieloopbaanbegeleiding. Zoals eerder vermeld zijn het meso- en macroniveau veeleer het voorwerp van *Academic Analytics.*

## A. Onderzoek naar onderwijsleerprocessen

### 1. Traditionele benadering

Onderzoek van onderwijsleerprocessen heeft een rijke geschiedenis. Het is lang gebaseerd geweest op gevalsstudies, anekdotische aanwijzingen en zelfreflectie. Het meer systematische onderzoek maakt ruwweg gebruik van twee benaderingen.

Een eerste benadering houdt in dat er vragen gesteld worden aan studenten en docenten met behulp van onder andere interviews, vragenlijsten en focusgroepen. Dat levert interessante inzichten in de perceptie en zelfgerapporteerde gedragingen van sleutelbetrokkenen op. Hoewel zeker in het geval van onderzoek naar percepties het bevragen van de betrokkenen zelf vanzelfsprekend is, is het dat veel minder bij het onderzoek naar de gedragingen die studenten en docenten stellen. Studenten en docenten blijken slechts in geringe mate in staat te zijn correct te rapporteren wat ze gedaan hebben, en nog minder wat ze zullen doen (Veenman, 2015). Het mag dan ook niet verwonderen dat er heel wat vragen rezen bij de uitkomsten van onderzoek naar door studenten of docenten gehanteerde en/of geprefereerde leer- en onderwijsstrategieën. De validiteit van vragenlijstonderzoek (evenals van interviews en focusgroepen) blijkt erg beperkt.

De tweede benadering gaat uit van (quasi-)experimenteel onderzoek naar het effect van specifieke interventies, zoals het gebruik van specifieke media, werkvormen, taken- en evaluatievormen, in veelal specifieke contexten en omstandigheden. Om onderzoeksredenen houdt dit onderzoek de variabiliteit in de verschillende condities sterk onder controle en om praktische redenen ging het veelal om eerder kortlopende interventies. De jongste jaren is in dit verband de aandacht voor *design-based*-onderzoek toegenomen, waarbij in meer realistische omstandigheden wordt getracht het effect van interventies systematisch op te volgen.

Toch zijn onderzoekers van onderwijsleerprocessen lang onvoldaan gebleven. Hun onderzoek heeft nood aan een direct zicht op de feitelijke leergedragingen van studenten. Zo kan duidelijk worden welke problemen zij ervaren en hoe ze precies omgaan met de hulp die in de (al dan niet technologisch ondersteunde) leeromgeving wordt geboden. Lang is getracht via 'hardop denken' zicht te krijgen op dergelijke leerprocessen. Dat lijkt vooral interessant in het geval dat een

onderzoek meer zicht wil krijgen op hoe een individuele lerende een probleem oplost of in een digitale leeromgeving aan de slag gaat. De aanpak heeft als nadeel dat hij moeilijk toepasbaar is in realistische omstandigheden en zich uitsluitend richt op het individu (waardoor onderzoek naar samenwerkend leren hier geen gebruik van kan maken). Daarnaast werden veel vragen gesteld over de impact die het hardop denken heeft op het proces dat juist wordt onderzocht. Jarenlang werd dan ook gepleit voor zogenaamde *unobtrusive measures*: metingen van onderwijsleerprocessen die het onderwijsleerproces zelf zo min mogelijk verstoren.

## 2. Opportuniteiten

Het gebruik van steeds meer technologische hulpmiddelen in het onderwijs heeft deze *unobtrusive measures* toegankelijker gemaakt voor onderwijsonderzoekers. Het gebruik van *log files* voor effectiviteitsonderzoek is niet nieuw: ze werden bijvoorbeeld al in de jaren 1990 ingezet in het domein van Computer Assisted Language Learning (Gelan et al., 2018). De huidige technologische ontwikkeling maakt de dataverzameling echter veel minder tijdrovend en meer haalbaar op grotere schaal. Bovendien is het nu meer en meer mogelijk om aan de hand van niet alleen *log files* maar bijvoorbeeld ook van *wearables* na te gaan wat studenten in technologisch ondersteunde leeromgevingen de facto doen. Dit versterkt tal van mogelijkheden, waarvan dit onderdeel er maar enkele schetst. Tegelijk neemt het onderzoek naar technologische hulpmiddelen de jongste decennia een specifieke plaats in in het domein van het pedagogisch onderzoek. Veel onderzoekers hebben moeten vaststellen dat theoretisch krachtige ondersteuningsmiddelen in leeromgevingen niet altijd het beoogde effect sorteren (zeker niet in realistische omstandigheden). Onderzoek naar het gebruik van diverse hulpmiddelen in een digitale leeromgeving toonde aan dat een behoorlijke groep van studenten de hulpmiddelen nauwelijks gebruiken, sommigen ze heel veel gebruiken en anderen er strategisch mee omgaan (Lust, 2012). Studenten die met succes hun gebruik afstemden op de vereisten van de cursus presteerden beter. Metingen van het gebruik van deze technologische hulpmiddelen tijdens het onderwijsleerproces bieden dan ook een belangrijke input voor LA-onderzoek dat het proces verder tracht te begrijpen.

Onderzoek toonde intussen aan dat enerzijds aan de hand van *log files* het studiegedrag van studenten in kaart kon worden gebracht, maar anderzijds ook dat er door fluctuaties tijdens de looptijd van een cursus veel meer nood is aan longitudinaal onderzoek. Vervolgonderzoek toonde aan dat over leeromgevingen heen relatief stabiele groepen van studenten kunnen worden onderscheiden op grond van hun zelfregulatiegedrag (Van Laer & Elen, 2018). Dit is beloftevol voor interventieonderzoek (zie II.B) waarin wordt getracht minder adequaat zelfregulatiegedrag bij te sturen.

### 3. Uitdagingen

Al dit onderzoek duidt op het potentieel van LA en geeft aan dat er belangrijke nieuwe inzichten kunnen worden verworven. Tegelijkertijd roept het gebruik van LA (en meer in het algemeen van *unobtrusive measures*) tal van nieuwe vragen op. Een eerste reeks vragen betreft privacy- (V.A) en ethische (V.B) aspecten. Daarnaast roept de grote hoeveelheid data die dikwijls volgen uit LA nieuwe vragen op die onderzoeksmethodologisch van aard zijn. Hoe kan een dergelijke dataset worden geïnterpreteerd en hoe kan er voldoende helder over de behandeling van de data worden gecommuniceerd om de repliceerbaarheid van het onderzoek te faciliteren? Daarnaast is de vraag naar de betekenis van de data essentieel. Een klik op 'volgende' interpreteren als 'het hebben gelezen van een webpagina' is mogelijk al te naïef. Vandaar dat meerdere elkaar bevestigende data nodig zijn.

In de analyse van de data onderscheiden we binnen het onderzoek naar onderwijsleerprocessen ruwweg een datagedreven en een theoriegedreven aanpak. In het geval van een datagedreven aanpak worden geen vooronderstellingen gemaakt en krijgen de data de kans te 'spreken'. Tegelijk dient aangestipt te worden dat in de algoritmes die dit 'spreken' toelaten natuurlijk tal van veronderstellingen worden gemaakt die niet steeds worden geëxpliciteerd. In het geval van een theoriegedreven aanpak worden data gebruikt om bepaalde theoretische veronderstellingen te toetsen. In dit geval is het gevaar van een selectief gebruik van data en het niet toetsen van alternatieve theoretische assumpties reëel. Het zal een uitdaging zijn voldoende aandacht te blijven spenderen aan deze twee verschillende aanpakken, maar vooral ook het onderzoek te proberen samen te brengen en te verenigen rond de gemeenschappelijke doelstelling.

---

*Aanbeveling 1: Dankzij nieuwe technologieën zijn er meer data van leergedragingen uit 'realistische' omstandigheden maar ook uit onlineleeromgevingen beschikbaar. Onderzoek op basis van onder andere concrete gedragsdata zal bestaande inzichten kunnen versterken maar ook nieuwe inzichten opleveren op basis van de nieuwe beschikbare data en technologieën. Die zullen mogelijks zelfs bestaande kennis ter discussie stellen. Blijvend inzetten op dit onderzoek is dan ook noodzakelijk.*

---

## B. Optimaliseren en personaliseren van de leeromgeving

LA vormt het snijpunt tussen onderzoek naar onderwijsleerprocessen en de onderwijspraktijk. Eén aspect van die praktijk is de aanpassing van de technologieondersteunde leeromgeving, afhankelijk van wat data over onderwijsleerprocessen vertellen over de leeromgeving en de gebruikers die erin actief zijn. Het doel is het ontwerp van de leeromgeving te optimaliseren en zodoende bij

de gebruiker een maximaal effect te bereiken. Bij een doorgedreven aanpassing aan individuele gebruikers is de leeromgeving **gepersonaliseerd**: er zijn evenveel leeromgevingen als gebruikers.

### 1. Wat, welke databron, hoe en waarom?

Bij de optimalisatie en personalisatie van de leeromgeving rijst de vraag **wat** er aangepast moet worden, op basis van **welke databron** en **hoe** (inclusief wanneer) (Vandewaetere, Desmet, & Clarebout, 2011). Bovendien moet bij elke aanpassing het **waarom** duidelijk zijn.
Het **wat** omvat in principe alle elementen die in het ontwerp van leeromgevingen in beschouwing worden genomen. Zo kan de volgorde van leertaken aangepast worden, maar ook de inhoud van de taken (bv. al dan niet met vereenvoudigde woordenschat), hun verschijningsvorm (bv. tijdsdruk, opmaak) of de leerondersteuning (zoals hints, bijkomende uitleg, feedback of ondertitels bij het bekijken van een video in een vreemde taal). Bij samenwerkend leren kunnen op basis van data-analyse gericht – homogene of net heterogene – groepen worden samengesteld.

Aanpassingen aan de leeromgeving zijn gebaseerd op de analyse van data uit verschillende **databronnen** (III.A): achtergrondkenmerken van gebruikers (bv. persoonlijkheid, leerstoornis, geslacht), gemeten activiteit op bijvoorbeeld leerplatformen, contextuele informatie over de leersituatie (programma, tijd, plaats of zelfs temperatuur of zuurstofgehalte in de ruimte). Interacties tussen studenten en taken leveren informatie op over de moeilijkheid of attractiviteit van leertaken of de (evoluerende) vaardigheden of interesses van studenten (zie ook IV.B.1). Interacties tussen studenten onderling of tussen studenten en experts geven dan weer informatie over de intensiteit of kwaliteit van samenwerkend leren; interacties met docenten kunnen iets leren over de effectiviteit van onderwijsinterventies. Meer in het algemeen kunnen data over de interacties tussen de studenten en hun omgeving informatie opleveren over de mate waarin de student in staat is het eigen leerproces te reguleren.

De optimalisatie van de leeromgeving kan volgen uit de analyse van een of meer gegevensbronnen. Zo kan de samenstelling van groepen gemaakt worden rekening houdend met persoonlijkheidskenmerken van de studenten, de kwaliteit van groepsprocessen en de taakuitkomsten, zoals geobserveerd tijdens eerdere (online- of technologie-gemedieerde) groepstaken, of beide.

De **hoe**-vraag slaat op de processen die worden gevolgd om de leeromgeving aan te passen. Die kunnen ruimte laten voor een tussenkomst door een docent, student of onderwijsontwerper, of ze kunnen volledig geautomatiseerd zijn. Daarnaast bepaalt een tijdsdimensie hoe de leeromgeving wordt aangepast: tussen langere periodes of leersessies in (met of zonder menselijke tussenkomst),

of *at runtime* (typisch geautomatiseerd). Zo kunnen docenten in een MOOC na een semester via een dashboard opvragen met welke oefeningen de studenten de grootste problemen hadden, en op basis van deze analyse de oefeningen aanpassen. Deze vorm van aanpassen is statisch en laat veel ruimte voor controle voor de beheerder. In een *intelligent tutoring system*, de allicht het meest tot de verbeelding sprekende gepersonaliseerde leeromgeving en heilige graal van de educatieve technologie, leidt de technologie de dans. *Learner models* houden voor elke individuele student de voortgang en andere parameters bij. Een *reasoner* kan na elke interactie beslissen welke taken wanneer en onder welke vorm aan welke lerende worden voorgeschoteld, of wat de meest gepaste ondersteuning is. Sommige systemen maken de leerdermodellen transparant voor de lerende en/of docent (*open learner models*) en maken het mogelijk om zelf in te grijpen in het adaptatieproces. Dit laatste veronderstelt dat de student over voldoende metacognitieve vaardigheden beschikt om die modellen te interpreteren en in te schatten wat de gevolgen zijn van een tussenkomst.

Ten slotte, en niet onbelangrijk, is er de vraag **waarom** leeromgevingen aangepast worden. Dit leidt opnieuw naar het onderwijskundig onderzoek (II.A). Het duidelijke antwoord op de overkoepelende vraag bestaat uit twee luiken. Enerzijds is elke leeromgeving het resultaat van een ontwerpproces waarin keuzes werden gemaakt, en kan zo'n leeromgeving dus ook verbeterd worden. LA biedt de onderwijsontwerpers informatie om de kwaliteit van de leeromgeving te verhogen. Anderzijds verschillen studenten van elkaar en moet de leeromgeving hier optimaal op inspelen. LA biedt de mogelijkheid om elke student de best mogelijke variant van de leeromgeving/cursus aan te bieden. Maar ook voor elke specifieke aanpassing moet de waarom-vraag steeds vooropstaan. Als het doel is om bijvoorbeeld de motivatie van studenten te verhogen door middel van een systeem van beloningen, dan is het essentieel dat hiervoor een beroep wordt gedaan op wetenschappelijke theorieën en evidentie rond (leer)motivatie.

### 2. Opportuniteiten

LA biedt het grote voordeel dat het leerproces meer nauwgezet kan worden opgevolgd en heeft daarom het potentieel om heel wat inzichten te verkrijgen om de leeromgeving te optimaliseren en personaliseren.

LA zorgt voor de noodzakelijke data om theorieën over onderwijsprocessen aan de praktijk te toetsen en te verfijnen, en leeromgevingen op een methodologisch betrouwbare en valide manier te verbeteren.

### 3. Uitdagingen

Een uitdaging bij het ontwerp van gepersonaliseerde leeromgevingen is dat alle betrokken actoren (student, docent, begeleider, technologie en onderwijsontwerper)

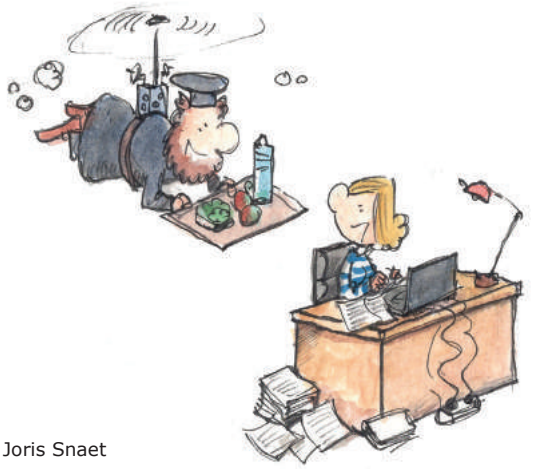
op eenzelfde lijn moeten staan tijdens het leerproces. De docent en de student moeten weten wat de technologie over hen 'weet'; de technologie moet weten welke acties de docent uitvoert om studenten te ondersteunen en hoe studenten met elkaar samenwerken. Wanneer de rollen van de docent, studenten en technologie op elkaar zijn afgestemd en het leerproces optimaal ondersteunen, is er sprake van *synergistic scaffolding* (Tabak, 2004). Een andere uitdaging is de algoritmes van gepersonaliseerde leeromgevingen transparant te maken voor docent en student. Algoritmes zijn niet neutraal maar werden ontworpen door mensen; in het ontwerpproces kunnen keuzes gemaakt zijn die voor sommige belanghebbenden nadelig zijn. De betrokkenen hebben het recht vooraf nuttige informatie te krijgen over de onderliggende logica en over het belang en de verwachte gevolgen van die verwerking. Ook de wettelijke (V.A) en ethische dimensies (V.B) zijn dus belangrijk.

Daarnaast vraagt de realisatie van effectieve gepersonaliseerde leeromgevingen een radicaal **interdisciplinaire samenwerking** van onderwijskundigen, cognitieve en motivatiepsychologen, specialisten in de desbetreffende leerdomeinen, technologen, datawetenschappers, privacydeskundigen en ethici.

---

*Aanbeveling 2: LA biedt heel wat potentieel voor de optimalisatie en personalisatie van de leeromgeving. Dit kan leiden tot een verbetering van het leerproces van studenten in het Vlaamse hoger onderwijs. De waarom-vraag, beantwoord vanuit onderwijskundig onderzoek, mag hierbij echter niet uit het oog verloren worden. Daarnaast vereist een effectieve realisatie een radicaal interdisciplinaire samenwerking van onderwijskundigen, cognitieve en motivatiepsychologen, specialisten in de diverse leerdomeinen, technologen, datawetenschappers, privacydeskundigen en ethici.*

---



© Joris Snaet

## C. Studieloopbaanbegeleiding

Instellingen in het hoger onderwijs hopen LA in te zetten voor een veelheid aan *doelen* op het vlak van studieloopbaanbegeleiding, zoals een betere studentenervaring, grotere studententevredenheid, een betere studiekeuze, de verhoging van studiesucces, de vermindering van de studieduur, minder drop-out, meer feedback tijdens het leerproces, een betere begeleiding naar bestaande ondersteuning en trainingen, data-gebaseerde studieloopbaanbegeleiding, identificatie en gerichte ondersteuning van studenten uit specifieke doelgroepen, een beter begrip van factoren die belangrijk zijn in studiesucces, een meer persoonsgerichte aanpak of studenten helpen bij het stellen van doelen. Internationaal zijn er instellingen in het hoger onderwijs die LA gebruiken voor studentenbegeleiding. Kunnen de Vlaamse instellingen deze toepassingen overnemen? Wat zijn de opportuniteiten en ook uitdagingen van LA in studieloopbaanbegeleiding die relevant zijn voor het Vlaamse hoger onderwijs?

### 1. Het belang van context

Eerst en vooral is het belangrijk te erkennen dat de context waarin instellingen in het hoger onderwijs werken de doelen op het vlak van studieloopbaanbegeleiding sterk beïnvloeden. Het Vlaamse hoger onderwijs typeert zich door een open toegang zonder selectie (met uitzondering van genees- en tandheelkunde en bepaalde kunstopleidingen) en lage inschrijvingsgelden. Daarnaast beïnvloeden de Vlaamse onderwijsdecreten rond onder meer studievoortgang en het financieringsmodel de visie en beslissingen van de instellingen. Instellingen in het hoger onderwijs van andere landen opereren dikwijls in een sterk verschillende context, waardoor de doelstellingen die ze mede willen realiseren met LA sterk verschillen. In de Angelsaksische landen focussen heel wat LA-interventies op 'retentie', het behoud van studenten: de hogeronderwijsinstellingen zijn immers vrij om studenten te selecteren, maar de overheid legt vervolgens via het financieringsmodel maxima op voor het aandeel studenten dat de opleiding mag verlaten zonder diploma. Het is duidelijk dat deze focus op 'retentie' niet houdbaar is in de Vlaamse hogeronderwijscontext, waar er geen selectie is 'aan de poort': hier kan een snelle heroriëntatie net wenselijk zijn naar een opleiding die beter is afgestemd op de vaardigheden en verwachtingen van de student.

Vlaamse hogeronderwijsinstellingen hebben daarnaast niet noodzakelijk dezelfde data beschikbaar als andere instellingen (III.C) en de onderwijsorganisatie en -visie kan per instelling ook grondig verschillen. Alvorens LA-interventies te willen transfereren naar de Vlaamse context is het daarom essentieel na te gaan of de context, beschikbare data, onderwijsorganisatie en -visie en de doelstellingen voldoende vergelijkbaar zijn.

---

*Aanbeveling 3: Context is belangrijk. Bekijk steeds of bestaande LA-initiatieven en -toepassingen wel overdraagbaar zijn naar de eigen context.*

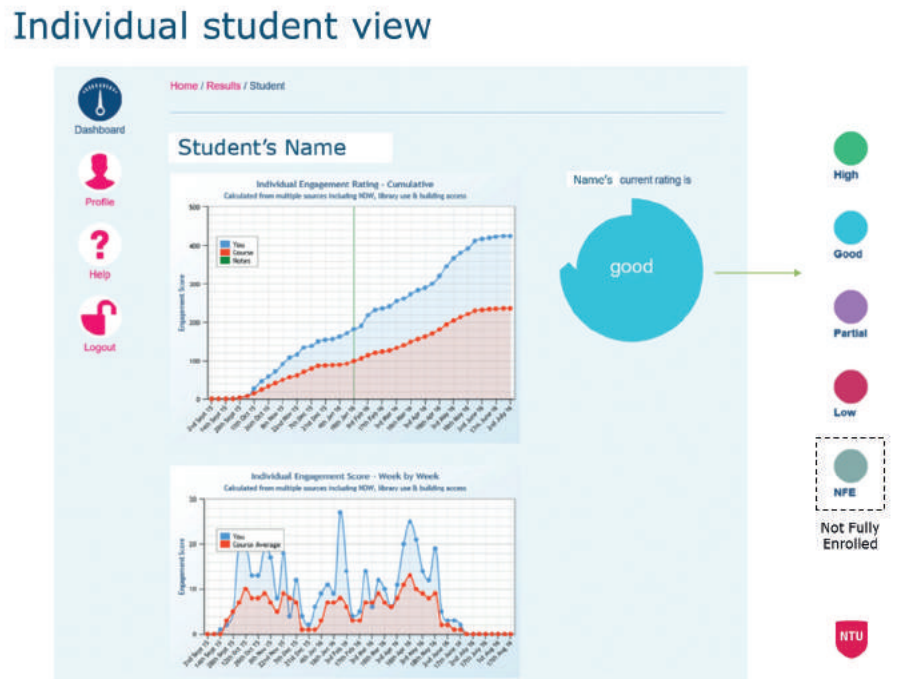
---

## 2. Learning dashboards voor studieloopbaanbegeleiding

Bij LA voor studieloopbaanbegeleiding komt meestal eerst het beeld voor ogen van een dashboard dat gericht is op feedback aan studenten, een zogenaamd **studentgericht dashboard** dat feedback geeft of raadgevingen aanbiedt met betrekking tot bijvoorbeeld studiekeuze, studie-inspanningen, doelstellingen of leermateriaal.

**Voorbeeld van een studentgericht dashboard: het *Engagement dashboard* van Nottingham Trent University (UK)**

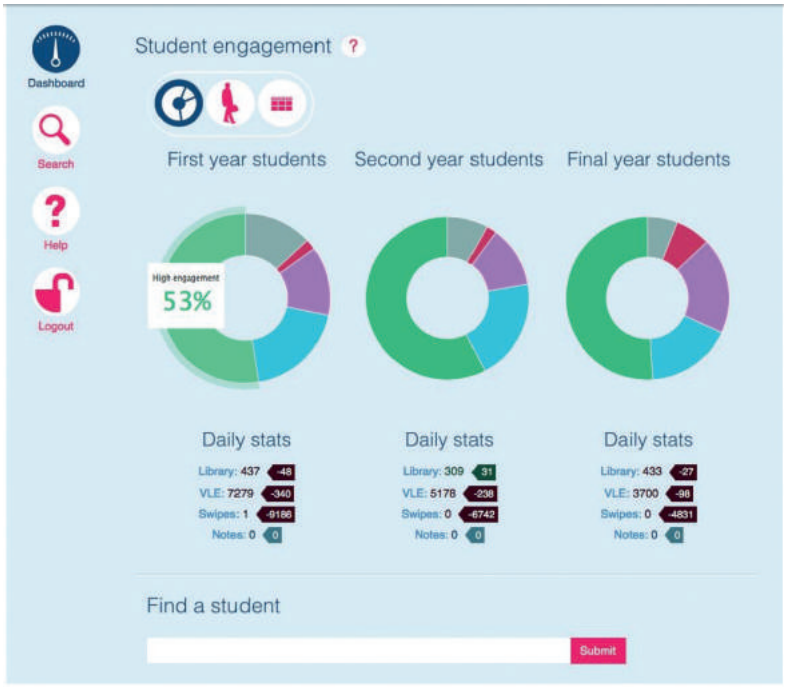
Het studentendashboard van Nottingham Trent University geeft studenten een blik op hun *student engagement*, de studiegerelateerde activiteiten die de student onderneemt. De studenten kunnen hun inzet en betrokkenheid ('*engagement*' in het Engels) over de tijd analyseren en vergelijken met andere studenten. Bijkomend krijgen studenten op basis van historische data een samenvattende score, gaande van hoge betrokkenheid tot lage betrokkenheid. Om de betrokkenheid te kwantificeren gebruikt het algoritme de activiteit op het leerplatform, het indienen van opdrachten, de aanwezigheid in contactmomenten en het binnengaan in de universitaire gebouwen met studentenkaart, het uitlenen van materiaal uit de bibliotheek en het gebruik van elektronische boeken.




Er zijn echter ook **personeelsgerichte dashboards** voor studieloopbaan-begeleiding: dashboards die in eerste instantie gericht zijn op de studieloop-baanbegeleiders en hen ondersteunen bij het begeleidingsproces. Denk aan een

**Voorbeeld van een personeelsgericht dashboard: het *Engagement dashboard* van Nottingham Trent University (UK)**

Nottingham Trent University heeft naast het studentgericht dashboard ook een dashboard voor zijn tutoren. Elke tutor volgt een 50-tal studenten op. Via het dashboard kan de tutor niet alleen de gegevens van de studenten opvragen maar ook hun inzet in en betrokkenheid bij de opleiding nagaan. Het dashboard heeft ook een *early warning system* en stuurt de tutor waarschuwingen als een student een aanhoudend lage inzet en betrokkenheid aan de dag legt. Ten slotte kan de tutor notities maken in het systeem en de student doorverwijzen naar andere diensten binnen de universiteit.

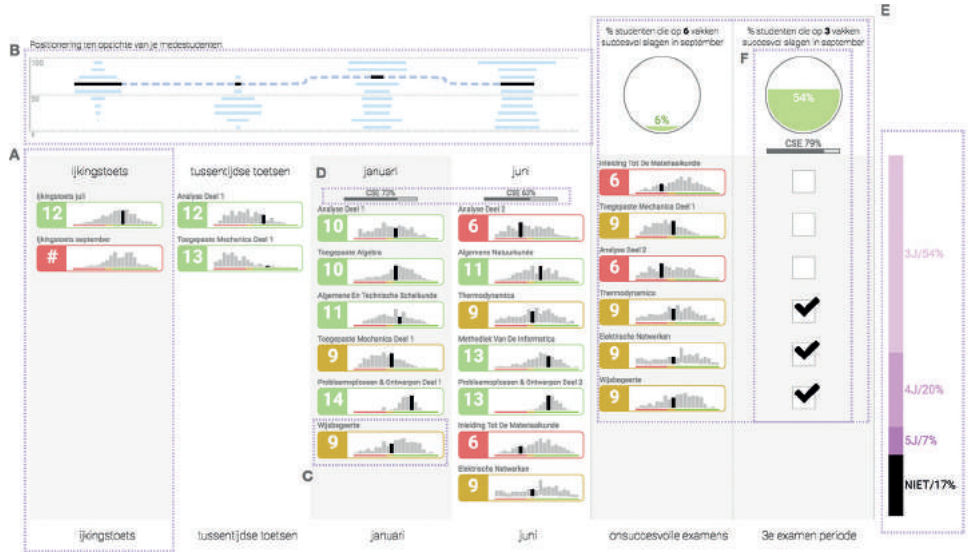



meldingssysteem (*early warning system*) dat een studieloopbaanbegeleider op de hoogte brengt van een student die geen leeractiviteiten aan de dag lijkt te leggen en vervolgens een meer gedetailleerd zicht biedt op de data van de student.

Een belangrijke recente evolutie zijn *dashboards* die de veelal live **interactie tussen studieloopbaanbegeleider en student** ondersteunen.

**Voorbeeld van een dashboard voor de interactie tussen de studieloopbaan-begeleider en de student: LISSA (KU Leuven, België)**

LISSA is een dashboard dat ontwikkeld is binnen de KU Leuven om de dialoog tussen een student en studietrajectleider te faciliteren (Charleer, Vande Moere, Klerkx, Verbert, & De Laet, 2017). Het dashboard beperkt zich tot het tonen van resultaten van studenten op verschillende evaluatieperiodes, inclusief de ijkingstoets, tussentijdse toetsen en de examens in de examenperiodes. In deel A van het dashboard worden de resultaten van de verschillende evaluatiemomenten in detail weergegeven. De resultaten van de studenten zijn ook gecontextualiseerd: studenten kunnen in het histogram naast hun resultaat voor een bepaald opleidingsonderdeel zien hoe hun scores zich verhouden tot de scores van een anonieme groep medestudenten. Boven het overzicht van de resultaten zien de studenten een globale positionering tegenover medestudenten (Deel B) op basis van het hele resultaat in een bepaalde evaluatieperiode. Deel G toont het aantal jaren dat studenten met vergelijkaardige resultaten nodig hebben om het bachelordiploma te behalen. Deel E van het dashboard bevat een specifieke module die de keuze van opleidingsonderdelen voor herkansing in de derde examenperiode ondersteunt.




## 3. Opportuniteiten

Naast de algemene opportuniteiten van LA zijn er ook enkele die specifiek zijn voor het gebruik van LA bij studieloopbaanbegeleiding.

LA kan helpen om studenten en docenten de data die de onderwijsinstelling over hen verzamelt op een **transparante** maar ook **nuttige wijze** ter beschikking te hebben. Studenten hebben wettelijk gezien immers het recht om deze data in te kijken (zie ook V.A). Eerder dan hen meteen een eenvoudige maar in de praktijk onbruikbare print-out te geven van de data kan een *learning dashboard* de informatie samenvatten en ordenen in relevante categorieën, zoals achtergrondgegevens,

persoonsgegevens, leeractiviteit op het leerplatform, studieresultaten, opgemeten leer- en studeervaardigheden, uitleengeschiedenis in de bibliotheek, toegang tot universitaire gebouwen en afspraken met studieloopbaanbegeleiders.

Een volgende stap in dit proces is de inzet van deze data om studenten **meer continu feedback** te geven tijdens het leerproces. Hierbij kan ook een vergelijking met medestudenten gemaakt worden, zowel studenten uit eenzelfde cohorte (bv. andere studenten in hetzelfde jaar in de opleiding) als uit eerdere cohorten (studenten die vroeger in een vergelijkbare studiesituatie zaten). De aangeboden informatie kan studenten uitdagen om te reflecteren over hun studiegedrag, maar hen ook stimuleren om dat studiegedrag aan te passen. Op basis van deze gegevens kunnen **automatische aanbevelingen** volgen, zoals advies bij het selecteren van keuzevakken en van de studierichting, aangepast studiemateriaal, nieuwe doelstellingen, relevante trainingen of een persoonlijk gesprek met een studieloopbaanbegeleider.

Daarnaast kan LA studieloopbaanbegeleiders ondersteunen bij **data-gebaseerde beslissingen en adviezen.** LA kan helpen om factoren te identificeren die belangrijk of voorspellend zijn voor studiesucces. Op basis hiervan kunnen studieloopbaanbegeleiders doelgroepen identificeren en een aangepaste begeleiding opzetten. Het kan hen de nodige data geven over de student, zodat er in het eigenlijke contact meer tijd vrijkomt om de persoonlijke aspecten te bespreken, die niet zichtbaar zijn in de data.

Vlaamse instellingen in het hoger onderwijs hebben een uniek menselijk kapitaal: daar horen ook de studieloopbaanbegeleiders bij. Zij hebben de nodige expertise en ervaring om binnen een multidisciplinair team LA-initiatieven te ontwerpen en implementeren. Ze brengen broodnodige kennis van de context en de huidige problematiek in, die de relevantie van de ontworpen oplossing ten goede komt. Daarnaast zorgt hun betrokkenheid voor de bewaking van het ethische karakter, het nodige draagvlak en uiteindelijk ook de acceptatie van het resultaat.

LA heeft aangetoond het potentieel te hebben om de tevredenheid te verhogen en studieresultaten te verbeteren, en zelfs om met kleine proefprojecten een verhoogd bewustzijn te creëren rond een data-gebaseerde aanpak, hierrond ook de discussie te starten en zelfs een aanknoping te zijn van een instellingsbrede strategie. Er kunnen ook positieve neveneffecten optreden, zoals een cultuuromslag naar het maken van data-gebaseerde beslissingen in studentenbegeleiding.

## 4. Uitdagingen

Naast de algemene uitdagingen van LA zijn er ook enkele die specifiek zijn voor het gebruik van LA bij studieloopbaanbegeleiding.

Eerst en vooral zijn de beschikbare data beperkt: heel wat relevante leeractiviteiten of persoonlijke aspecten die het leren beïnvloeden laten immers geen digitaal

spoor na, en niet alle vakken bieden digitale ondersteuning. Het beeld dat op basis van de data van een student ontstaat is per definitie onvolledig. Er bestaat ook een gevaar voor depersonificatie, hoewel er dankzij LA ook net meer ruimte en tijd kan zijn voor persoonlijke aspecten (V.B). Daarnaast ontstaat er wrijving tussen privacy (zie ook V.A) en een persoonlijke aanpak.

---

*Aanbeveling 4: Zet Learning Analytics in om de persoonlijke aanpak van studieloopbaanbegeleiding te versterken, met ruimte voor de inbreng van persoonlijke omstandigheden.*

---

LA vindt dikwijls dataverbanden die de betrokkenen niets nieuws leren: het is geen verrassing dat studenten die meer actief zijn op het leerplatform ook beter presteren. De vraag is echter ook of de gevonden verbanden wel enige causaliteit vertonen of een aanknopingspunt bieden voor begeleiding. Is het relevant studenten te adviseren om meer actief te zijn op het leerplatform? Zal dit leiden tot meer studiesucces? Bij de interpretatie is de kennis van de docent nodig, die de context en de gebruikte leermiddelen kan kaderen. De gevonden verbanden zijn ook inherent onzeker, waarbij het gevaar ontstaat dat de betrokkenen deze onzekerheden gaan interpreteren als kansen: in een stelling als "student X heeft 76% kans op slagen" wordt het onderwijsproces afgedaan als een kansspel waaraan de student onderworpen is. Daarnaast is er ook de vraag of geïdentificeerde doelgroepen en daarvoor ontworpen interventies wel ethisch zijn (V.B).

---

*Aanbeveling 5: Betrek de belanghebbenden vanaf de eerste stap van de ontwikkeling bij Learning Analytics. Vergeet naast onderwijsdeskundigen, studenten en docenten ook de studieloopbaanbegeleiders niet: zij beschikken over de broodnodige expertise inzake de context, de huidige problematiek en individuele en ethische studentenbegeleiding.*

---

Zoals eerder vermeld is de concrete context waarin de onderwijsinstelling werkt van groot belang voor LA. Dit zet meteen een rem op de schaalbaarheid en transfereerbaarheid van vele LA-initiatieven.

LA lijdt onder onrealistische verwachtingen. Het kan slechts één onderdeel van een instellingsbrede strategie voor studentenbegeleiding zijn.
Niet alle tot nu toe uitgerolde LA-initiatieven hebben geleid tot positieve resultaten. Waargenomen ongewenste effecten zijn bijvoorbeeld de demotivering van studenten, het negatief bijstellen van doelen en het onnodig uitschrijven uit de opleiding. Het is dan ook essentieel om in een goede opvolging en monitoring te voorzien.

LA op een doelgerichte manier inzetten is niet vanzelfsprekend en vereist de nodige opleiding van de betrokkenen. Het is daarom belangrijk om vanaf de start

in de nodige professionaliseringstrajecten te voorzien, zodat de betrokkenen de nodige competenties (zie de zes kritische dimensies van LA in figuur 2) kunnen ontwikkelen. Daarnaast moeten de data-analyse (IV.B) en de visualisatie (IV.C) goed zijn afgestemd op het eigenlijke gebruik en het doelpubliek. Hoewel LA kan leiden tot het meer doelgericht inzetten van middelen, kan het ook de werklast van de studieloopbaanbegeleiding verhogen.

---

*Aanbeveling 6: Het gebruik van LA kan enkel succesvol zijn als de gebruikers over de nodige competenties beschikken. Stem daarom de LA-toepassingen af op het doelpubliek en voorzie in informatie, opleidingen en richtlijnen die zijn afgestemd op de gebruikersgroep. Enkel zo zal LA op een correcte en gewenste manier ingezet en gebruikt worden.*

---

# III. Data

Centraal in de definities van LA staat de term 'gegevens' of 'data'. Elk LA-initiatief start met de gegevens van de gebruikers in een leercontext of -omgeving. De gebruikers kunnen zowel docenten en studenten zijn, maar bijvoorbeeld ook studieloopbaanbegeleiders. Dit onderdeel bespreekt voornamelijk de data die betrekking hebben op de studenten.

Het wijdverspreide gebruik van internet, computers, mobiele apparaten en softwaresystemen zoals *learning management*-systemen (LMS) zorgt voor een data-explosie (Long & Siemens, 2011) in onderwijsomgevingen. Terwijl vroeger het bijwonen van een les, het lezen van een boek of een discussie in de gang bijna onmiddellijk ontastbaar werd, zijn er nu expliciete data beschikbaar die in leeromgevingen worden geproduceerd en bijgehouden. Studenten, docenten en studieloopbaanbegeleiders laten sporen na via de digitale systemen die ze gebruiken of waarmee ze interageren. Mobiele apparaten, studenten- of personeelskaarten met technologie als RFID (*Radio-Frequency Identification*) of NFC (*Near-Field Communication*) en allerlei andere sensoren zorgen ervoor dat die datasporen alleen maar talrijker en ook rijker zullen worden. De geproduceerde data maken het mogelijk om via slimme analysetechnieken inzicht te krijgen in hoe leerprocessen effectief verlopen en op welke manier ze verbeterd kunnen worden (II.B).

Data die als startpunt fungeren voor LA zijn erg divers. De analyse van hun diversiteit kan vanuit verschillende standpunten gebeuren. Het volgende onderdeel gebruikt een opdeling op basis van het soort gegevens, gerelateerd aan hun betekenis. Andere opsplitsingen zijn mogelijk op basis van het dataformaat (in welke verschijningsvorm bestaan de data?), de databron en de plaats en het beheer van de opslag.

## A. Soorten data

De inhoudelijke betekenis van LA-data kan zeer verscheiden zijn (The University of Edinburgh Information Services, 2017). Er zijn minstens vier grote categorieën identificeerbaar op basis van het type.

Ten eerste zijn er **socio-demografische data** – leeftijd, geslacht, burgerlijke staat, gezinssamenstelling... – en de nauw hiermee verwante socio-economische data, zoals inkomen, en sociale klasse. Deze data komen dikwijls uit het zogenaamde *Student Information System* (SIS).

Ten tweede zijn er ook **academische data**. Hier gaat het om studieprestatiedata, zoals informatie over studiekeuzes, resultaten en behaalde diploma's, en ook specifieke relevante documenten die tot stand zijn gekomen tijdens het algemene

leerproces, zoals essays, blogs en andere opdrachten. De prestatiescores bij de evaluatie of de toetsing van studenten zijn veelgebruikte systematische metingen. Toetsen zijn erg informatief voor alle betrokkenen van educatieve omgevingen. Ze geven namelijk niet enkel informatie over het niveau en/of de vorderingen van de betrokken studenten, maar ook over de effectiviteit van het leermateriaal en de aanbiedingswijze ervan. Hierdoor leveren ze relevante input voor docenten, studieloopbaanbegeleiders, aanbieders van educatief materiaal en beleidsmakers, die elk op hun niveau het onderwijsleerproces proberen te optimaliseren. Deze data komen dikwijls uit een combinatie van het *Student Information System* (SIS) en het *Learning Management System* (LMS), ook virtueel leerplatform (VLE, *Virtual Learning Environment*) genoemd. Deze data kunnen ook zelf-gerapporteerd zijn als ze verkregen werden via onder andere vragenlijsten, portfolio's of interviews.

Ten derde is er de belangrijke bron van de **data over leeractiviteiten***.* Een eerste mogelijkheid is dat de aanwezigheid van studenten in leeractiviteiten of leercentra en bibliotheken wordt gebruikt. Dit kan niet enkel op basis van traditionele aanwezigheidslijsten, maar ook van de deelname aan digitale activiteiten tijdens contactmomenten (bijvoorbeeld het gebruik van *clickers* en de bezochte webpagina's, informatie van wificonnecties en gps-data). Vanwege het toenemende gebruik van digitale leersystemen produceren meer en meer onderwijsomgevingen data die weerspiegelen welke leeractiviteiten werden uitgevoerd. Deze elektronische leeromgevingen zijn bijzonder geschikt om diverse soorten data over het leerproces veelal automatisch te verzamelen. Volledig uitgewerkte elektronische leeromgevingen (zoals MOOCs) omvatten doorgaans ook (korte) toetsen om de student feedback te geven over de voortgang of om een gepast leerpad uit te tekenen. Daarnaast kunnen ook heel wat andere data geregistreerd worden, zoals welke oefeningen studenten maken, welke antwoorden ze geven en welke hulpmiddelen ze hierbij gebruiken, met inbegrip van de tijdstippen waarop de acties plaatsvinden. Deze logging data bevatten potentieel interessante informatie over de gebruiker (onder meer over evoluerende interesses en vaardigheden) en de omgeving (onder meer over de aantrekkelijkheid en de moeilijkheid van bepaalde inhouden). Automatisch loggen biedt de mogelijkheid om data te verzamelen zonder dat hierbij het gebruik van de omgeving of het leerproces wordt verstoord, wat mogelijk wel het geval is als er gebruik wordt gemaakt van expliciete tussentijdse toetsen of vragen. Bovendien biedt dit bijkomende perspectieven voor het permanent meten van evoluerende variabelen, terwijl systematische metingen eerder momentopnamen zijn. Naast logging data uit het LMS, MOOCs en *Small and Private Online Courses* (SPOCs) maken ook meer en meer docenten gebruik van e-books, videoplatformen, websites, apps, *serious games* enzovoort. *Wearables* bieden ook een nieuwe bron van data over leergedragingen: studenten kunnen uitgerust worden met sensoren die onder andere hun bewegingen, hartslag of hersenactiviteit meten. Heel wat data over leeractiviteiten hebben een lage granulariteit, waarbij er

doorgaans heel wat verwerking en aggregatie moet gebeuren alvorens de data ontsloten kunnen worden en tot concrete inzichten kunnen leiden. Behalve door ze automatisch te verzamelen kunnen data ook beschikbaar komen uit interviews (zelf-gerapporteerd) en observaties.

Ten vierde zijn er **contextuele onderwijskundige data**, over onder andere het curriculum, vakontwerp, lestijden, planning en werkvormen. Ze zijn veelal beschikbaar in het SIS, LMS en andere universitaire databanken, maar kunnen ook verzameld worden via vragenlijsten, interviews en observaties.

### B. Small of big data

LA wordt soms onterecht vereenzelvigd met *big data*. Om van *big data* te kunnen spreken moet het gaan om een groot volume, een diversiteit aan gegevenstypes en -bronnen en een hoge snelheid bij de datageneratie. LA kan ook al toepassing vinden bij *small* en zelfs traditionele data. Een voorbeeld is een studentendashboard op basis van academische data, dat studenten afhankelijk van hun scores op opleidingsonderdelen informatie geeft over de verwachte studievoortgang. Nieuwe databronnen, die dikwijls leiden tot *big data*, openen uiteraard wel een nieuw scala aan mogelijkheden voor een meer continue opvolging van de betrokkenen.

### C. Belang van context

De context heeft een grote impact op de LA-data. Zo zullen er afhankelijk van de context andere data beschikbaar zijn. Onlinecursussen bieden bijvoorbeeld veel meer datasporen dan meer traditionele lessenreeksen.

Op het microniveau moet daarbij bovendien het cursusdesign in rekening genomen worden bij het verbinden van conclusies aan het gedrag dat is gemeten op basis van de leeractiviteit binnen een bepaald opleidingsonderdeel. Zo zal contextuele informatie over het cursusdesign, zoals het lessenschema in een *flipped learning* opzet, het organiseren van evaluaties op bepaalde tijdstippen in een *mastery learning* opzet of het online plaatsen van feedback over opdrachten in een taakgebaseerd vakontwerp, het onlineleergedrag sterk bepalen.

Onderzoek legde onder andere de beperkingen bloot van predictieve LA die algoritmes toepasten op grote gedecontextualiseerde datasets van klikgedrag in LMS en examenresultaten (Rogers, Dawson, & Gaevi, 2014). Voorspellende indicatoren voor leersucces in bepaalde studies bleken geen voorspellende waarde te hebben voor andere studies. Belangrijke verschillen bestonden er ook tussen instellingen, disciplines en leerontwerp. Bovendien lieten bepaalde statische voorspellers van 'risicogedrag', zoals eerdere testresultaten van studenten, vaak ook geen ruimte voor pedagogische interventies om kansen op leersucces te verhogen. De pedagogische context van het gebruik van onlineleeromgevingen

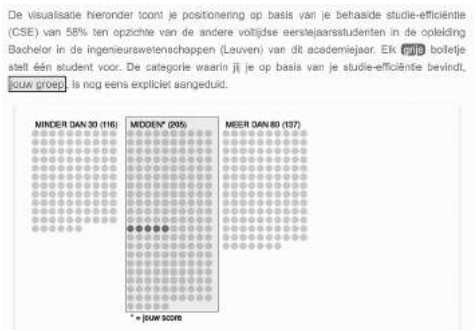
**Learning dashboard voor small data: feedback op academische resultaten (KU Leuven, België)**

REX is een learning dashboard dat studenten feedback geeft over hun academische resultaten na een examenperiode. Het tracht zelfreflectie uit te lokken over de impact van de resultaten op het studietraject (Tom Broos, Peeters, et al., 2017). Het dashboard biedt de volgende mogelijkheden:

- een zelfreflectie-instrument waarbij de studenten hun tevredenheid over de behaalde score op elk opleidingsonderdeel rapporteren. Met de mogelijkheid om per opleidingsonderdeel te zien waar de behaalde score zich positioneert ten opzichte van die van andere studenten;




- een globale positionering op basis van de studie-efficiëntie (CSE) ten opzichte van andere studenten in de opleiding;




- de mogelijke impact op het studietraject aan de hand van het aantal jaren dat vroegere studentenen nodig hadden om hun bachelor af te ronden, afhankelijk van hun studie-efficiëntie;




- concrete tips en aanbevelingen voor de verdere studie en
- een samenvatting van de relevante regelgeving rond studievoortgang en de mogelijke impact op het studietraject.

moet daarom steeds in rekening gebracht worden, zowel voor een correcte interpretatie van het gedrag in onlineleeromgevingen als voor het kunnen organiseren van aanbevelingen of interventies met het oog op het verhogen van studiesucces.

Op het macroniveau hebben ook de regionale en nationale regelgeving impact op de beschikbare data. Zo vraagt de overheid typisch studievoortgangsgegevens op van de universiteiten. Dit zorgt ervoor dat de universiteiten systemen opzetten waardoor deze gegevens beschikbaar zijn.

**De impact van nationale regelgeving op data: swipe cards in Nottingham Trent University (UK)**

Engelse universiteiten moeten bewijzen dat hun studenten een belangrijk deel van de tijd aanwezig zijn op de campus, om uit te sluiten dat ze illegaal aan het werk zijn. Om hierop een antwoord te geven heeft Nottingham Trent University een systeem opgezet waarbij ze de aanwezigheid van studenten nauwkeuriger opvolgt met onder andere een systeem waarbij studenten hun studentenkaart gebruiken om in gebouwen binnen te gaan. Deze swipe cards zijn intussen een belangrijke databron voor het bepalen van het *student engagement* in hun LA dashboards.

Ook de geschiedenis kan een grote impact hebben. In landen als Duitsland en Oostenrijk is er een grote bezorgdheid rond ethiek en privacy van LA, die mede is ontstaan door het misbruik van persoonlijke gegevens tijdens de Tweede Wereldoorlog. Deze bezorgdheid heeft zich niet enkel vertaald in een grote waakzaamheid bij de belanghebbenden, maar ook in regelgeving, zowel bij de overheid als bij de instellingen zelf, onder andere maar niet uitsluitend omtrent het gebruik van persoonsgegevens.

# IV. Hoe?

Hoewel hierboven blijkt dat data tegenwoordig overal zijn en hoe langer hoe meer beschikbaar komen, is het uiteraard essentieel om ze eerst en vooral te verzamelen, op te slaan en voor te verwerken. Pas daarna kunnen de data-analyse (IV.B) en -visualisatie (IV.C) starten die aan de grond liggen van de toepassing van LA (II). Daarbij moet erover gewaakt worden dat bij de profilering en geautomatiseerde besluitvorming de betrokkenen ook nuttige informatie over de logica en de verwachte gevolgen kunnen opvragen en een beroep kunnen doen op het recht op inzage (V).

## A. Dataverzameling, -opslag en -voorverwerking

### 1. Dataverzameling

Zoals blijkt uit III is er een veelheid aan data beschikbaar. Alvorens te beslissen welke data te gebruiken moet de doelstelling duidelijk zijn. Het doel (zie II) is immers bepalend voor het antwoord op de vraag welke data noodzakelijk of nuttig zijn. Daarnaast moet het doel ook volgens het gegevensbeschermingsrecht vooraf worden bepaald en nadien worden nageleefd (V).

Binnen LA zijn data die afkomstig zijn uit het LMS, *Student Information System* (SIS) en onlinecursussen zoals MOOCs en SPOCs het gangbaarst, mede omdat ze meestal al beschikbaar zijn of met een beperkte inspanning verzameld kunnen worden. Starten met de beschikbare data is een te verantwoorden stap, zolang het bijdraagt tot het bereiken van het doel dat gesteld is met LA. Een bijkomend voordeel is dat dit al een eerste en mogelijk noodzakelijke ervaring kan aanreiken voor het gebruik van LA. Beter is echter om al vanaf het begin van een LA-project vast te leggen welke soorten data nodig zijn om mogelijke nuttige inzichten rond het leerproces in een bepaalde onderwijscontext te verkrijgen en om data gericht en actief te verzamelen (zie ook verder over verplichtingen voor instellingen rond het gegevensbeschermingsrecht in V.A). Al van bij het bepalen van de datakeuzes is het cruciaal daar onderwijskundige expertise en de betrokken docenten bij te betrekken voor een correcte vraagstelling en interpretatie van de data-analyse. Dit is ook cruciaal voor het verkrijgen van *actionable feedback* die het leerproces effectief kan ondersteunen en versterken.

---

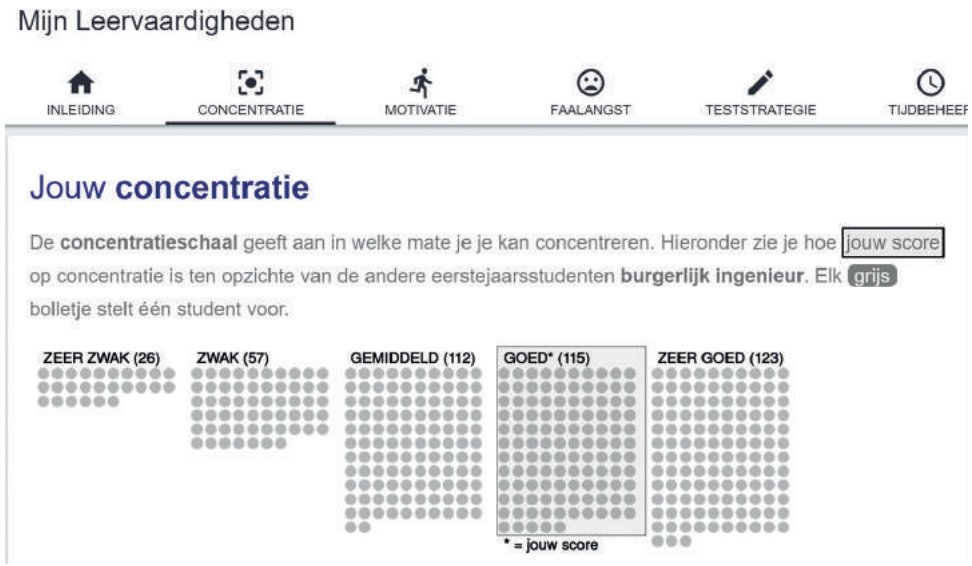
*Aanvulling aanbeveling 5: Betrek de verschillende belanghebbenden vanaf de eerste stap van de ontwikkeling bij Learning Analytics. De docenten van het vak, onderwijskundigen en studieloopbaanbegeleiders zijn onmisbaar voor een correctie interpretatie en het bereiken van de doelstellingen van de data-analyse. Raadpleeg ook de studenten voor het bepalen/evalueren van interventies die hun leerproces kunnen ondersteunen.*

---

Hoewel de verzameling van data uit andere bronnen, zoals *wearables*, camera's of gps, een grotere uitdaging kan zijn, kan de doelstelling van het LA-project zo'n inspanning vereisen. Daarnaast kunnen ook minder voor de hand liggende data, zoals de resultaten van bevragingen of onderzoeksenquêtes, een nuttige en eenvoudig beschikbare databron zijn.

**Data van een gevalideerde vragenlijst voor feedback op leer- en studeervaardigheden (KU Leuven, België)**

Het LASSI-dashboard gebruikt de gegevens uit een wetenschappelijk gevalideerde vragenlijst om eerstejaarsstudenten al in de eerste weken van het academiejaar feedback te geven op hun concentratie, motivatie, faalangst, gebruik van teststrategieën en tijdbeheer (Broos et al., 2017; Broos, Verbert, Vansoom, Langie, & De Laet, 2017; De Laet et al., 2017). Studenten krijgen niet enkel de eigen score te zien, maar ook hun positionering ten opzichte van medestudenten. Daarnaast geeft het dashboard ook weer hoe de leer- en studeervaardigheden het studiesucces van studenten uit eerdere cohortes hebben beïnvloed, geeft het concrete tips om de vaardigheden te verbeteren en verwijst het naar bijkomende ondersteuning binnen de universiteit.




Samen met de groei van de hoeveelheid data die bijgehouden wordt, wordt het ook des te belangrijker na te gaan wie de data heeft en wie zeggenschap heeft over de data (zie ook V.A).

---

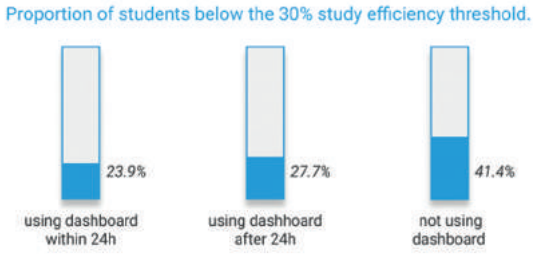
*Aanbeveling 7: De hogeronderwijsinstelling heeft de verantwoordelijkheid erover te waken dat ze, zelfs bij gebruik van externe diensten, toch de controle op en het eigenaarschap van de data behoudt, en dat deze ook beschikbaar blijven voor verder intern gebruik.*

---

Ten slotte kan de uitrol van *learning dashboards*, als resultaat van een LA-toepassing, net ook bijkomende datasporen opleveren. Of en hoe de belanghebbenden interageren met de dashboards levert immers opnieuw informatie op, die bijvoorbeeld nuttig is voor de meting van hun inspanningen of het zelfregulatiegedrag. Een en ander vereist evenwel dat dit steeds transparant gebeurt en er een wettelijke basis is voor de data-inzameling (zie ook V).

**Gebruik van het LASSI-dashboard als bijkomend dataspoor (KU Leuven, België)**

Het gebruik van het LASSI-dashboard (zie vorig kader) dat studenten van 26 opleidingen van de KU Leuven en drie opleidingen van de TU Delft in de eerste week ontvangen, creëert al een bijkomend dataspoor. Het aandeel van studenten die op het einde van het academiejaar een erg lage studie-efficiëntie halen is immers significant hoger bij de studenten die het dashboard niet gebruiken dan bij de studenten die het wel gebruiken. Het gebruik van het dashboard in de eerste week van het academiejaar heeft dus een bijkomend dataspoor gecreëerd, dat bruikbaar is voor de identificatie van studenten die het risico lopen op het behalen van slechte studieresultaten.




*Figuur 3: Aandeel van studenten die later een lager studie-efficiëntie behaalden voor drie gebruikersgroepen van het LASSI-dashboard: studenten die binnen de 24 uur het dashboard gebruiken, degenen die het na 24 uur doen en degenen die het niet gebruiken.*

## 2. Dataopslag

Om de interoperabiliteit van data uit leeromgevingen te bewerkstelligen, zowel van *small* als *big data*, bestaan e-learning-standaarden die het mogelijk maken verschillende leerplatformen te laten samenwerken. Het gebruik van datastandaarden biedt heel wat voordelen voor de verzameling en uitwisseling van data (del Blanco, Serrano, Freire, Martinez-Ortiz, & Fernandez-Manjon, 2013). De *Learning Tools Interoperability*-standaard (LTI, van IMS Global Learning Consortium), de AICC learning-standaard (*Aviation Industry Computer-Based Training Committee*) of Scorm (*Sharable Content Object Reference Model*, *Advanced Distributed Learning* (ADL), ondersteund door het Amerikaanse ministerie van Defensie, maken het onder andere mogelijk om externe leerapplicaties te integreren in een LMS of andere leeromgeving. Leermateriaal dat ontwikkeld is in een bepaalde auteursapplicatie kan zo in verschillende LMS'en geïmporteerd of geïntegreerd worden, en aangeboden aan gebruikers.

Twee nieuwe standaarden voor het registreren van ervaringen binnen leeromgevingen in het kader van LA winnen op dit moment wereldwijd aan belang. Wanneer een student handelingen uitvoert in een leeromgeving, worden die aan de hand van een gestandaardiseerd dataformaat gelogd als events met een tijdsaanduiding.

Enerzijds is er de *IMS Caliper*-specificatie[4] (ontwikkeld door het IMS Global Learning Consortium) en anderzijds de *Experience API* (of xAPI) van ADL. *IMS Caliper*, dat onder meer wordt gebruikt door *Blackboard LMS*, wordt ontwikkeld binnen een gesloten consortium. xAPI (ontstaan als Tincan API) is een open standaard die is ontwikkeld door de opensourcegemeenschap. Open standaarden kunnen als voordeel hebben dat de data-inzameling en -transfer door de gemeenschap gemakkelijker kunnen geverifieerd of gecontroleerd worden, meer dan bij gesloten standaarden. De standaarden definiëren regels voor het loggen van ervaringen van studenten en van activiteiten (zowel online- als fysieke prestaties). De twee standaarden hebben verschillende oorsprongen en regels. Hoewel ze zich parallel ontwikkelen in concurrentiële toepassingen, zijn er toenaderingen en initiatieven om vergelijking, complementariteit en samenwerking te onderzoeken.

**Voorbeelden van activiteitenstatement op basis van de xAPI-standaard**

Het voorbeeld hieronder toont drie gelogde leeractiviteiten. De student uit rij 1 heeft een bepaalde interactieve oefening over het persoonlijk voornaamwoord afgerond met een resultaat van 80%. De student uit rij 2 heeft een reactie gepost op het discussieforum met titel b, als reactie op een specifiek onderwerp en op een specifieke andere student. De student uit rij 3 heeft de taak van week 2 van een medestudent beoordeeld.

| <Actor> | <Verb> | <Object> | (<Result>) | (<Context>) |
|---|---|---|---|---|
| learner 63151 | completed | Interaction on "EC0102 Le pronom personnel" | with result 80 % | |
| learner x | posted | discussion b | | in discussion thread i, in reply to learner z |
| learner y | assessed | learner z | | on week 2 assignment about … |

[4] Er is geen unanimiteit over het label 'specificatie' of 'standaard'. Officiële standaarden zijn erkend door organisaties als de International Standards Organisation. DISC is een initiatief dat de opdracht gekregen heeft om xAPI-interoperabiliteit en -conformiteit te ondersteunen (http://datainteroperability.org/)

Alle formele en informele leertoepassingen die compatibel zijn met de standaarden (in opmars, maar niet altijd even conform), of ze nu binnen een LMS of los daarvan gebruikt worden, kunnen informatie over de leerervaringen synchroon or asynchroon doorsturen naar een centraal opslagsysteem, een *Learning Record Store* (LRS).

Het flexibele gebruik van werkwoorden en leerobjecten in de standaarden maakt het in principe mogelijk om een breed gamma aan leerervaringen uit een even breed gamma aan leeromgevingen te loggen in een natuurlijke taal (statements zijn zowel leesbaar voor machines als mensen dankzij het het json-dataformaat). Het gedeelde formaat maakt het dan ook mogelijk om leerervaringen in verschillende toepassingen te vergelijken. Andere voordelen van het gebruik van standaarden zijn een meer efficiënte dataverzameling (minder werk aan het uitzuiveren van data), flexibiliteit in de toegang tot de eigen data (deze kunnen bewaard worden in een LRS die in eigen beheer gehouden wordt en niet in het beheer van de LMS-aanbieder), de mogelijkheid om de toegang tot LA-infrastructuur en de keuze van LA niet te laten afhangen van een specifieke aanbieder en ten slotte de mogelijkheden voor longitudinaal onderzoek. Het academisch onderzoek en de instellingen van hoger onderwijs hebben daarom alle belang bij interoperabiliteit tussen verschillende leeroplossingen en gestandaardiseerde data die gedeeld en vergeleken kunnen worden.

Op dit moment zijn de LA-standaarden weliswaar *'work in progress'*. Samenwerkingsinitiatieven om de conformiteit te controleren en duidelijke richtlijnen te formuleren voor de conforme implementatie van datatracking zijn nodig.

---

*Aanbeveling 8: Vlaamse instellingen van hoger onderwijs zetten in op een gedragen strategie rond het gebruik en de implementatie van internationale standaarden voor LA-data. Samenwerking tussen instellingen versterkt de positie van elke instelling ten opzichte van commerciële aanbieders bij de onderhandeling over het gebruik van open standaarden.*

---

Een *Learning Record Store* is, zoals hierboven al aan bod kwam, een centraal *student record warehouse* voor het opslaan van leerdata uit verschillende bronnen. Van hieruit kunnen data ook gedeeld worden met bestaande externe rapporteringstools (of andere LRS) (Del Blanco et al., 2013).

---

*Aanbeveling 9: LA kan zowel op small als op big data gebaseerd zijn. De Vlaamse instellingen van hoger onderwijs zijn technologisch alvast voorbereid voor small data. De opslag en verwerking van big data vragen echter bijkomende expertise en infrastructuur, en de bijbehorende investeringen. De data moeten daarbij beschikbaar zijn, zodat de betrokkenen ook vragen kunnen beantwoorden of zelfs nieuwe vragen formuleren, die dan vervolgens tot inzichten kunnen leiden. Hierbij is niet enkel de toegang belangrijk maar ook de toegankelijkheid.*

---

De opgeslagen informatie kan dus ook voor analysedoeleinden worden opgehaald uit de LRS.

### 3. Datavoorverwerking

De verschillende soorten verkregen data zijn 'ruw' en waarschijnlijk imperfect. Ze kunnen inconsistenties en redundanties bevatten. Daarnaast kunnen de data mogelijk heel uitgebreid zijn. Daarom zijn bijkomende voorverwerkingsstappen noodzakelijk vooraleer de data bruikbaar zijn voor LA. De datavoorverwerking is verre van te verwaarlozen en is zelfs een van de grootste struikelblokken bij de huidige implementatie van LA wereldwijd (Leitner et al., 2017). Datavoorverwerking bestaat typisch uit *data-cleaning*, datanormalisatie, datatransformatie, *missing value imputation* (bij ontbrekende gegevens), data-integratie en het identificeren van storingsbronnen. Bijkomende grote uitdagingen voor LA zijn de koppeling van gegevens uit verschillende bronnen en de koppeling van de gegevens aan concrete betekenisvolle elementen of acties (bv. welke oefening in de leeromgeving). Het gebruik van unieke identificatiecodes kan de koppeling vereenvoudigen, maar zal niet altijd mogelijk zijn. Heel wat bedrijven bieden technische oplossingen aan voor de datakoppeling, maar ook de wettelijke (V.A) en ethische (V.B) aspecten mogen hierbij niet uit het oog verloren worden. Daarnaast zal een instelling mogelijk niet enkel individueel de eigen beschikbare data willen koppelen, maar ook een instellingsoverstijgende koppeling willen realiseren.

## B. Data-analyse

### 1. Statistische en psychometrische technieken

In dit onderdeel gaan we in detail in op enkele analysetechnieken uit de statistiek en psychometrie. Eenvoudigere technieken (zoals het gebruik van beschrijvende maten en grafieken, of standaardtechnieken, zoals lineaire regressie en variantie-analyse) komen niet aan bod, zonder hiermee afbreuk te willen doen aan hun waarde.

Technieken uit de psychometrie zijn inzetbaar bij de analyse van **toetsen of andere systematische metingen**. Eén mogelijkheid is de itemresponstheorie (*item response theory* of IRT; Embretson & Reise, 2000), een verzameling statistische modellen die de kans op een bepaalde respons op het item (oefening, taak, vraag) modelleren als functie van kenmerken van de persoon en het betrokken item. Bij toepassing op toetsdata geeft het eenvoudigste IRT-model, het Rasch-model, de kans op een juist antwoord op een item als functie van onderliggende (latente) persoons- en itemkenmerken (doorgaans geïnterpreteerd als respectievelijk de vaardigheid van de student en de moeilijkheidsgraad van het item).

Dit model heeft verschillende nuttige toepassingen. Een eerste is het **kalibreren van items**, meer bepaald om hun moeilijkheidsgraad in kaart te brengen.

Dit leidt tot een soort van meetschaal van bijvoorbeeld 'vaardigheid' waarop items gepositioneerd zijn, te vergelijken met een meetlat. Door vervolgens de betrokkene, meestal een student, een aantal gekalibreerde items aan te bieden kan de plaats van de student op de schaal ingeschat worden, net zoals de lengte van een persoon op een meetlat aangeduid zou kunnen worden. Indien een student hoger op de schaal staat dan een item, betekent dit dat de student het item beheerst. Een interessant kenmerk van een dergelijke meting is dat de verwachte waarde voor de schatting van de vaardigheid van een student niet afhangt van het aangeboden item. Dit leidt meteen tot een tweede toepassing: IRT is inzetbaar om de voortgang van studenten te volgen, zonder zich zorgen te hoeven maken over de vergelijkbaarheid van de moeilijkheidsgraad van de items op de verschillende meetmomenten. Meer zelfs: naargelang van de gemeten vaardigheid is het mogelijk om items met een meer gepaste moeilijkheidsgraad aan te bieden. Dit leidt tot **adaptief testen**. IRT wordt dan ook gebruikt bij *computerized adaptive testing* (CAT), waarbij na elk antwoord van een student de inschatting van de vaardigheid van de student wordt geüpdatet en deze inschatting wordt gebruikt voor de selectie van het volgende item. Dat is het item waarvan de verwachte informatiewaarde over de vaardigheid het grootst is. Op die manier kunnen kortere tests met adaptieve items even nauwkeurig de vaardigheid van studenten inschatten dan langere tests met vaste items.

Digitale leeromgevingen leveren meer op dan data over de correctheid van de antwoorden op taken. Typisch zijn er heel wat meer **logging data** beschikbaar. Een eerste uitdaging is inzicht krijgen in de samenhang tussen de informatiebronnen en ze op een verstandige manier combineren om zo nauwkeuriger en meer valide uitspraken en voorspellingen te kunnen doen. Methoden die het aantal dimensies reduceren (zoals factoranalyses) of regressiemethoden voor situaties met veel predictoren (zoals variabele *selection methods* en *penalty methods*, zoals *ridge* en *Lasso-regressie*) kunnen hier hun nut bewijzen.

Daarnaast vormen logging data ook een uitdaging voor IRT-modellen. IRT-modellen voor toetsdata nemen immers aan dat de vaardigheid tijdens de toets constant blijft, terwijl de vaardigheid in een leeromgeving waarin studenten bij het maken van oefeningen onmiddellijk feedback krijgen typisch evolueert over de tijd. Een toepassing van IRT in leeromgevingen vraagt dan ook om dynamische uitbreidingen van de modellen (Kadengye, Ceulemans, & Van den Noortgate, 2015). Een andere toepassing van IRT-analyses in leeromgevingen waarin studenten zelf een keuze kunnen maken in de oefeningen is zicht krijgen op welke oefeningen studenten kiezen.

IRT-modellen zijn specifieke gevallen van *generalized linear mixed models* (GLMMs) (De Boeck & Wilson, 2004). De flexibiliteit van het GLMM-raamwerk laat toe modellen te *fitten* die rekening houden met de specificiteit van het type data en de interesse van de gebruiker. Ze kunnen bijvoorbeeld rekening houden

met meerdere vaardigheden tegelijk (gebruik makend van multidimensionale modellen), met complexere antwoordtypes, zoals bij polytome items, met antwoordtijden enzovoort. Door predictoren toe te voegen kan het effect van bepaalde itemkenmerken op de moeilijkheidsgraad van de items of het effect van persoonskenmerken op de vaardigheid van de studenten nagegaan worden. IRT is ook bruikbaar voor het meten van andere latente persoonsvariabelen dan vaardigheid, zoals interesse of motivatie.

Andere modellen die voor de analyse van LA relevant zijn, zijn multiniveau-modellen (*multilevel models* of *mixed models* (Raudenbush & Bryk, 2002)). Multiniveaumodellen zijn bruikbaar bij de analyse van geclusterde data. Aan de hand van deze modellen kan bijvoorbeeld bij grootschalige onderwijsassessments nagegaan worden in welke mate er verschillen zijn tussen scholen, tussen klassen binnen scholen en tussen leerlingen binnen klassen. Bovendien bieden ze de mogelijkheid om systematisch te zoeken naar verklaringen voor eventuele verschillen, door school-, klas- en studentkenmerken als parameters op te nemen. Deze verklaringen kunnen niet enkel theoretisch interessant zijn, ze kunnen ook handvatten aanreiken voor een optimalisatie van de onderwijspraktijk. De flexibiliteit van deze modellen en de mogelijkheid om rekening te houden met allerlei soorten van afhankelijkheden maken deze modellen nuttig voor een hele waaier aan onderwijskundig onderzoek, inclusief interventieonderzoek. Andere statistische modellen die erg relevant zijn in onderwijskundig onderzoek zijn (multiniveau) structurele vergelijkingsmodellen (*structural equation models* of SEMs (Kaplan, 2009)). Terwijl traditionele analyses dikwijls werken met geobserveerde variabelen en daarbij een onderscheid maken tussen afhankelijke en onafhankelijke variabelen, geeft SEMS aan hoe de geobserveerde variabelen indicatoren zijn van latente variabelen en kunnen ook complexere verbanden tussen latente variabelen worden gemodelleerd, waarbij bijvoorbeeld een variabele tegelijk de rol van afhankelijke en onafhankelijke variabele kan opnemen.

Met psychometrische modellen kun je niet enkel uitspraken doen over gebruikers (bijvoorbeeld: zijn er kenmerken van gebruikers die voorspellend zijn voor hun interesse, vaardigheid of leercurve?) of de omgeving (bijvoorbeeld: zijn er kenmerken van oefeningen die voorspellend zijn voor de populariteit ervan of voor de mate waarin er leerresultaten worden geboekt?), maar ook over interactie-effecten (bijvoorbeeld: welk soort oefeningen leidt bij welke type gebruikers tot de beste resultaten?). Ook kan het model rekening houden met de kenmerken van de context waarin wordt gewerkt (bijvoorbeeld: welk soort oefeningen is onder deze conditie voor dit soort gebruiker het meest geschikt?). Deze analyses zijn niet alleen interessant vanuit theoretisch en onderzoeksoogpunt (II.A), maar ze kunnen ook leiden tot globale aanpassingen van de leeromgeving (II.B). Zo kan een leeromgeving gebouwd worden die de motivatie en het leren van de gebruiker zo hoog mogelijk houdt, eventueel rekening houdend met kenmerken van die gebruiker (Wauters, Desmet, & Van den Noortgate, 2010). Door voort te gaan

op de acties die studenten ondernemen en de antwoorden die zij geven ontstaat er geleidelijk aan een scherper zicht op de mogelijkheden, noden en wensen van individuele studenten. Een uitdaging is om via data-analyse het beeld permanent bij te sturen en op basis hiervan de leeromgeving aan te passen. Een mogelijkheid is de leeromgeving voldoende uitdagend en leerrijk te houden, door oefeningen aan te bieden met een aangepaste moeilijkheidsgraad, bijvoorbeeld waarvan de verwachting is dat de student een kans heeft van 0.70 om tot een correct antwoord te komen. Het spreekt voor zich dat de aanpassing van de omgeving haast onmiddellijk dient te gebeuren, zodat de gebruiker hier geen hinder van ondervindt.

**Adaptief testen met het *Elo rating system* (ERS, Nederland)**

Het *Elo rating system* (ERS; Klinkenberg, Straatemeier, & van der Maas, 2011) is een voorbeeld van een eenvoudig en dus snel algoritme dat op basis van een voorlopige inschatting van de vaardigheid van de student en de moeilijkheidsgraad van het item voorspelt of de respons correct zal zijn, en dat afhankelijk van de werkelijke respons deze inschattingen naar boven of naar onder bijstelt. Een uitdaging bij deze werkwijze is het zoeken naar goede startwaarden, zodat het algoritme snel convergeert naar de echte vaardigheid (het *cold start*-probleem). Een andere bekommernis is dat dit algoritme geen rekening houdt met de informatie die het voorgaande traject van de student of trajecten van andere studenten kunnen bieden. Voor beide beperkingen kan een combinatie van het Elo-algoritme met analyses van voorafgaande logging data zoals hierboven beschreven een oplossing bieden.

## 2. Machinaal leren

De eenvoudigste, maar ook een heldere, definitie van machinaal leren (*machine learning*) luidt: "Een machine leert met betrekking tot een specifieke taak T, prestatiemaatstaf P en type ervaring E, als het systeem de eigen prestatie P, ten aanzien van taak T, naar aanleiding van ervaring E, op betrouwbare wijze verhoogt." (Mitchell, 1997). Merk op dat dit Standpunt de term 'machinaal leren' hanteert en niet 'automatisch leren', om niet de verkeerde indruk te wekken dat de mens niet betrokken zou zijn in het proces.

Belangrijk is dat computersystemen die kunnen leren, in staat moeten zijn 'ervaringen' op te doen op basis van verkregen data, zodat ze de invloed van hun eigen gedrag kunnen doormeten en dat gedrag naar aanleiding daarvan kunnen bijstellen. Dat gaat volautomatisch op basis van zogenaamde algoritmes. Een algoritme is een eindige reeks instructies die vanuit een gegeven begintoestand naar een beoogd doel leidt. Vergelijk het met een stappenplan waarmee computers een taak kunnen verrichten. Bij machinaal leren bestaan die stappenplannen deels uit wiskundige functies die verbanden zoeken tussen datapunten of profielen van personen of groepen van personen of organisaties. Zo kan bijvoorbeeld blijken dat studenten die in de laatste week voor de examens een piekactiviteit op de onlineleeromgeving vertonen, veel vaker studievertraging oplopen.

Machinaal leren en statistiek voor LA delen hetzelfde doel: leren aan de hand van data over leren. De oorsprong van de twee is echter sterk verschillend. Statistiek is een discipline binnen de wiskunde, terwijl machinaal leren een onderdeel is van de computerwetenschappen en artificiële intelligentie. Een statistische aanpak binnen LA gaat in een eerste stap uit van het formuleren van hypotheses en stelt vervolgens statistische modellen op om deze hypotheses te bevestigen of ontkrachten. Machinaal leren en *knowledge discovery in databases* gaan typisch verder dan beschrijvende statistiek, in de zin dat ze ook op zoek gaan naar valide, ongekende, potentieel bruikbare en verklaarbare modellen/patronen. Dat laatste is ook gerelateerd aan *data mining*, waarin op een geautomatiseerde manier patronen en relaties worden gezocht in grote hoeveelheden gegevens.

Er zijn twee grote takken binnen het machinaal leren: gesuperviseerd leren en ongesuperviseerd leren. Bij **gesuperviseerd leren** krijgt het algoritme voorbeelden (trainingset) van invoer en de bijbehorende uitvoer. Het leert op basis van deze voorbeelden hoe de eigenschappen van de invoer bepalend zijn voor de uitvoer. Na de leerfase kan het algoritme ook voor nieuwe invoer zelfstandig de juiste uitvoer produceren (generaliseren). Een veel gebruikte probleemvorm is classificatie, waarbij de invoerelementen ingedeeld moeten worden in groepen. Een typisch voorbeeld uit LA is de classificatie van studenten in een risicogroep en een niet-risicogroep, of het voorspellen van het risico dat een student zal uitvallen. Voorbeelden van algoritmes zijn beslissingsbomen, artificiële neurale netwerken en *Support Vector Machines*. De jongste jaren zijn heel wat performante systemen ontwikkeld op basis van *deep learning*.

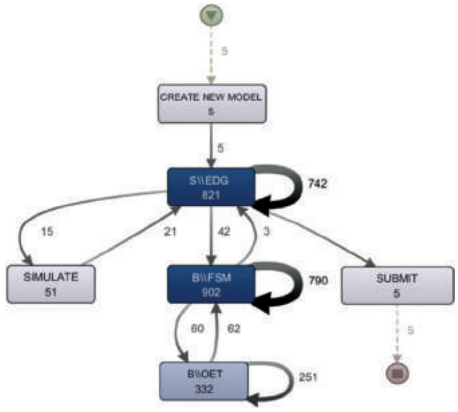
**Voorbeeld van gesuperviseerd leren: voorspellen van *drop-out* aan de hand van logging data in een MOOC**

Data-analyse aan de hand van gesuperviseerd leren kan uitval van studenten (*drop-out)* voorspellen op basis van de uitgebreide logging data die worden geproduceerd door onlineactiviteiten van studenten in een MOOC (Deeva, De Smedt, De Koninck, & De Weerdt, 2018). De grote uitdaging is dat de logging veel en redelijk ongestructureerde data oplevert. Uit de vergelijking van verschillende technieken bleken vooral de analytische modellen, en in het bijzonder *random forests* voor sequentieclassificatie, in staat om met deze data om te gaan en nauwkeurige voorspellingen te produceren.

Bij **ongesuperviseerd leren** zijn er geen voorbeelden van de gewenste uitvoer; het algoritme ontdekt zelf een structuur in de gegeven invoer. Dit kan bijvoorbeeld door de invoer te verdelen in groepen van elementen die op elkaar lijken, zoals hoogperformante studenten met faalangst. Voorbeelden van algoritmes zijn *k-nearest neighbour of Expectation-Maximization clustering* met *Gaussian Mixture Models*.

**Voorbeeld van ongesuperviseerd leren: process mining voor het analyseren en ontdekken van leerpatronen**

In de context van een vak rond domeinmodelleren werd ongesuperviseerd leren gebruikt om typische patronen te ontdekken die studenten (nieuw in het domein) gebruiken om te modelleren en vervolgens hun modellen te valideren (Sedrakyan, De Weerdt, & Snoeck, 2016). De patronen zijn gebaseerd op event logs tijdens het gebruik van de modelleringssoftware MERODE en de resultaten van studenten. Zo concludeerden de onderzoekers dat studenten die slechter presteren meer sequentieel werken (één model tegelijk), terwijl de best presterende studenten iteratief te werk gaan en meer frequent overstappen tussen verschillende modellen. De ontdekte patronen kunnen docenten inzicht geven in het leerproces van de studenten en hen helpen om de begeleiding van studenten te verbeteren. Dankzij de ontdekte patronen kan men studenten ook procesgeoriënteerde feedback geven, eerder dan enkel feedback op het resultaat.



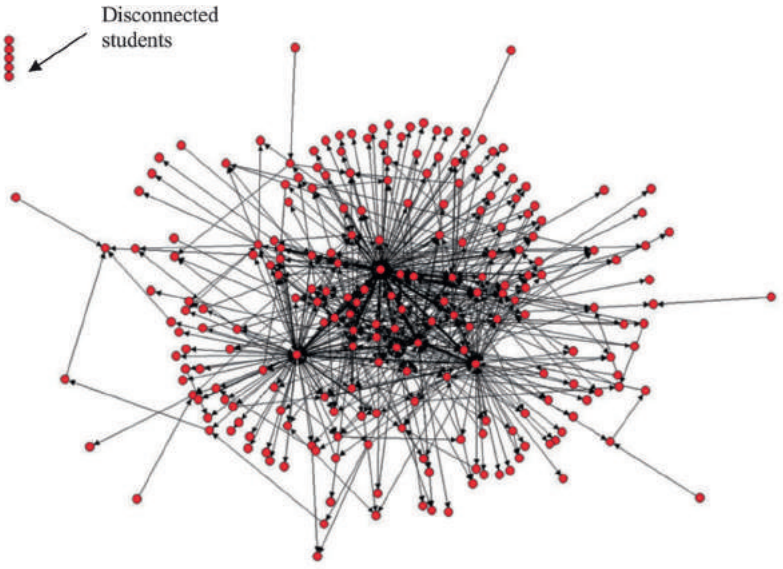

*Figuur 4: Procesmap voor modelleeractiviteiten van de vijf slechts presterende groepen* (Sedrakyan et al., 2016).

## C. Visualisatie

Informatievisualisatie is het gebruik van interactieve visuele representaties om cognitie te versterken (Card, Mackinlay, & Shneiderman, 1999). In *visual analytics* ligt de nadruk specifiek op de constructie van modellen en de visualisatie hiervan om de modellen beter te begrijpen of te verfijnen (Keim, Mansmann, Schneidewind, Thomas, & Ziegler, 2008). Een relevant doel van informatievisualisatie is het ontdekken van patronen door gebruik te maken van het perceptuele vermogen van mensen. De patronen kunnen vaak gemakkelijker ontdekt worden uit een visuele representatie (Ware, 2012). Technieken uit informatievisualisatie zijn daarnaast geschikt om complexe, multidimensionale data meer interpreteerbaar te maken, zoals data over het volgen van studenten in een onlineleeromgeving (Mazza & Milani, 2005).

**Voorbeeld van visualisatie: analyse van de online-interactie tussen studenten**

De onderstaande visualisatie toont de interacties tussen studenten (rode punten) in een onlinecursus op een gegeven tijdstip (Dawson, 2010). Door de interacties over de tijd te volgen is het niet enkel mogelijk om ze op te volgen maar ook om veranderingen in het studentengedrag over de tijd te ontdekken en zelfs om de impact van interventies na te gaan. Daarnaast kan men studenten die niet of weinig interageren gemakkelijk identificeren.




Visualisatie helpt niet enkel bij het beantwoorden van vooropgestelde vragen, maar ook bij het identificeren van nieuwe vragen. Daarnaast biedt het de mogelijkheid om problemen met de gegevens zelf te onthullen, bijvoorbeeld over de manier waarop de gegevens zijn verzameld, gekoppeld of voorverwerkt. Vooral bij het gebruik van (semi)geautomatiseerde *trackers* om data van studenten te verzamelen is een dergelijke kwaliteitscontrole nodig (S. Charleer, Klerkx, Duval, De Laet, & Verbert, 2017).

Statische visualisaties bieden typisch antwoorden op een beperkt aantal vragen die een gebruiker kan hebben over data (Klerkx, Verbert, & Duval, 2017). Infographics worden bijvoorbeeld vaak gebruikt voor het vertellen van verhalen in de journalistiek. Het kijken naar een visualisatie leidt echter vaak tot nieuwe vragen, die alleen kunnen worden beantwoord door interactie met de gegevens (Few, 2006). Het is daarom vaak nodig om dynamische interactietechnieken aan de visualisatie toe te voegen en om zinvolle visualisaties te ontwerpen die verkennende data-analyse stimuleren.

In het LA-domein worden interactieve visualisaties van relevante data voor eindgebruikers en betrokkenen vaak onderzocht onder de term *'learning*

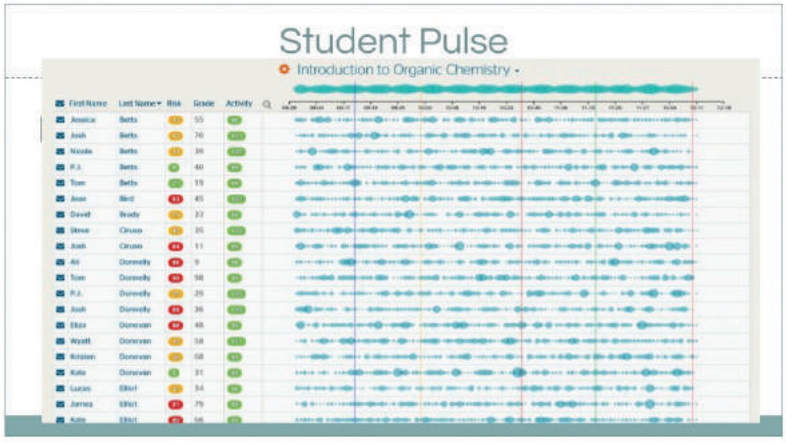
*dashboards'*. Dat zijn interactieve schermen die aan de betrokkenen inzicht geven in de data die werden verzameld en de data-analyses die werden uitgevoerd.

Een eerste groep dashboards is vooral gericht op het activeren van studenten tijdens hoorcolleges of groepswerk (Verbert et al., 2014). Een prominent voorbeeld van het eerste is *Classroom Salon* (Barr & Gunawardena, 2012), waarin studenten gezamenlijk slides annoteren. Dit instrument activeert studenten tijdens hoorcolleges en geeft docenten inzicht in potentiële moeilijkheden tijdens een college.

De meeste dashboards zijn ontwikkeld om studenten en docenten te ondersteunen tijdens *blended* of onlineleren (Verbert et al., 2014). Deze groep is gericht op het faciliteren van feedback naar studenten en wil docenten meer inzicht geven in de vooruitgang van studenten.

**Voorbeeld van een dashboard voor de ondersteuning van de docent bij de opvolging van studenten in een onlinecursus**

*Student Pulse* is een verzamelnaam voor visualisaties die zijn ontwikkeld voor de opvolging van studenten in een onlinecursus (Apereo foundation, 2017). De visualisaties tonen of en in welke mate studenten interageren met het onlineleermateriaal.




Een laatste categorie zijn dashboards voor specifieke toepassingen of contexten, bijvoorbeeld voor de interactie tussen een studieloopbaanbegeleider en een student (zie het voorbeeld in II.C.2).

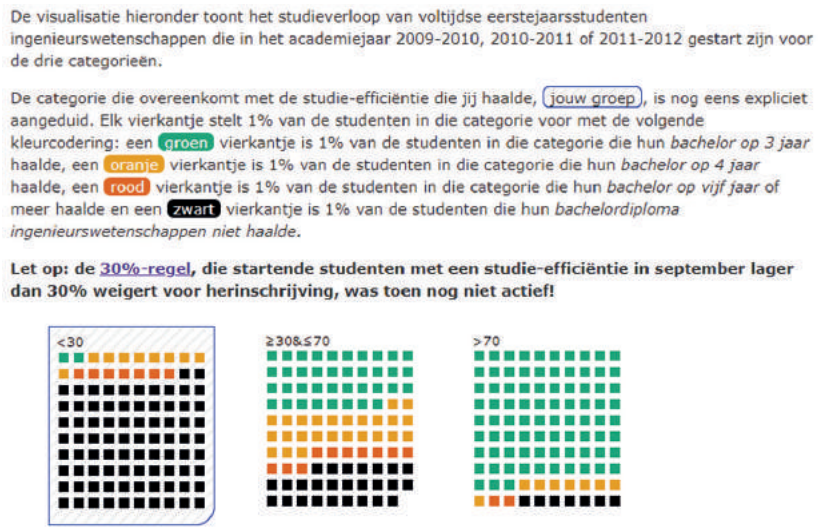
### D. Descriptief, predictief of prescriptief

Dit onderdeel gaat meer in detail in op drie mogelijke gebruikswijzen van LA: descriptief, predictief of prescriptief. Hoewel de scheidingen niet strikt zijn, biedt deze opdeling toch inzicht in verschillende gebruikswijzen van LA.

**Descriptieve LA** beschrijft of vat de beschikbare data samen, zodat ze interpreteerbaar en bruikbaar zijn voor de gebruikers. In essentie beschrijft descriptieve LA het verleden: dit kan gaan over een zeer recent verleden (seconden, minuten) maar ook over een verder verleden (weken, maanden, jaren). Aan de hand van de beschrijving van vroegere data kan het verleden beter begrepen worden en kunnen er (historische) inzichten ontstaan die beslissingen of acties in de toekomst zullen beïnvloeden.

**Voorbeeld van descriptieve LA (KU Leuven, België)**

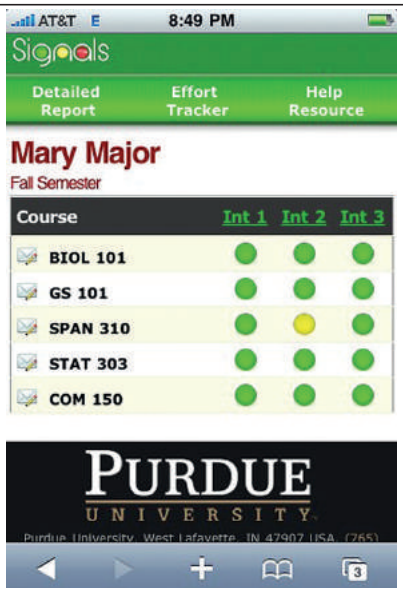
Deze visualisatie uit het REX-dashboard (Broos, Peeters, et al., 2017) toont de studievoortgang van vroegere studenten, afhankelijk van hun behaalde studie-efficiëntie in het eerste jaar. Doelstelling van de visualisatie is de huidige eerstejaarsstudenten inzicht te geven in hun mogelijke studievoortgang en op deze manier hun beslissingen en gedrag te beïnvloeden.




**Predictieve LA** wil aan de hand van de beschikbare data, die per definitie enkel iets over het verleden vertellen, de toekomst voorspellen. Deze voorspelling kan mogelijke toekomstige scenario's identificeren of verduidelijken. De voorspellingen kunnen uiteraard nooit 100% nauwkeurig zijn en bevatten dus nog een niveau van onzekerheid. Mogelijke toepassingen zijn de voorspelling van het individuele studiesucces van een student en de voorspelling van de groep van studenten die zullen falen voor een bepaald opleidingsonderdeel.

**Voorbeeld van predictieve LA (USA)**

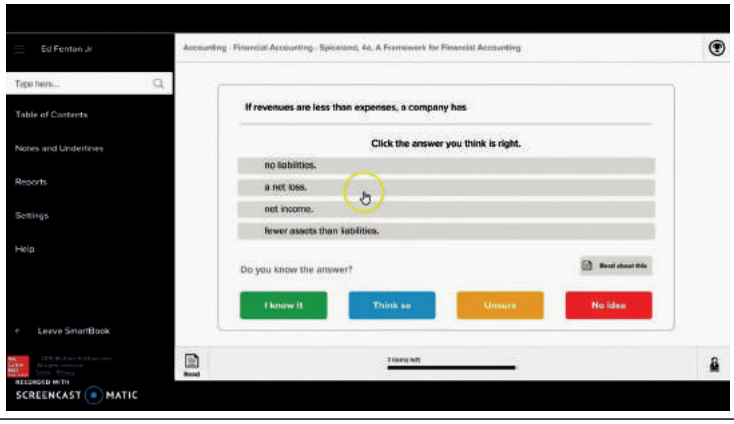
Dit is een visualisatie van Course Signals zoals gebruikt werd aan de Purdue University (Arnold & Pistilli, 2012). De visualisatie toont voor verschillende tussenliggende tijdstippen (Int 1, Int 2, Int 3) en voor verschillende vakken (bv. BIOL 101) de voorspelling van een algoritme: groen als het algoritme voorspelt dat de student vermoedelijk succesvol zal zijn, geel als er een redelijk risico op falen is en rood als het voorspelde risico op falen groot is. Het algoritme houdt rekening met de eerdere academische resultaten van de student, de interactie met het LMS, achtergrondinformatie over de vooropleiding (resultaten in secundair onderwijs, andere vooropleiding, resultaten op gestandaardiseerde tests) en persoonsgegevens. (woonplaats, leeftijd, opgenomen studiepunten).




**Prescriptieve LA** wil aan de hand van de beschikbare data iets 'voorschrijven', bijvoorbeeld een beslissing of gedrag aanbevelen. Dit gaat dus een stapje verder dan een voorspelling: het koppelt hier ook een aanbeveling aan.

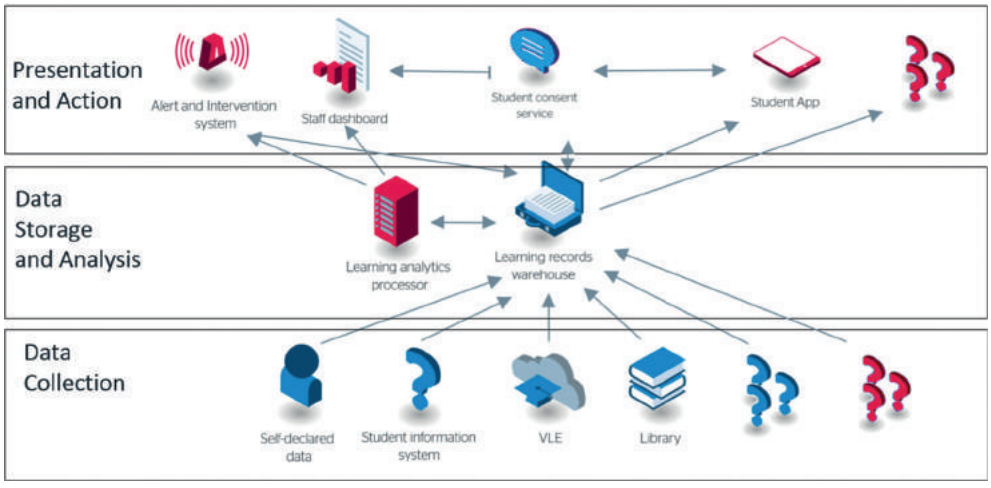
**Voorbeeld van prescriptieve LA (USA)**

LearnSmart van Mc Graw Hill is een interactief studeerinstrument dat de vaardigheden en kennis van een student binnen een onlinecursus evalueert en opvolgt. Het instrument identificeert de onderdelen die de student al beheerst en andere die nog oefening vergen. Afhankelijk van de antwoorden van studenten op vragen in de onlinecursus en het vertrouwen dat ze hierbij rapporteerden, biedt het instrument individuele leerpaden aan.



## E. Technologische oplossingen voor LA

Gezien de grote interesse in LA komen er geleidelijk aan meer oplossingen op de markt en zijn er ook meer oplossingen beschikbaar via de opensourcegemeenschap. Er zijn toepassingen die een deeloplossing bieden voor een van de componenten van het LA-proces en er zijn er die verschillende aspecten combineren in een volledige oplossing (Sclater, 2017). Het is daarom belangrijk verschillende componenten van een LA-omgeving nader te bekijken. Het schema hieronder toont de lagen en processen van een LA-architectuur zoals die opgezet werd in het Verenigd Koninkrijk door JISC[5] en opengesteld werd voor alle instellingen van hoger onderwijs, die deel wensten te nemen aan LA-pilootstudies (Baylis, 2019).




*Figuur 5: Verschillende lagen en processen van een LA-architectuur zoals die opgezet werd in het Verenigd Koninkrijk door JISC en opengesteld werd voor alle instellingen van hoger onderwijs die daar deel wensten te nemen aan LA pilots* (Baylis, 2019).

In deel III werden de verschillende soorten data van de eerste laag, dataverzameling, al besproken. Data komen dikwijls uit een LMS en SIS, maar ook van andere databronnen, zoals bibliotheken. Deze data worden bij voorkeur gelogd aan de hand van een datastandaard en worden per betrokkene verzameld in een centrale *Learning Record Store* (LRS) (IV.A.2). Datasporen uit de verschillende bronnen kunnen hier aan elkaar gekoppeld worden.

In de tweede laag zorgt een LA-processor voor het aggregeren, analyseren (zie IV.B en IV.C) en naverwerken van de data, zodat ze bruikbaar zijn voor de bovenste laag met de toepassingen (zie II).

[5] JISC is in het Verenigd Koninkrijk een non-profitorganisatie voor digitale diensten en oplossingen in het hoger en voortgezet onderwijs (https://www.jisc.ac.uk/).

Er zijn verschillende soorten technologische oplossingen beschikbaar (ook hier telkens zowel open source als gepatenteerde mogelijkheden) die één of meer van de technologische modules (SIS, LMS, LRS, data-analyse, visualisatie, dashboards) aanbieden:

- Producten die **LA** aanbieden **vanuit het LMS** (betreffende het leren dat hierbinnen plaatsvindt) en onder andere aan de hand van dashboards rapporteren over de voortgang in de cursus (gespendeerde tijd, sociale interactie, gebruik van leermiddelen) of prestaties/outputs voor tests en opdrachten. Een LA-module kan ook als externe plug-in extra functionaliteit toevoegen aan de leeromgeving.
- ***Student Information System*** **(SIS) geïntegreerd in het LMS**: dit maakt het mogelijk om leerdata rechtstreeks te koppelen aan studentengegevens om inzichten te genereren.
- ***Learning Record Store*** **(LRS) geïntegreerd in het LMS**. Veel LMS-providers hebben de overstap gemaakt naar het gebruik van een centraal datawarehouse voor leerdata uit verschillende bronnen. De meeste van deze providers hebben gekozen om deze te integreren in hun leeromgeving. Sommigen hebben hun eigen LRS ontwikkeld, terwijl andere een bestaande LRS hebben geïntegreerd.
- In plaats van een LRS te integreren zijn er ook LMS-systemen die een API-gebaseerde toegang bieden tot een extern centraal datawarehouse. Voordeel hiervan is dat de gebruiker een eigen LRS met leerdata van studenten kan gebruiken. De gebruiker moet het LMS configureren om bepaalde statements in deze externe LRS op te slaan. Alle communicatie tussen LMS en LRS wordt in dit geval afgehandeld via API's.
- Een **LRS** kan worden aangeboden **met of zonder ingebouwde data-analysetoepassingen**. Ook ***Business Intelligence*** **(BI)-systemen**, die vaak al in gebruik zijn van universiteiten om onder meer studievoortgang en *drop-out* op het niveau van de instelling te meten, breiden uit met nieuwe functionaliteiten om ook leerdata te integreren en te visualiseren in dashboards.

Het is niet evident voor instellingen van hoger onderwijs om in hun eentje een volledige LA-architectuur op te zetten en te kiezen voor specifieke oplossingen of applicaties. Daarom zijn nationale organisaties voor digitale ondersteuning en diensten, zoals JISC in het Verenigd Koninkrijk, SURF[6] in Nederland of het wereldwijde netwerk van de Apereo Foundation, zo belangrijk. Ze bieden diensten, infrastructuur en expertise aan voor het opzetten van LA binnen instellingen van hoger onderwijs en onderhandelen met grote spelers, zoals LMS-aanbieders. De samenwerking tussen JISC en Blackboard heeft bijvoorbeeld geleid tot de ontwikkeling van een *xAPI building block* voor het verzamelen van leerdata binnen *Blackboard Learn*-leeromgevingen en het zenden van de data naar een centrale *Learning Record Store*. Nationale organisaties laten instellingen mogelijk toe om

[6] SURF is de ICT-samenwerkingsorganisatie van het onderwijs en onderzoek in Nederland (https://www.surf.nl/).

concrete ervaring op te doen in het domein van LA en beter geïnformeerde keuzes te maken over de diensten en oplossingen die voor de organisatie het meest geschikt zijn.

---

*Aanbeveling 10: Nationale samenwerkingsorganisaties op het vlak van onderwijstechnologie hebben in Nederland en Engeland aangetoond dat ze een belangrijke steun kunnen zijn voor instellingen van hoger onderwijs bij de overweging en implementatie van LA. Ze hebben in deze landen hierdoor niet enkel een impuls gegeven aan de concrete toepassing van LA, maar ook aan het onderzoek hierrond. Een Vlaamse samenwerkingsorganisatie – of aansluiting bij een andere nationale samenwerkingsorganisatie – kan voor Vlaanderen een gelijkaardige rol spelen en zorgen voor opschaling.*

---

Een belangrijke overweging is alvast de modulariteit van de technologische oplossingen. Hoewel een geïntegreerde LA-oplossing in eerste instantie aantrekkelijk is, beperkt die de flexibiliteit en keuzevrijheid. Dankzij standaarden die in ontwikkeling zijn wordt het realistischer dat instellingen van hoger onderwijs zelf hun technologische LA-ondersteuning samenstellen, aan de hand van modules van verschillende aanbieders en opensourcebronnen.

Bij het overwegen van technologische oplossingen die zo ver gaan dat ze concrete LA-implementaties aanbieden dient de instelling van hoger onderwijs eerst belangrijke vragen te beantwoorden: past dit in de context en visie, zowel qua doelstellingen als qua aanpak? Komt de aangeboden oplossing ook tegemoet aan een bestaande vraag? Is er een draagvlak bij de toekomstige gebruikers, zoals studieloopbaanbegeleiders en studenten? Is er een aangetoonde impact? Kan de instelling met de eigen expertise en middelen een eigen en misschien beter afgestemde oplossing realiseren? Hierbij zal het ook belangrijk zijn dat zij het beheer van en de verantwoordelijkheid over de LA-data kunnen waarnemen.

# V. Randvoorwaarden

Dit onderdeel behandelt twee belangrijke dimensies in de randvoorwaarden voor LA: privacy- en wettelijke aspecten, en vervolgens ethische aspecten

## A. Privacy- en wettelijke aspecten: juridische en organisatorische instrumenten

LA houdt in dat data of informatie over geïdentificeerde of identificeerbare personen worden verwerkt. Deze data of informatie worden beschouwd als 'persoonsgegevens' .

**Wat zijn persoonsgegevens?**

Persoonsgegevens zijn alle gegevens die direct of indirect herleidbaar zijn tot een persoon. Een naam of adres is een persoonsgegeven, maar ook gegevens over gedrag vallen hieronder. Bijhouden wat iemand in een leeromgeving doet, is een vorm van persoonsgegevens verzamelen.

Bron: Engelfriet et al., 2017. Zie ook de definitie van persoonsgegevens in art. 4 van de AVG (EU 2016/679)

De betrokken personen zullen doorgaans studenten zijn, maar ook docenten, studieloopbaanbegeleiders, cursusverantwoordelijken of andere gebruikers van digitale leeromgevingen kunnen betrokkenen zijn. Over deze personen zal rechtstreeks of onrechtstreeks informatie ingezameld en verwerkt worden, bijvoorbeeld via de logging van navigatie-informatie.

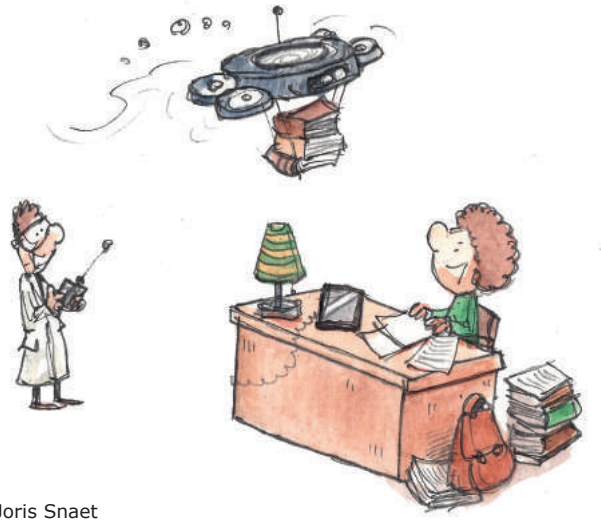
Persoonsgegevens vallen onder de wetgeving Gegevensbescherming. Die werd aangescherpt door de Europese Algemene Verordening Gegevensbescherming (EU 2016/679: de AVG, die met haar Engelse afkorting ook bekendstaat als de GDPR). De AVG, aangevuld met nationale bepalingen, biedt een algemeen kader voor de wettelijke bescherming voor alle gebruik en de verwerking van persoonsgegevens in de Europese Unie. De AVG is dus ook van toepassing op LA: zowel overheids(instellingen) als private partijen die LA gebruiken en voor de verwerking verantwoordelijk zijn, moeten de AVG naleven. Gegevensbescherming bij LA vergt overigens bijzondere aandacht omdat LA doorgaans gebruik maakt van een combinatie van doorgedreven data-inzameling en -opslag, de analyse van kleine en grote hoeveelheden data, profilering en artificiële intelligentie (AI). Deze nieuwe technieken blijven vooralsnog een uitdaging voor het recht. De AVG verplicht om de risico's bij dergelijke gegevensverwerkingen beter in te schatten (risico-gebaseerde aanpak) (Kindt, 2017) (V.A.4). Bij het gebruik van cookies op digitale platformen, locatie- en verkeersgegevens van elektronische communicatie zijn daarenboven ook nog de wettelijke verplichtingen van de Europese ePrivacy-wetgeving (2002/58/EG) van toepassing. Deze wetgeving wordt momenteel herzien. Op haar specifieke bepalingen gaan we niet verder in.

Voor we de verplichtingen onder het gegevensbeschermingsrecht, zoals vastgelegd in de AVG, bespreken, merken we nog op dat LA ook en bovenal steeds de fundamentele rechten en vrijheden zoals neergelegd in het *Handvest van de grondrechten van de Europese Unie* moet respecteren. Die kunnen door LA-technieken in het gedrang komen. De fundamentele rechten omvatten onder andere het recht op privacy (art. 7 *Handvest*) en op gegevensbescherming (art. 8 *Handvest*). Indien LA raakt aan deze rechten, bijvoorbeeld bij voortdurende monitoring – een risico dat bestaat bij gebruik van onder meer *wearables*, een allesomvattende opslag in *Learning Record Stores* en/of een doorgedreven koppeling, combinatie en gebruik van beschikbare data over diverse platformen heen, prescriptief gebruik of supergedetailleerde leerprofielen van studenten (zie IV) – zal men hiervoor een wettelijke basis nodig hebben. Men zal een legitiem doel en de **noodzaak en evenredigheid** moeten kunnen aantonen. Ook al kunnen bepaalde LA-data een waardevol instrument zijn bij het meten van studie-efficiënte en dus geschikt zijn voor een doel van algemeen belang, en al zijn algoritmes nuttig om verschillen te meten of leeromgevingen aan te passen, beperkingen van het recht op respect voor het privéleven en gegevensbescherming door LA verwerkingen dienen steeds binnen de grenzen van het strikt noodzakelijke te blijven. Bij conflicterende rechten of belangen, zoals het belang dat een overheid heeft bij het meten van studievoortgang en -efficiëntie tegenover privacy- en gegevensbescherming en de vrijheid van studenten om hun leer- en studiegedrag zelf te organiseren, zal een zorgvuldige afweging nodig zijn. Het daadwerkelijk beantwoorden aan een voor België of de Europese Unie algemeen belang, zoals bijvoorbeeld het nastreven van gelijkekansenonderwijs, kan een legitiem doel zijn.

Maar elke onderwijsinstelling moet zich dus ook steeds de vraag stellen of een dataverwerking mogelijk een inmenging is, en indien dit zo is, wat de strikte noodzaak, proportionaliteit en effectiviteit daarvan zijn. Zo zal men de strikte noodzaak, proportionaliteit en effectiviteit én ook de volledige centralisatie van alle studentengegevens en -gedragingen uit verschillende bronnen en van alle studenten – in plaats van bijvoorbeeld alleen de studenten die studiebegeleiding hebben aangevraagd – voor LA-doeleinden moeten kunnen aantonen en bewijzen. Ook moet de verantwoordelijke onderwijsinstelling zich afvragen in welke mate en onder welke voorwaarden het rechtmatig is om bijvoorbeeld op basis van LA studiekeuzes van studenten te beperken. Instellingen zullen ook bij het induceren van gedragsveranderingen omzichtig te werk moeten gaan.

Dit zijn slechts enkele voorbeelden van het risico op een inbreuk op fundamentele rechten, zoals privacy- en gegevensbescherming, het recht op vrijheid van meningsuiting en het recht op onderwijs. De instelling is zelf verantwoordelijk om deze risico's in te schatten en de nodige waarborgen en maatregelen te nemen om ze te beperken. Ook discriminatie loert om de hoek, net als andere ethische vragen (zie V.B). De toetsing aan fundamentele rechten komt enigszins aan bod in de AVG en in de effectbeoordeling (zie hieronder), maar moet hier toch los van gezien worden.

Onder de AVG moet elke instelling en elke andere (rechts)persoon of instantie die toegang heeft en/of gebruik maakt van LA-persoonsgegevens in de eerste plaats nagaan wie de (mede)verantwoordelijke is voor de dataverwerkingen. Om dit te bepalen kijkt men wie alleen dan wel met andere instellingen of derden de beslissingen neemt over (1) de doeleinden van het gebruik van de LA-data en (2) de hiervoor gebruikte (technische) middelen. De verantwoordelijke zal in vele gevallen de onderwijsinstelling zijn die LA gebruikt, maar soms ook de platformuitbaters of zelfs IT-leveranciers, al dan niet samen met de instelling. Op deze verantwoordelijke(n) zijn de verplichtingen van de AVG van toepassing, op straffe van hoge boetes en van bijvoorbeeld ook de blokkering van de data. Bij digitale platformen en instrumenten, waarop soms diverse instanties aangesloten zijn die er toegang toe hebben, zoals bij sommige commerciële oplossingen, is bepalen wie de verantwoordelijke(n) zijn mogelijk niet eenvoudig. Indien meerdere verantwoordelijken samen de beslissingen nemen en de data gebruiken, bijvoorbeeld een onderwijsinstelling samen met een onderzoeksorganisatie, of meerdere hogeronderwijsinstellingen, zullen afspraken gemaakt moeten worden wie de verplichte informatie aan de betrokkenen zal verstrekken en waar zij kunnen aankloppen voor de uitoefening van hun rechten. De verantwoordelijke mag ook niet verward worden met de rol van de verwerkers van LA-data, zoals bijvoorbeeld ondersteunende commerciële IT-bedrijven, die uitsluitend ten behoeve van verantwoordelijke(n) de data verwerken. Zij worden soms ingeschakeld voor bijvoorbeeld de opslag van de gegevens. Op deze verwerkers rust in hoofdzaak de verplichting om de data te beveiligen en de bescherming van de rechten van betrokkenen te waarborgen, naast nog enkele andere nieuwe verplichtingen onder de AVG. Een schriftelijke overeenkomst tussen de verantwoordelijke en de verwerker is in dit geval nodig.



© Joris Snaet

In dit onderdeel vermelden we eerst beknopt de transversale principes en enkele belangrijke verplichtingen voor elke LA-verwerking. Vervolgens komen enkele AVG-verplichtingen aan bod waarmee specifiek LA rekening moet houden: de verplichtingen van de verwerkingsverantwoordelijke bij uitsluitend geautomatiseerde besluitvorming en profilering, en de verplichting om een effectbeoordeling te maken. Daarnaast is er ook kort aandacht voor het gebruik van LA voor statistiek of wetenschappelijk onderzoek.

## 1. Transversale principes van het gegevensbeschermingsrecht

*a) Doelbepaling en -binding*

Het is van cruciaal belang dat de verantwoordelijke in de eerste plaats voor elke verwerking van LA-persoonsgegevens (inzameling, opslag, gebruik, doorgeven enzovoort) (een) uitdrukkelijk omschreven en gerechtvaardigd(e) doeleinde(n) bepaalt (art. 5.1.b AVG; zie ook II). Dit geldt zowel voor een afzonderlijke verwerking, bijvoorbeeld doorgeven, maar ook voor een geheel van verwerkingen. Een precieze omschrijving van elke doelstelling is dus nodig (**doelspecificatie**). Zomaar persoonsgegevens verzamelen kan niet. Dit is anders voor data die niet aan geïdentificeerde of met redelijke inspanningen identificeerbare personen gekoppeld kunnen worden en die dus geen persoonsgegevens zijn.

Cruciaal bij de register- en informatieverplichting (zie hieronder) maar ook voor andere AVG- verplichtingen zijn dus de noodzaak en de verplichting om voor elk gebruik te bepalen – en in het register en ook in de informatie aan betrokkenen dus nauwkeurig te beschrijven – voor welke doeleinden de data gebruikt worden. Niet alleen het wat en van wie, maar ook het gebruik waarvoor staat centraal, voor elke verwerking. Twee voorbeelden: met welk doel wordt een uitleengeschiedenis van de bibliotheek bij LA betrokken en gebruikt? En: wat is het doel van het zichtbaar maken van LA-data aan andere betrokkenen dan studieloopbaanbegeleiders? Als die doeleinden eenmaal zijn bepaald en omschreven, moeten de verantwoordelijke en iedereen die met de data binnen de organisatie omgaat, de doelomschrijving(en) ook te allen tijde respecteren bij elke verwerking (**doelbinding**). Doelspecificatie en doelbinding zijn spilverplichtingen onder de AVG waaraan vele andere verplichtingen gekoppeld zijn, waaronder de minimale gegevensverwerking en -opslag en de informatieverplichting. Het is dus steeds het vooraf vastgelegde doel dat zal bepalen welke gegevens nodig zijn en verwerkt kunnen worden. Niet andersom.

*b) Rechtmatige, behoorlijke en transparante LA-verwerkingen*

Een tweede, zeer belangrijk transversaal principe houdt in dat data steeds op **rechtmatige, behoorlijke en transparante wijze verzameld en verwerkt** moeten worden (art. 5.1.a AVG). De verantwoordelijke zal zich dus ook de vraag moeten stellen of de dataverwerking rechtmatig is, en wel voor elke LA-verwerking

of geheel van verwerkingen, zoals bijvoorbeeld bij de inzameling bij een bepaalde bron of op een bepaalde wijze (bv. met *wearables*), of ook bij de opslag. Onrechtmatig zou bijvoorbeeld kunnen zijn: voortdurend toezicht (*surveillance*) houden over alle leeractiviteiten van alle studenten of docenten. Omdat de 'digitale voetafdruk' toeneemt, neigen het verzamelen, koppelen en de centrale opslag van alle digitale activiteiten van alle studenten op leerplatformen naar *surveillance*. Volgens vaste rechtspraak van het Hof van Justitie vereisen de bescherming van het fundamenteel recht op privacy en de bescherming van persoonsgegevens dat uitzonderingen en beperkingen enkel bij wet mogen, voor een wettelijk legitiem doeleinde, en dat ze binnen de grenzen van het strikt noodzakelijke moeten blijven (zie ook hierboven). Verder liggen automatische besluitvorming en profilering (zie verder) erg gevoelig. Omdat over de rechtmatigheid de meningen kunnen verschillen, is het wenselijk dat er een breder debat komt over wat bij LA rechtmatig kan zijn en wat niet, wat behoorlijk is en wat niet, en hoe de rechten van de betrokkenen gevrijwaard kunnen worden. Anderzijds is duidelijk dat transparantie steeds en altijd van primordiaal belang blijft. Heimelijk LA-data inzamelen en gebruiken kan als algemene regel niet. De verantwoordelijke zal dus informatie moeten verstrekken, zoals voor elke logging van data op platformen, bij *games* enzovoort, voor de koppelingen, en voor het descriptieve, prescriptieve en predictieve gebruik van de data.

*c) Nog vijf andere algemene gegevensbeschermingsprincipes*

Er gelden nog vijf andere algemene principes waaraan alle LA-verwerkingen moeten beantwoorden (art. 5 AVG). Vooreerst is er het belangrijke principe van de **minimale gegevensverwerking** (art. 5.1.c AVG). Dit houdt in dat alle voor LA bestemde persoonsgegevens toereikend moeten zijn, maar ook ter zake dienend en beperkt tot wat noodzakelijk is voor de LA-doeleinden. Een onderwijsinstelling kan dus niet met het oog op LA massaal veel persoonsgegevens inzamelen, opslaan of gebruiken zonder dat ze noodzakelijk zijn, laat staan niet eens nuttig zijn om de vastgestelde LA-doeleinden te bereiken. Voor de ontwikkeling van LA-profielen of -groepsprofielen betekent dit bijvoorbeeld in dit geval ook het gebruik van gepseudonimiseerde en waar mogelijk geanonimiseerde gegevens. Ook moet de instelling van hoger onderwijs als verantwoordelijke de vraag stellen of een welbepaalde databron – zoals de studentenadministratie, logs uit leeromgevingen, het gebruik van mobiele *apps*, data uit *games* en bibliotheekbezoeken – wel toegevoegde waarde biedt voor het beoogde en gespecificeerde LA-doel.

Een ander algemeen principe is dat de gebruikte **gegevens** ook **juist** moeten zijn en dat de gegevens **geactualiseerd** moeten zijn (art. 5.1.d AVG). Een derde principe zegt dat ze slechts **beperkt opgeslagen** mogen worden en dat de betrokkenen niet langer geïdentificeerd mogen worden dan nodig (art. 5.1.e AVG). Bij LA is er vaak nood aan datavoorverwerking en is er sprake van het gebruik van data afgeleid van andere datagroepen en van data gebaseerd op

verbanden en kansberekeningen. Deze data zijn mogelijk minder accuraat en kunnen dus een probleem vormen. De onderwijsinstellingen moeten zich in elk geval behoeden voor het '*garbage in, garbage out*'-fenomeen. Sommige van de genoemde principes, evenals de correcte opname van de afgeleide en predictieve data in het register en de inschatting in de effectbeoordeling (zie hieronder) zullen voor LA een ware uitdaging vormen.

Een vierde principe stelt dat voor elke LA-verwerking de **integriteit en vertrouwelijkheid** gerespecteerd moeten worden (art. 5.1.f AVG). Dit gaat over de vraag wie toegang dient te hebben tot de LA-gegevens en of die gegevens voldoende beveiligd zijn. Aangepaste technische en organisatorische veiligheidsmaatregelen zijn dan ook onontbeerlijk. Ten slotte zal de verantwoordelijke aansprakelijk en verantwoordelijk gehouden worden voor de naleving van al deze principes en moet hij/zij ook kunnen bewijzen dat deze principes en elk van de AVG-verplichtingen worden nageleefd (art. 5.2 AVG) (Kindt, 2017).

## 2. Andere verplichtingen onder het gegevensbeschermingsrecht

*a) Wettelijke grond en beveiliging*

Persoonsgegevens voor LA kunnen niet zomaar ingezameld en gebruikt worden zonder één (of meerdere) wettelijke gronden, zoals omschreven in de AVG. Daarop moeten elke verwerking en het gebruik dus gebaseerd worden. Is er bijvoorbeeld een ondubbelzinnige, vrije, geïnformeerde en specifieke toestemming? Of een gerechtvaardigd belang van de instelling, weliswaar met respect voor de rechten en vrijheden van de betrokkenen? De instelling zal een *informed consent* ook moeten kunnen bewijzen. Indien de toestemming bijvoorbeeld online werd gegeven, zal ze dit moeten bijhouden, net als een kopie van de informatie die aan de betrokkene verstrekt werd. Ook de toestemming voor dataverwerkingen moet steeds vrij zijn en daarom ook losstaan van een eventuele toestemming voor een andere (contractuele) aangelegenheid. Daar mag ze dus niet meer aan gekoppeld worden. Voor uitsluitend geautomatiseerde besluitvorming, waaronder profilering, is er bovendien een uitdrukkelijke (expliciete) toestemming vereist (zie hieronder).

In totaal omschrijft de AVG zes gronden, waarvan er dus voor elke (of groep van) verwerking(en) voor een specifiek LA-doel minstens één toepasselijk moet zijn. Andere wettelijke gronden die voor welbepaalde LA-dataverwerkingen een rol kunnen spelen zijn ook de noodzaak van de verwerkingen voor een wettelijke verplichting, voor een taak van algemeen belang (zoals het meten van de studie-efficiëntie) of voor de uitvoering van een overeenkomst met de betrokkenen (zoals het bijhouden van academische data). Voor speciale categorieën van persoonsgegevens (zie ook hierna) geldt een verbod op verwerking, behalve als er een uitzondering van toepassing is.

Behalve wettelijke gronden voor de verwerkingen van LA-data is ook het treffen van passende technische en organisatorische maatregelen die zijn afgestemd op de risico's verplicht. Dit omvat, doorgaans en waar passend, onder meer de pseudonimisering en versleuteling van persoonsgegevens, en een regelmatige evaluatie van de doeltreffendheid van de beveiligingsmaatregelen.

*b) Vastleggen van LA-verwerkingen in een register*

Elke verantwoordelijke en verwerker van LA-data moet in een register een gedetailleerde beschrijving opnemen van de LA-verwerkingen, met daarin onder andere de categorieën van de persoonsgegevens en van personen waarop de gegevens betrekking hebben, en de LA-doeleinden. Verwerking slaat op elke bewerking of geheel van bewerkingen van de data, met inbegrip van het inzamelen, het vastleggen en het gebruiken, maar ook het ordenen, bijwerken, raadplegen, wissen, doorzenden of op enige andere wijze ter beschikking stellen.

Een verantwoordelijke instelling zal met andere woorden de **data(stromen) in kaart moeten brengen** en beschrijven welke datacategorieën voor LA worden ingezameld en gebruikt, en over wie. Het kan hierbij dus gaan over onder andere data die rechtstreeks bij de betrokkenen zijn verkregen, zichtbaar of onzichtbaar ingezameld, of die onrechtstreeks werden verkregen, bijvoorbeeld door ze te ontvangen van derden, maar het betreft ook het gebruik van data waarover de instelling zelf beschikt en de combinatie daarvan. Verschillende mogelijke bronnen en types van data kwamen aan bod in III. De wijze van het verkrijgen van de data is ook van belang voor het tijdstip en de informatie die aan de betrokkenen gegeven worden over die data. Ook het eventuele doorgeven van LA buiten de Europese Unie, de bewaartermijnen en de beveiligingsmaatregelen voor de persoonsgegevens moeten in het register schriftelijk worden vastgelegd. Hoewel er in het algemeen voor kleinere organisaties een uitzondering is op de verplichting om een register bij te houden, is een register voor LA-verwerkingen aanbevolen en allicht vereist, omdat LA profilering inhoudt en niet zonder risico's is voor de rechten en vrijheden van de betrokkenen. Enkele risico's bij het inschatten van de rechtmatigheid kwamen al aan bod, maar er zijn ook andere, zoals het opsporen of het gebruik van regressietechnieken, psychometrie en de risico's van impliciete vooroordelen en discriminatie. De Belgische Gegevensautoriteit (zie https://www.gegevensbeschermingsautoriteit.be/), die nu ook voor het eerst de mogelijkheid heeft om administratieve boetes op te leggen, heeft nuttige richtlijnen gegeven over de verplichting om een register op te stellen, bij te houden en regelmatig bij te werken (Aanbeveling 06/2017). Het interne register moet het voor de verantwoordelijke instelling mogelijk maken om een goed inzicht te verkrijgen in alle dataverwerkingen, maar ook in de daarmee gepaard gaande AVG-verplichtingen. Het moet op verzoek of bij een onderzoek ook ter beschikking worden gesteld van de autoriteit.

*c) Aanstelling van een gegevensfunctionaris*

De instellingen van hoger onderwijs moeten – indien ze een overheidsinstelling zijn – een gegevensbeschermingsfunctionaris aanstellen die betrokken wordt bij alle aangelegenheden met LA-data en die toeziet op de naleving van de AVG. Andere instellingen, verantwoordelijke(n) en verwerker(s) van LA-data zullen eveneens het best een functionaris aanstellen.

---

*Aanbeveling 11: Onderwijsinstellingen moeten een register met een nauwkeurige beschrijving van alle data(stromen) en doeleinden van LA bijhouden en regelmatig bijwerken. Laat de gegevensbeschermingsfunctionaris hierbij een centrale rol spelen en laat de functionarissen van verschillende instellingen hun ervaringen uitwisselen.*

---

*d) Informatie aan betrokken studenten en docenten en hun rechten*

De LA-verwerkingen moeten transparant zijn voor alle betrokkenen. Dit betekent in de eerste plaats dat volgens de AVG wettelijke en welbepaalde informatie moet gegeven worden bij de inzameling, onder andere over wie, wat, waarvoor en hoe de LA-data gebruikt worden. Dat is ook van belang bij een eventuele *informed consent*. Geheime inzameling is dus – zoals reeds vermeld – uit den boze.

**AVG en het einde van de standaardinstemming**

AVG vereist dat er voor uitsluitend geautomatiseerde besluitvorming met rechtsgevolgen of gelijkaardige aanzienlijke effecten, waaronder profilering, een uitdrukkelijke vrije, geïnformeerde, specifieke en ondubbelzinnige wilsuiting is van de betrokkene (artikel 22.2 AVG), en ook passende waarborgen. Deze toestemming zal het best blijken uit een uitdrukkelijke verklaring, duidelijk afgescheiden van enige andere verklaringen, in duidelijke en eenvoudige taal. Dit verdient de voorkeur boven een ondubbelzinnige actieve *handeling*. In een onlineomgeving kan een uitdrukkelijke toestemming gegeven worden via bijvoorbeeld een afzonderlijk toestemmingsscherm, waar de betrokkenen een vakje moet aanvinken waarin duidelijk omschreven is waarmee ze akkoord gaan.

De AVG legt verder strengere voorwaarden op voor de toestemming en heeft dus ook een einde gemaakt aan de 'standaardinstemming' die vroeger regelmatig werd gehanteerd, zoals bijvoorbeeld de beweerde toestemming door de algemene verplichte aanvaarding van algemene voorwaarden of door het verdere gebruik van een website terwijl er geen echte keuze is. Daarnaast moet de betrokkene vóór de vraag tot instemming de nodige informatie hebben gekregen, in duidelijke en eenvoudige taal. Ten slotte hebben betrokkenen het recht om hun toestemming steeds in te trekken. In dit geval zal de instelling de verwerking(en) onverwijld stopzetten.

De AVG bepaalt in detail over welke informatie het precies gaat, ook als de data onrechtstreeks verkregen worden van derden of binnen de instelling. In dit tweede geval moeten de betrokkenen binnen een redelijke termijn, maar

uiterlijk binnen één maand, geïnformeerd worden. De informatieverplichting geldt dus ook voor alle data over geïdentificeerde of identificeerbare studenten, docenten of begeleiders die zijn verkregen via digitale tools, inclusief platformen. Alle informatie, ook over hun rechten, moet tegelijk beknopt en transparant zijn, begrijpelijk en in gemakkelijk toegankelijke vorm, en in duidelijke en eenvoudige taal. De verantwoordelijke instelling zal dus de nodige aandacht moeten geven aan deze informatieverplichting. Bij geautomatiseerde besluitvorming en profilering geldt de informatieverplichting overigens ook voor het gebruik hiervan en vereist ze bijvoorbeeld nuttige informatie over de gehanteerde logica (zie hieronder). Dit ligt niet steeds voor de hand bij de soms complexe LA-verwerkingen. Betrokkenen hebben verder het recht op inzage in al hun gegevens en op bijkomende informatie over elk gebruik, evenals op de rectificatie van onjuiste persoonsgegevens en de beperking van verwerkingen. Betrokkenen kunnen ook te allen tijde bezwaar aantekenen om redenen die eigen zijn aan hun specifieke situatie indien de instelling van hoger onderwijs zich zou beroepen op de noodzaak van LA-verwerkingen voor een taak van algemeen belang of op de noodzaak voor een eigen legitiem belang, inclusief bij profilering. In dit geval is er een omkering van de bewijslast en zal de instelling *moeten aantonen dat haar dwingende gerechtvaardigde gronden zwaarder wegen dan de belangen, rechten en vrijheden van de betrokkene. (art 21.1 AVG)*

*e) Recht op gegevenswissing ('recht op vergeten') en kennisgeving bij een 'datalek'*

Instellingen moeten er zich ook op voorbereiden dat ze data kunnen wissen. Een nieuw recht onder de AVG vereist immers dat men onder voorwaarden persoonsgegevens kan verwijderen op verzoek, bijvoorbeeld bij de intrekking van de toestemming. Het wissen moet dan gebeuren zonder onredelijke vertraging.

Ingeval een inbreuk op de beveiliging per ongeluk of op onrechtmatige wijze leidt tot bijvoorbeeld het verlies of het vrijgeven van data, dient de instelling de autoriteit zonder vertraging en uiterlijk binnen de 72 uur in te lichten, tenzij een risico niet waarschijnlijk is. Ook de betrokkenen moeten onverwijld op de hoogte gebracht worden als er voor hun rechten en vrijheden waarschijnlijk een hoog risico is.

*f) Gegevensbescherming by design en by default*

De AVG legt ook de verplichting op om van bij het ontwerp van LA-systemen gegevensbescherming toe te passen en gepaste (technische) beschermings-maatregelen te nemen. Gegevensbescherming wordt dan ingebouwd *by design*, zodat privacy- en gegevensbescherming ook de uitgangspositie wordt *by default*. Anonimisering en pseudonimisering om dataminimisatie te bewerkstelligen zijn hiervan voorbeelden.

*Aanbeveling 12: Gegevensbescherming en privacy moeten van bij het begin ingebouwd worden in de architectuur van LA-systemen en -procedures, zodat bescherming het uitgangspunt is by default, en privacy- en gegevensbescherming ingebakken zijn by design.*

*g) Gedragscode en certificering*

Het is aan te bevelen dat de onderwijssector voor LA-verwerkingen in overleg een gedragscode opstelt. Zo'n code kan immers rekening houden met de specifieke kenmerken en noden van LA-verwerkingen en verduidelijken hoe de verplichtingen van de AVG nageleefd worden, bijvoorbeeld inzake de pseudonimisering van LA-gegevens. Die kunnen dan voor goedkeuring voorgelegd worden. Gedragscodes inzake LA die over instellingen heen worden ontwikkeld, kunnen dus bijdragen tot een correct gebruik van LA.

Ook de certificering van gebruikte technieken is een te onderzoeken piste.

*Aanbeveling 13: Vlaamse instellingen van hoger onderwijs moeten, gezien de complexiteit van de wettelijke aspecten en het facet 'privacy', ondersteuning bieden aan alle personeelsleden (docenten, technologen, onderzoekers, studentenbegeleiders) die betrokken zijn bij LA. De ondersteuning mag niet enkel bestaan uit een concretisering van het bestaande wettelijke kader, maar moet ook gedragscodes en persoonlijk advies voor welbepaalde dossiers bevatten. Hiertoe is samenwerking met de ethische commissies en met de gegevensbeschermingsfunctionarissen van belang.*

## 3. Profilering en geautomatiseerde individuele besluitvorming

*a) Algemeen: het recht om niet onderworpen te worden aan besluiten uitsluitend gebaseerd op geautomatiseerde verwerking, inclusief profilering*

Als uitgangspunt geldt dat studenten, docenten en anderen over wie LA-informatie verwerkt wordt, het recht hebben om niet onderworpen te worden aan besluiten die uitsluitend zijn gebaseerd op geautomatiseerde verwerking en waaraan (i) rechtsgevolgen zijn verbonden of (ii) die de betrokkene vergelijkbaar in aanzienlijke mate treffen (art. 22.1 AVG). Het opstellen en gebruiken van profielen wordt hierbij expliciet vermeld. Slechts in drie gevallen is geautomatiseerde besluitvorming uitzonderlijk mogelijk: (a) wanneer een dergelijk besluit noodzakelijk is voor een overeenkomst, zoals bijvoorbeeld het onderwijscontract, (b) bij een wettelijke bepaling of (c) indien er een uitdrukkelijke vrije, geïnformeerde, specifieke en ondubbelzinnige wilsuiting is van de betrokkene (artikel 22.2 AVG).

Elk LA-systeem zal bijgevolg zeer omzichtig moeten omspringen met het opmaken en toepassen van profielen en van besluitvorming. Bij mogelijke wezenlijke (rechts)

gevolgen zal de onderwijsinstelling dus steeds in een menselijke tussenkomst of een uitzonderingsprocedure moeten voorzien. Een heikel punt is de kwestie wanneer de LA-gebruiker zich kan beroepen op een 'gerechtvaardigd belang' in plaats van de uitdrukkelijke toestemming, omdat er minder zwaarwichtige gevolgen aan de profilering en besluitvorming verbonden zouden zijn, zoals bijvoorbeeld voor automatische aanpassingen in leeromgevingen, het indelen bij bepaalde groepen, automatische feedback op LA platformen enz. Men kan echter argumenteren dat dit onder dit verbod valt, behoudens uitzondering, omdat deze beslissingen geautomatiseerd zijn en men ervan kan uitgaan dat dit de betrokkenen vergelijkbaar aanmerkelijk treft.

De uitdrukkelijke toestemming met een uitsluitend geautomatiseerde besluitvorming moet dus geïnformeerd en expliciet zijn. De voorkeur wordt gegeven om deze toestemming doorgaans door een verklaring (eerder dan een handeling) en schriftelijk te verkrijgen, en eventueel te voorzien van een handtekening, die ook later als bewijs kan dienen. In een uitsluitende onlinecontext kan een uitdrukkelijke toestemming gegeven en bewezen worden door het invullen van een elektronisch document, het terugsturen van een email of met een gescand gehandtekend document, maar het kan ook met een afzonderlijk toestemmingsscherm met daarop het verzoek tot (vrijwillig) akkoord door het aanvinken van een vakje waarnaast duidelijk omschreven is waarmee men akkoord gaat (Groep Gegevensbescherming, WP251rev01). Bovendien moeten de betrokken studenten of docenten er ook over geïnformeerd zijn dat ze hun toestemming te allen tijde kunnen intrekken. De betrokkenen moeten ook steeds van bij de aanvang informatie ontvangen over het bestaan van de geautomatiseerde besluitvorming en profilering, en van hun recht op bezwaar. Dit mag dus niet heimelijk gebeuren. Hierbij moet ook uitleg over de logica verstrekt worden, en over het belang en de verwachte gevolgen voor de betrokken studenten en docenten (art. 13.2.f) en art. 14.2.G; AVG). De betrokkenen beschikken over een recht van bezwaar (zie hierboven).

*b) Andere verplichtingen*

Bij de toepassing van geautomatiseerde profilering en besluitvorming mogen in principe geen 'gevoelige' gegevens gebruikt worden. Dat zijn bijzondere categorieën van gegevens, zoals deze waaruit ras of etnische afkomst, politieke of religieuze of levensbeschouwelijke overtuiging, of het lidmaatschap van een vakbond blijkt, evenals gegevens over de gezondheid, iemands seksuele gedrag of seksuele gerichtheid, maar ook biometrische gegevens voor unieke identificatie, tenzij er hiervoor eveneens een uitdrukkelijke toestemming is of op grond van een wettelijke bepaling in geval van een zwaarwegend algemeen belang. De reden is dat dergelijke gegevens gevoelig zijn en kunnen leiden tot discriminatie.

De instelling moet zich dus hoeden voor het gebruik van informatie waaruit bijvoorbeeld de etnische afkomst al dan niet onrechtstreeks tot uiting komt,

zowel bij het opstellen van individuele of groepsprofielen op grond van (grote hoeveelheden) data als bij het toepassen van de profielen**.** Gender staat niet in dit lijstje, maar het gebruik van genderinformatie bij LA mag evenmin leiden tot discriminatie. Als LA-belanghebbenden zich beroepen op een afwijking, zullen steeds 'passende waarborgen' genomen moeten worden ter bescherming van de rechten, vrijheden en gerechtvaardigde belangen van de betrokkenen. Dit houdt in dat elke betrokkene minstens nog steeds het recht moeten hebben (a) op een menselijke tussenkomst, (b) om het eigen standpunt kenbaar te maken en (c) om het besluit aan te vechten. Ook het opstellen van groepsprofielen en het gebruik daarvan vergen specifieke aandacht. Zo zal de instelling moeten nagaan of en in welke mate de gegevens geanonimiseerd kunnen worden. Het gebruik van (groeps)profilering zal ook steeds een effectbeoordeling vereisen (zie hieronder).

### 4. Effectbeoordeling en inschatting van de risico's

De AVG legt aan de verantwoordelijke van verwerkingen op om zelf een inschatting te maken van de effecten en risico's die de verwerkingen van persoonsgegevens met zich meebrengen (art. 35.3.a) AVG). Dat is het geval wanneer er systematisch uitgebreide beoordelingen van persoonlijke aspecten worden gemaakt, (i) gebaseerd op geautomatiseerde verwerkingen, zoals profilering, en (ii) waarop besluiten worden gebaseerd waaraan rechtsgevolgen zijn verbonden of die de persoon op vergelijkbare wijze wezenlijk treffen. Dit geldt eveneens voor alle andere verwerkingen die 'waarschijnlijk een hoog risico' inhouden, in het bijzonder wanneer nieuwe technologieën worden gebruikt (art. 25.1 AVG). LA kan hieronder vallen, zowel wanneer dit toegepast wordt op studenten als op docenten en studieloopbaanbegeleiders. De bedoeling van deze systematische beschrijving is een beoordeling te maken van de noodzaak en de evenredigheid van de verwerkingen om het vooropgestelde doel te realiseren. Indien LA bijvoorbeeld enkel en alleen tot doel heeft studenten de leerstof te laten inoefenen en ze te laten leren van hun fouten, is het mogelijk niet noodzakelijk om alle oefenresultaten op te slaan. Dit kan anders zijn indien LA zou worden aangewend om studenten hun leerprogressie te laten opvolgen, maar in dat geval is het bijvoorbeeld mogelijk niet noodzakelijk om de docent inzage te geven in de persoonlijke voortgang.

Hoewel LA gebaat is bij grote datasets, blijft een zo minimaal als nodige gegevens-verzameling en -verwerking voor de betrokkenen een algemeen beginsel. Verder moet het effect op de rechten en vrijheden van de betrokkenen, zoals het recht op niet-discriminatie, op gegevensbescherming of op privacy, ingeschat worden. Als de persoonsgegevens en profielen bijvoorbeeld niet op een veilige manier worden verwerkt, houdt dit een risico in op lekken. Hiervoor zal de verantwoordelijke in passende waarborgen moeten voorzien.

Als een beoordeling van het gegevensbeschermingseffect hoge restrisico's aan het licht brengt, is de verwerkingsverantwoordelijke verder verplicht om

de toezichthoudende autoriteit voorafgaand aan de verwerking te raadplegen (art. 36, lid 1 AVG). De publicatie van zo'n beoordeling is geen wettelijke AVG-verplichting. Wel wordt aangeraden te overwegen om ten minste onderdelen te publiceren, zoals een samenvatting of conclusie van de beoordeling. Verder kunnen gedragscodes omtrent LA die over instellingen heen ontwikkeld worden, bijdragen tot een correct gebruik van LA.

---

*Aanbeveling 14: Gedragscodes inzake LA worden het best ontwikkeld over de Vlaamse instellingen van hoger onderwijs heen. Ze zullen bijdragen tot een optimaal en correct gebruik van LA.*

---

## 5. Gebruik van LA voor statistiek of wetenschappelijk onderzoek

LA voor wetenschappelijk onderzoek, bijvoorbeeld voor het testen van algoritmes, waarbij geen profielen op het niveau van het individu worden aangemaakt of toegepast, valt onder enkele meer soepele bepalingen in verband met doelgebondenheid en opslag, en de uitoefening van de rechten van de betrokkenen. Toch zullen elke keer passende waarborgen getroffen moeten worden, waaronder technische en organisatorische maatregelen zoals het aggregeren, pseudonimiseren of anonimiseren van de gegevens. Doel is het risico voor individuen op traceerbaarheid te reduceren.

Voor het gebruik van LA voor louter statistische doeleinden op instellingsniveau of voor mededelingen aan de overheid, bijvoorbeeld in het kader van zogenoemde *Academic Analytics*, is de onderwijsinstelling in principe eveneens niet gebonden aan een ander en eerder bepaald doel, maar kan ze gegevens 'hergebruiken' voor dit bijkomende LA-doeleinde. Ook hier zullen passende waarborgen nodig zijn, gekoppeld aan technische en organisatorische maatregelen, waaronder in de eerste plaats het gebruik van geanonimiseerde gegevens. De nationale (Belgische) wetgeving ter vervollediging van de AVG, vastgelegd bij Wet van 30 juli 2018 betreffende de bescherming van natuurlijke personen met betrekking tot de verwerking van persoonsgegevens (*Belgisch Staatsblad* van 5 sept 2018), formuleert voor het onderzoeksgebruik bijkomende verduidelijkingen en eisen (zie Titel IV).

Vlaamse instellingen moeten zich ook hoeden voor hun (onbewuste) medewerking aan massale data-inzameling door buitenlandse (transatlantische) platformaanbieders, zonder afdoende bescherming van de persoonsgegevens.

## 6. Besluit

LA verwerkt in de meeste gevallen persoonsgegevens van studenten, studieloopbaanbegeleiders en docenten. Hierdoor is een gedetailleerd regelgevend

kader inzake gegevensbescherming van toepassing bij gebruik van LA in Vlaamse hogeronderwijsinstellingen. Hierboven werd de aandacht gevestigd op enkele verplichtingen, zonder exhaustiviteit na te streven. We steunden daarbij op de aangewezen bronnen.

In het algemeen kan men besluiten dat LA vooraf de doeleinden expliciet moet omschrijven en naleven, en dat er voor elke verwerking (of groep verwerkingen) steeds een wettelijke grond moet zijn. LA moet te allen tijde ook transparant zijn en bij belangrijke gevolgen voor betrokkenen is doorgaans vooraf een uitdrukkelijke toestemming vereist, die de betrokkene steeds kan intrekken. Ook bezwaar is mogelijk. Zelfs met een uitdrukkelijke toestemming moeten profilering en andere technieken voor geautomatiseerde besluitvorming dus met de grootste omzichtigheid worden toegepast, zeker als ze rechts- of andere gevolgen kunnen hebben die de persoon vergelijkbaar aanzienlijk treffen. De strengste bepalingen gelden dus voor LA bij profilering en (automatische) besluitvorming. Vanwege de Europese en nationale wettelijke voorwaarden voor LA-dataverwerking, die in dashboards en platformen ook ingebouwd moeten worden, is de overname van LA-toepassingen of -platformen die in het buitenland worden gebruikt, waaronder de Verenigde Staten, ook niet zonder meer mogelijk. LA in wetenschappelijk onderzoek en bij statistische verwerkingen op instellingsniveau is aan andere voorwaarden onderworpen, die nationaal verder werden uitgewerkt. Deze voorwaarden zijn ook belangrijk, maar ze zijn soepeler, om onderzoek niet uit te sluiten.

### B. Ethische aspecten

In de vorige hoofdstukken werd duidelijk dat er een inherente aanname is dat LA een positieve impact kan hebben binnen een instelling van hoger onderwijs, zowel voor de individuele student, studentenbegeleider of docent als voor de instelling in haar geheel. LA heeft immers het potentieel om het leer- en doceerproces beter te begrijpen. Op basis daarvan kunnen gerichte interventies een positieve impact hebben op het studiesucces en de studievoortgang. Maar naast de wettelijke aspecten die in het vorige deel zijn besproken, brengt LA ook ethische vraagstukken met zich mee (Sclater, 2015). Ideologische aannames en epistemologie kleuren de behandeling van de ethische vraagstukken (Slade & Prinsloo, 2013). Daarnaast spelen de onderwijsvisie maar ook de intrinsieke machtsverhoudingen die heersen binnen een instelling van hoger onderwijs een belangrijke rol.

---

*Aanbeveling 15: Elke instelling, die met LA aan de slag wil gaan, doet dat het best vanuit een beleidstekst voor LA, met inbegrip van de ethische aspecten en hoe die kaderen in de onderwijsvisie en het bestaande kwaliteitszorgsysteem. Het is ook aan te raden om in voldoende ruimte te voorzien voor bottom-upinitiatieven.*

---

## 1. Verzameling, gebruik en interpretatie van data

LA kan potentieel heel wat verschillende databronnen (III.A) hanteren, elk met hun eigen ethische uitdagingen. Het is dan ook moeilijk een eenduidige en omvattende set van richtlijnen op te stellen. Ondanks de veelheid aan databronnen zal het nooit mogelijk zijn elk leergedrag of relevante actie te meten, precies omdat ze niet allemaal digitaal zijn. De **reikwijdte en representativiteit** van de data zijn dus beperkt en de data zullen in het beste geval een beperkt zicht bieden op sommige aspecten van leergedrag. De **kwaliteit** van de data staat centraal: de resultaten van de data-analyse kan maar zo goed zijn als de gebruikte data. Daarnaast blijven bepaalde databronnen **buiten bereik** om ethische redenen, zoals het grootschalig verzamelen van de hersenactiviteit van studenten en docenten of de continue lokalisatie van studenten. **Ontbrekende data** kunnen correlaties introduceren tussen variabelen die slechts onrechtstreeks met elkaar verbonden zijn. Dit kan op zijn beurt leiden tot foute of misleidende interpretaties. Zelfs als de data volledig zijn, is er gevaar voor bevooroordeelde of zelfs foute interpretaties. Zo kunnen *swipe cards* (zie het voorbeeld in III.C) overgeïnterpreteerd worden als effectieve aanwezigheid in de lesmomenten. Een andere typische fout is het interpreteren van verbanden tussen variabelen (correlaties) als oorzaak en gevolg (causaties).

**Voorbeeld: correlatie versus causatie**

LA kan dit verband vaststellen: 'studenten die intensiever gebruik maken van het onlineleermateriaal behalen hogere scores op het examen'. Dit is louter een correlatie. De causale interpretatie 'intensiever gebruik van het onlineleermateriaal leidt tot een hogere score op het examen' geldt niet automatisch. Het is immers mogelijk dat de meer gemotiveerde studenten gemiddeld meer gebruik maken van het onlineleermateriaal en dat meer gemotiveerde studenten gemiddeld ook hogere scores halen op het examen. Deze gemotiveerde studenten zouden de hogere scores evengoed kunnen halen zonder intensief het onlineleermateriaal te gebruiken.

Bijgevolg is het ook mogelijk dat, indien minder gemotiveerde studenten op basis van de melding 'intensiever gebruik van het onlineleermateriaal leidt tot een hogere score op het examen' het onlinemateriaal meer intensief zullen gebruiken, ze geen hogere scores zullen halen.

Een fout die nog verder gaat is het foutief verbinden van intenties – gewild of ongewild – met metingen van gedrag.

Een ander gevaar van LA is **depersonificatie**: het niet langer zien van *datasubjects* (studenten of docenten) als individuen met hun persoonlijke factoren en omstandigheden. Bij depersonificatie worden ze herleid tot de beschikbare data rond hun achtergrond, leer- en doceerprestaties en opgemeten leer- of doceergedrag. In dit kader is er ook een oproep voor menselijke tussenpersonen (*human intermediation*): dashboards die gesprekken met studenten of docenten

**Voorbeeld: foutief verbinden van intenties aan metingen van gedrag**

Stel: er wordt vastgesteld dat een docent voor een bepaald opleidingsonderdeel gemiddeld 16 op 20 als score gaf. Een mogelijke interpretatie is dat deze docent prima heeft lesgegeven en de leerstof zo duidelijk heeft uitgelegd dat studenten in hoge mate de doelstellingen van het opleidingsonderdeel hebben bereikt. Andere mogelijke interpretaties zijn dat de docent bang is voor het toekennen van lage scores aan zijn studenten (om bijvoorbeeld beroepsprocedures te vermijden), dat hij geen tweede zittijd wil afnemen of dat hij goede studentenevaluaties wil 'afkopen' met hoge punten (zogenaamde *grading leniency*). In het eerste geval is de conclusie dat de docent zijn job prima heeft gedaan wat in de andere gevallen uiteraard niet zo is. De loutere analyse van de opgemeten data, zonder bijkomende informatie, volstaat niet om een conclusie over de intenties van de docent te trekken.

ondersteunen, eerder dan systemen die rechtstreeks de belanghebbenden aanspreken. Een verwante uitdaging is **oververeenvoudiging**. Statistische modellen en algoritmen analyseren vooral gemiddelde verbanden, eerder dan individuele gevallen. Een andere veelgebruikte vereenvoudiging is het opdelen van studenten of docenten in groepen, profielen en 'etiketten', de zogenaamde profilering. Eerst en vooral zijn er mogelijk groepen, profielen of etiketten die met de nodige voorzichtigheid gebruikt moeten worden. Beleidsmatig kan het immers nuttig zijn te weten of studenten met een bepaalde etnie, geslacht of handicap afwijkend leergedrag vertonen. Ethisch is het een heel andere kwestie om hier correct en gepast mee om te gaan. Het gebruik van gevoelige gegevens is ook wettelijk een grote uitdaging (zie V.A).

Dat laatste brengt ons bij een ander belangrijk aspect: de toegang tot ruwe en verwerkte data, zowel geanonimiseerd als niet-geanonimiseerd. Elke onderwijsinstelling is verplicht bij te houden wie tot welke gegevens toegang heeft. Maar de vraag rijst vooral in welke mate de bronnen (student of docent) nog eigenaar zijn van de data die over hen verzameld worden en welke mogelijkheden zij hebben om de data in te zien dan wel te wijzigen. Bijzondere aandacht moet bovendien gaan naar het risico van een dubbelrol van beoordelaar en begeleider. Van achtergrondinformatie over studenten die noodzakelijk is om hen te begeleiden kan niet worden verondersteld dat ze 'vergeten' is op het evaluatiemoment. Er bestaat dus een risico dat de achtergrondinformatie een ongewenste invloed heeft op de evaluatie.

### 2. Actie en interventies

LA-interventies die gebaseerd zijn op data van lage kwaliteit, bevooroordeelde of foute interpretaties, oververeenvoudiging in groepen enzovoort kunnen een negatieve impact hebben die gaat van inefficiëntie, ineffectiviteit tot demotivatie en zelfs afkeer. Interventies op maat van bepaalde groepen leiden automatisch tot ethische uitdagingen, omdat enerzijds de opdeling gebaseerd is op beperkte gegevens en dus imperfect zal zijn, omdat de populatie binnen een groep nog

divers kan zijn en omdat er andere informatie of aanbevelingen komen naar verschillende studentengroepen.

Het principe van *actionable feedback* ondersteunt ethische LA-interventies: feedback op basis waarvan de ontvanger actie kan ondernemen.

**Voorbeelden van *actionable* en *non-actionable feedback***

Het is bijvoorbeeld potentieel zinvol feedback te geven op het gebruik van het onlineleermateriaal, want studenten en docenten kunnen hun leer- of doceergedrag hieraan aanpassen. Hun geslacht of socio-economische achtergrond daarentegen kunnen studenten niet veranderen. De feedback hierop toespitsen (bijvoorbeeld 'jongens hebben 10% minder slaagkans dan meisjes') is dan ook niet zinvol. Dit sluit niet uit dat het wel opportuun kan zijn om de feedback aan te passen op basis van de kenmerken van een student, als aangetoond is dat dit noodzakelijk is om de feedback meer effectief te maken. Het blijft daarentegen niet zinvol om rechtstreeks feedback te geven op het geslacht of de socio-economische achtergrond an sich.

Veel interventies leiden studenten naar bijkomende ondersteuning of aangepaste leertrajecten. Zijn studenten of docenten echter 'verplicht' hierop in te tekenen of de raadgevingen te volgen, zelfs als die niet overeenkomen met hun eigen voorkeuren of doelstellingen? Voor de instelling zelf is er ook een ethisch dilemma rond de plicht van handelen*:* heeft een hogeronderwijsinstelling de plicht om in te grijpen als er bewijs is dat een student of docent geholpen zou zijn met een interventie? Hoe kan de instelling vervolgens een reeks van doelgerichte, effectieve en consistente interventies bepalen? Wat zijn de gevolgen als een onderwijsinstelling niet ingrijpt, hoewel LA dit had gesuggereerd?

Mogelijk kunnen op basis van de resultaten van LA ook interventies volgen die de belangen van de betrokkenen schaden, bijvoorbeeld de uitsluiting van studenten voor een herinschrijving in de opleiding als een voorspellend LA-model toont dat ze een lage slaagkans hebben en zich niet hebben ingezet in de opleiding.

Een van de kerndoelen van LA is zelfreflectie en zelfregulatie. Toch bestaat het gevaar dat LA net het omgekeerde gaat bewerkstelligen en de autonomie reduceert of zelfs leidt tot betutteling, dat het creatieve oplossingen en kritisch denken afleert in plaats van aanmoedigt en aanleert. Dit kan in de uitkomsten leiden tot het meer conformiteit in plaats van meer diversiteit.

Hoewel gedragsverandering een gewenste impact is van LA, is het de vraag of die gedragsverandering effectief leidt tot beter leren of onderwijzen. Enerzijds is al aangetoond dat personen hun gedrag veranderen zodra ze weten dat ze geobserveerd worden en anderzijds is het mogelijk dat ze niet langer deelnemen aan activiteiten of zich terugtrekken. Daarnaast rijst de vraag of het gedrag wel in positieve zin wordt bijgesteld, zodat het ook effectief leidt tot beter leren?

Of zullen mensen gedrag stellen dat vermijdt of net ervoor zorgt dat ze in een bepaalde doelgroep komen. Denk aan veelvuldig klikken op een leerplatform om niet langer als risicostudent geïdentificeerd te worden. Of omgekeerd: niet langer zelf naar een leerplatform gaan, om aanspraak te maken op de extra begeleiding voor risicostudenten.

Een menselijke interventie op basis van LA, en predictieve en prescriptieve analytics in het bijzonder (IV.D), roept bijkomende uitdagingen op. De betrokken personen zullen de nodige opleidingen en ondersteuning nodig hebben om de LA-toepassing op een correcte en gewenste manier in te zetten.

---

*Herhaling aanbeveling 6: Het gebruik van LA kan enkel succesvol zijn als de gebruikers de nodige competenties hebben. Stem daarom de LA-toepassingen af op het doelpubliek en voorzie in informatie, opleidingen en richtlijnen die zijn afgestemd op de gebruikersgroep. Enkel zo zullen de LA op een correcte en gewenste manier ingezet en gebruikt worden.*

---

### 3. Toestemming, beschikbaarheid en deïdentificatie van data

Toestemming, beschikbaarheid en deïdentificatie van data zijn nauw verbonden met de eerder behandelde privacy- en wettelijke aspecten van LA (V.A). Een extra vermelding waard is dat *learning dashboards* een opportuniteit kunnen bieden om de belanghebbenden een zinvolle inkijk te geven in de verzamelde data die de louter juridische beschikbaarheid van data overstijgt.

LA kan ook onopzettelijk gevoelige data creëren: op basis van onlinegedrag op het leerplatform kan je misschien wel de religie of etniciteit afleiden. Zelfs als de data geanonimiseerd zijn, is het soms toch mogelijk om op basis van die data de persoon 'af te bakenen' of te identificeren. Hoewel het niet eenvoudig is om deze onopzettelijke creatie of identificatie vooraf goed in te schatten, moet dit aan bod komen in een ethische risicoanalyse. Deze analyse houdt nauw verband met de wettelijk opgelegde beoordeling van het gegevensbeschermingseffect (0).

---

*Aanbeveling 16: Laat de bestaande ethische commissies een centrale rol spelen bij het adviseren over ethische aspecten van zowel LA-onderzoek als concrete onderwijskundige interventies.*

---

### 4. Afmelding

Een bijkomende uitdaging is het voorzien in de mogelijkheid om zich af te melden voor of terug te trekken uit dataverzameling, data-analyse en interventies (*opt-out* en *withdraw*). Het afmelden of terugtrekken kan echter potentieel een negatieve

impact hebben op het individu (bv. niet op de hoogte zijn van de aangeboden ondersteuning) en op de hele groep (incomplete en mogelijke bias op datasets, waardoor de nauwkeurigheid, correctheid en effectiviteit onder druk staat). Transparantie over de gevolgen is een ethische vereiste, maar bijkomend is het belangrijk na te gaan of er wel echt een keuze is om zich af te melden of terug te trekken, en of er geen druk is vanuit de instellingen om wel deel te nemen.

### 5. Transparantie en betrekken van belanghebbenden

Als een rode draad door de hele ethische (maar ook wettelijke) problematiek loopt transparantie. Transparantie over dataverzameling, dataopslag, data-analyse, datagebruik, interventies, afmelding enzovoort. Een minimale interpretatie van transparantie is de vraag naar expliciete toestemming. Verrijkender is de diverse belanghebbenden te betrekken in de verschillende fases van LA.

---

*Aanvulling aanbeveling 5: Betrek de verschillende belanghebbenden vanaf de eerste stap van de ontwikkeling bij Learning Analytics. Vergeet naast de onderwijskundigen, studenten en docenten ook de studieloopbaanbegeleiders niet. Zij brengen broodnodige expertise over de context, de huidige problematiek en de individuele en ethische studentenbegeleiding aan.*

*Een grote betrokkenheid van de belanghebbenden komt de relevantie van de ontworpen oplossing ten goede, ondersteunt de bewaking van het ethische karakter, creëert het nodige draagvlak en leidt uiteindelijk ook de acceptatie van het resultaat.*

---

# VI. Besluit en aanbevelingen

Dit KVAB-Standpunt heeft als ambitie een helikopterzicht aan te reiken op het domein van Learning Analytics (LA). Aan de hand van vier thema's – Waarom?, Data, Hoe? en Randvoorwaarden – verkende het een aantal opportuniteiten en gevaren van LA voor het Vlaams hoger onderwijs. Zo reikt het Standpunt een houvast aan voor diverse belanghebbenden: beleidsmakers, studenten, docenten, studieloopbaanbegeleiders en onderzoekers.

LA biedt heel wat potentieel voor Vlaamse instellingen van hoger onderwijs, in drie toepassingsdomeinen: onderzoek naar onderwijsleerprocessen (II.A), optimalisatie en personalisatie van de leeromgeving (II.B) en studieloopbaanbegeleiding (II.C). Dit potentieel mag ons echter niet verleiden tot een overhaaste toepassing. In het bijzonder de doelspecificatie en dus de waarom-vraag, die wordt beantwoord vanuit het onderwijskundig onderzoek, mag niet uit het oog worden verloren (Aanbeveling 2, blz. 25).

LA raakt aan veel domeinen, gaande van onderwijskunde tot *machine learning*. De effectieve realisatie van LA vergt dan ook een radicaal interdisciplinaire samenwerking van onderwijskundigen, cognitieve en motivatiepsychologen, specialisten in de concrete leerdomeinen, technologen, datawetenschappers, privacydeskundigen en ethici (Aanbeveling 2, blz. 25). Daarnaast legt LA de brug tussen onderzoek en praktijk. Daarom moeten zowel onderzoekers als de *practitioners* (docenten, studieloopbaanbegeleiders, studenten) van bij de eerste stap van de ontwikkeling van LA-initiatieven mee aan tafel zitten. De *practitioners* brengen broodnodige kennis van de context en de huidige uitdagingen ter tafel, die de relevantie van de ontworpen initiatieven ten goede komt. Daarnaast helpen zij het ethische karakter van de initiatieven te bewaken en kan hun vroege betrokkenheid het nodige draagvlak creëren en ook de uiteindelijke acceptatie van het resultaat ten goede komen (Aanbeveling 5, blz. 31, 38 en 74).

De inzet op multidisciplinaire samenwerking moet ook een kenmerk zijn van het LA-onderzoek dat een essentiële voedingsbodem is voor de onderzoeksgebaseerde toepassing van LA. Het onderzoek kan dankzij nieuwe technologieën leergedragingen in meer realistische leeromgevingen bestuderen. Onderzoek op basis van onder andere concrete gedragsdata zal bestaande inzichten kunnen versterken en ook nieuwe inzichten leveren op basis van de nieuw beschikbare data en technologieën, die mogelijk bestaande kennis in vraag zullen stellen. Voldoende financiering voor multidisciplinaire projecten is dan ook een noodzaak (Aanbeveling 1, blz. 22).

LA komt in essentie uit beschikbare data over leergedragingen en kan de belanghebbenden ondersteunen in het onderwijs- en leerproces. LA kan zorgen voor automatische aanbevelingen, oefentrajecten op maat, aanbevelingen voor vakken, waarschuwingen bij een lage activiteitsgraad enzovoort. Iedere

instelling moet er echter over waken dat LA ingezet wordt om de persoonlijke aanpak en inzichten binnen het onderwijs te versterken, en zeker niet om het menselijk contact en de ruimte voor persoonlijke omstandigheden te reduceren (Aanbeveling 4, blz. 31). Elke instelling die met LA aan de slag wil gaan dient een richtinggevende beleidstekst op te stellen, met aandacht voor de privacy, de wettelijke en de ethische aspecten van LA én de kadering van LA binnen de meer algemene onderwijsvisie (Aanbeveling 15, blz. 69). Zo'n beleidstekst moet een klimaat scheppen dat enerzijds richtinggevend is (top-down) maar anderzijds zorgt voor een kader en ruimte laat voor bottom-upinitiatieven (Aanbeveling 15, blz. 69). Hoewel bestaande beleidsteksten en ervaringen van andere onderwijsinstellingen relevante inspiratie bieden, is een aanpassing naar de eigen context, regelgeving en specifieke onderwijsvisie belangrijk. De specifieke context van de onderwijsinstelling heeft immers een grote invloed op veel aspecten van LA, zoals de noodzaak van beschikbare data (bv. *swipe cards* en meten van aanwezigheid), de onderwijscontext (bv. vrije toegang tot de opleiding of selectie), de doelstellingen van LA (bv. retentie of snelle heroriëntering), de nationale en instellingsspecifieke regelgeving en de onderwijs- en begeleidingspraktijken (bv. al dan niet een persoonlijke en professionele studieloopbaanbegeleider voor elke student). Voor elk bestaand LA-initiatief of voor al dan niet commerciële toepassingen moet de instelling nagaan of ze wel toepasbaar en transfereerbaar is naar de eigen context (Aanbeveling 3, blz. 26).

Een veel voorkomende misvatting is dat LA enkel toepasbaar is op *big data*. LA kan ook bouwen op *small data* die al beschikbaar zijn binnen de Vlaamse instellingen van hoger onderwijs. Denk aan data zoals academische resultaten, de studievoortgang en zelf-gerapporteerde gegevens in verband met leer- en studeervaardigheden. De instellingen zijn technologisch alvast voorbereid voor de *small data*. De opslag en verwerking van *big data* vergen bijkomende expertise, infrastructuur en dus investeringen (Aanbeveling 9, blz. 42). De hogeronderwijsinstelling heeft de verantwoordelijkheid te bewaken dat ze, zelfs bij gebruik van externe diensten, toch de controle over de data behoudt en dat die bijvoorbeeld ook beschikbaar blijven voor intern gebruik (Aanbeveling 7, blz. 39). De toegang en toegankelijkheid van de data voor LA moeten op zo'n manier georganiseerd worden dat de betrokkenen hierdoor vragen kunnen beantwoorden, maar dat ze ook nieuwe vragen ontdekken (Aanbeveling 9, blz. 42). Het gebruik van standaarden biedt opportuniteiten voor onder andere de efficiëntie van de dataverzameling, de toegang tot de data en de modulariteit van de LA-infrastructuur. Vlaamse instellingen van hoger onderwijs moeten daarom inzetten op een gedragen strategie rond het gebruik en de implementatie van internationale standaarden voor LA-data. De samenwerking van deze instellingen zal de positie versterken van elke individuele instelling bij de onderhandelingen over het gebruik van open standaarden met commerciële aanbieders, die weigerachtig kunnen staan tegenover de implementatie daarvan (Aanbeveling 8, blz. 42).

Daarnaast zal het gebruik van LA enkel succesvol zijn als de belanghebbenden over de nodige competenties beschikken. LA-initiatieven moeten daarom op hen afgestemd zijn en moeten ondersteund worden met de nodige informatie, opleidingen, richtlijnen (Aanbeveling 6, blz. 32 en 73) en gedragscodes (Aanbeveling 14, blz. 68).

Ethische en wettelijke aspecten en privacy spelen een grote rol bij LA. Op ethisch vlak moet een instelling niet alleen een beroep doen op haar richttinggevende beleidstekst (Aanbeveling 15, blz. 69), maar ook op de ervaring van de *practitioners* (Aanbeveling 5, blz. 31, 38 en 74) en de expertise van de bestaande ethische commissies (Aanbeveling 16, blz. 73). Op wettelijk vlak heeft de Europese Algemene Verordening Gegevensbescherming (AVG, in het Engels GDPR) een aanzienlijke impact op het domein van LA, aangezien LA dikwijls persoonsgegevens van betrokken studenten, studieloopbaanbegeleiders en docenten gebruikt. LA moet doelspecifiek en rechtmatig zijn, transparant en behoorlijk blijven en rechten van betrokkenen respecteren. Bij profilering en geautomatiseerde besluiten is doorgaans hun uitdrukkelijke toestemming vereist. Zelfs met deze toestemming moeten profilering en andere geautomatiseerde besluitvorming met de grootste omzichtigheid worden ingevoerd, zeker als ze rechts- of andere gevolgen kunnen hebben die de persoon aanzienlijk treffen. Deze impact zal zorgvuldig geanalyseerd en gedocumenteerd worden. Gezien de complexiteit van de wettelijke problematiek is de oprichting van een overlegorgaan voor de verantwoordelijken en de gegevensbeschermingsfunctionarissen van de Vlaamse instellingen van hoger onderwijs aan te bevelen. Dit overlegorgaan maakt de uitwisseling van ervaringen en best practices mogelijk (Aanbeveling 11, blz. 63).

Vlaamse instellingen van hoger onderwijs moeten op ethisch en wettelijk vlak de nodige ondersteuning bieden aan al hun personeelsleden – docenten, technologie-ontwikkelaars, onderzoekers, studentenbegeleiders – in de vorm van het concretiseren van de wettelijke en ethische kaders, gedragscodes en persoonlijk advies voor specifieke dossiers (Aanbeveling 13, blz. 65). Deze ondersteuning moet de implementatie van bescherming *by default* en de realisatie van privacy- en gegevensbescherming *by design* mogelijk maken (Aanbeveling 12, blz. 65).

LA, en bij uitbreiding educatieve technologie, belangt elke Vlaamse instelling van hoger onderwijs aan. De gemeenschappelijke context waarin deze instellingen opereren creëert gelijkaardige uitdagingen en opportuniteiten op het vlak van de inzet ervan. Een overlegorgaan in de schoot van de VLUHR, waar ervaringen uitgewisseld worden en een gezamenlijk Vlaams richtinggevend kader wordt uitgewerkt, is dan ook aan te bevelen (Aanbeveling 14, blz. 68). Internationale voorbeelden uit het Verenigd Koninkrijk (JISC) en Nederland (SURF) tonen aan dat nationale samenwerkingsorganisaties rond educatieve technologie een belangrijke steun zijn voor instellingen van hoger onderwijs bij de overweging en implementatie van LA. Ze geven niet alleen een impuls aan de concrete toepassing

van LA, maar ook aan het LA-onderzoek zelf. De oprichting van een Vlaamse of Belgische samenwerkingsorganisatie rond educatieve technologie of de aansluiting bij een bestaande organisatie strekt dan ook tot aanbeveling (Aanbeveling 10, blz. 55).

# VII. Terminologie

| | |
|---|---|
| Academic Analytics | *Academic Analytics* focust op beleidsuitdagingen en/of politieke en economische uitdagingen: hoe kunnen leeropportuniteiten en onderwijsresultaten verbeterd worden op opleidings- of instellingsniveau of zelfs op regionaal, nationaal en internationaal niveau? Het betreft vooral cijfers over studiesucces en uitval die opleidingen, instellingen en overheden gebruiken om hun effectiviteit en efficiëntie te meten en te verbeteren. |
| Actionable | *Actionable* is een Engels bijvoeglijk naamwoord dat aangeeft dat iets aanleiding geeft tot actie. |
| Anonimisering | Anonimisering is de procedure waarmee identificerende gegevens definitief verwijderd worden of zodanig verwerkt dat de gegevens niet langer herleid kunnen worden tot een geïdentificeerde of identificeerbare persoon. Anonimisering is per definitie in principe niet omkeerbaar. Hierin onderscheidt het zich van pseudonimisering, doordat de betrokkene niet of niet meer identificeerbaar is of waarbij het koppelen van informatie uit verschillende bronnen op de persoon niet mogelijk is. Zie ook overweging 26 AVG. |
| AVG | Algemene Verordening Gegevensbescherming (EU) 2016/679 (zie https://eur-lex.europa.eu/legal-content/NL/TXT/?uri=celex%3A32016R0679), in het Engels ook 'GDPR'. Europese regelgeving voor de bescherming van persoonsgegevens. Ook van toepassing op de meeste LA-data. Deze Europese regelgeving is sinds 25 mei 2018 van toepassing en vervangt de eerdere Richtlijn 1995/46. |
| Big data | *Big data* zijn datasets die voldoen aan deze kenmerken: groot volume, een diversiteit aan gegevenstypes en -bronnen en een hoge snelheid van datageneratie. |
| Blended Learning | *Blended Learning* verwijst naar de combinatie van verschillende didactische werkvormen, in het bijzonder naar de combinatie van onlinewerkvormen en traditioneel onderwijs. |

| | |
|---|---|
| Drop-out | Engelse term die in een onderwijscontext verwijst naar de uitval van studenten uit een opleiding, faculteit of instelling. |
| EDM | *Educational Data Mining* |
| Educational Data Mining | *Educational Data Mining* focust op de ontdekking van nieuwe patronen, verbanden en trends door de analyse van grote hoeveelheden data over leren. Deze patronen, verbanden en trends zouden met andere technieken moeilijk of onmogelijk te analyseren zijn. |
| GDPR | *General Data Protection Regulation* (EU) 2016/679). In het Nederlands de AVG. Zie AVG. |
| IRT | *Item Response Theory* |
| Item Response Theory | *Item response theory* (IRT) is een techniek uit de psychometrie en staat ook bekend als *latent trait theory*, *strong true score theory* of *modern mental test theory*. Het gaat om een paradigma voor het ontwerp, de analyse en het scoren van tests, vragenlijsten of andere instrumenten die vaardigheden, attitudes of andere kenmerken van personen opmeten. |
| JISC | JISC is een non-profitorganisatie in het Verenigd Koninkrijk. Ze biedt digitale diensten en oplossingen aan voor het hoger en voortgezet onderwijs (https://www.jisc.ac.uk/). |
| LA | *Learning Analytics* (zie hieronder) |
| Learning Analytics | *Learning Analytics* is het meten, verzamelen, analyseren en rapporteren van gegevens over studenten en hun context, met als doel het begrijpen en optimaliseren van het leren en de omgeving waarin dit gebeurt. |
| Learning Management System | Een *Learning Management System* (LMS) automatiseert de administratie en het beheer van onder andere cursussen en leermateriaal. Hoewel een LMS heel gelijkaardig is aan VLE, zijn er toch subtiele verschillen die ontstaan door hoe de technologische omgevingen worden gebruikt. Dit standpunt gebruikt de termen op een uitwisselbare manier. |

| | |
|---|---|
| Leerkrediet | Het leerkrediet is een instrument van het Vlaamse ministerie van Onderwijs en Vorming dat van toepassing is op het Vlaamse hoger onderwijs. De maatregel werd ingevoerd vanaf 2008-2009 om het studieduurverlenging en het onbeperkt (blijven) studeren en herkansen op kosten van de overheid tegen te gaan. |
| LMS | *Learning Management System* (zie hierboven) |
| LRS | *Learning Record Store*, *student record warehouse* voor het opslaan van leerdata uit verschillende bronnen van waaruit data ook gedeeld kunnen worden met of geëxporteerd naar externe rapporteringstools |
| Machinaal leren | Machineleren of in het Engels *machine learning* is een breed onderzoeksveld binnen de computerwetenschappen en kunstmatige intelligentie in het bijzonder. Het houdt zich bezig met de ontwikkeling van algoritmes en technieken waarmee computers kunnen leren |
| Massive Open Online Course | Een *Massive Open Online Course* (MOOC) is een onlinecursus waarvoor iedereen zich mag inschrijven. *Massive* verwijst naar het grote aantal deelnemers dat MOOCs dikwijls hebben. Hoewel MOOCs oorspronkelijk gratis waren, zijn betalende systemen of onderdelen (bijvoorbeeld betaling voor een certificaat) in opmars. |
| MOOC | *Massive Open Online Course* (zie hierboven) |
| Onderwijsleerproces | Het onderwijsleerproces is het proces waarbij studenten kennis en vaardigheden verwerven. |
| Open source | *Open source* beschrijft de praktijk die in de productie en de ontwikkeling vrije toegang geeft tot de bronmaterialen (de *source*) van het eindproduct. |
| Opleidingsonderdeel | Term gebruikt in het Vlaamse hoger onderwijs om naar een 'vak' te verwijzen. Een opleidingsonderdeel is een afgebakend geheel van onderwijsleer- en evalua-tieactiviteiten dat gericht is op het verwerven van doelstellingen wat betreft kennis, inzicht, vaardigheden en attitudes. |

| | |
|---|---|
| Profilering | De geautomatiseerde verwerking van (persoons) gegevens met als doel persoonlijke aspecten van een individu te evalueren. Het wordt gebruikt om o.a. beroepsprestaties, de economische situatie, gezondheid, persoonlijke voorkeuren, interesses, betrouwbaarheid, gedrag, locatie of verplaatsingen te analyseren of te voorspellen, en om eventueel te classificeren, bijvoor-beeld door persoonsgegevens te vergelijken en samen te brengen. Zie ook de definitie in art. 4.4 AVG. |
| Pseudonimisering | Pseudonimisering is de procedure waardoor persoonsgegevens niet meer aan een specifieke persoon kunnen worden gekoppeld zonder aanvullende gegevens, die apart worden bewaard. Identificerende gegevens kunnen bijvoorbeeld worden vervangen door een code of worden bewerkt met een versleutelingsalgoritme (het pseudoniem). Het algoritme kan eventueel voor elke persoon altijd hetzelfde pseudoniem construeren, waardoor informatie over de persoon uit verschillende bronnen kan worden gecombineerd. Zie ook overweging 26 AVG en de definitie in art. 4.5 AVG. |
| Psychometrie | Psychometrie is de wetenschap die zich bezighoudt met de technieken van het meten van psychologische fenomenen, zoals kennis, vaardigheden, attitudes, eigenschappen en persoonskenmerken. |
| Retentie | Letterlijke vertaling van de Engelse term *retention,* die in een onderwijscontext verwijst naar het behouden van studenten in een opleiding, faculteit of instelling. |
| SIS | *Student Information System* (zie hieronder) |
| Small data | Staat in tegenstelling tot *big data*. |
| Small Private Online Course | Een *Small Private Online Course* (SPOC) is een term die is afgeleid van de *Massive Open Online Course* (MOOC). Net als een MOOC is het een onlinecursus, maar een SPOC staat in tegenstelling tot een MOOC niet open voor iedereen. Een SPOC is afgeschermd voor bepaalde doelgroepen, is enkel op uitnodiging en/of heeft inschrijvingskosten. |
| SPOC | *Small Private Online Course* (zie hierboven) |

| | |
|---|---|
| Student Information System | Een *Student Information System* (SIS) verwijst naar een systeem op basis van administratieve software of een informatiemanagementsysteem voor het beheren van studentinformatie in een onderwijscontext. |
| Studieloopbaanbegeleider | Een studieloopbaanbegeleider is in het Vlaams hoger onderwijs een verzamelterm die verwijst naar de personeelsleden die instaan voor de begeleiding van studenten in hun studieloopbaan of studietraject. Het betreft onder andere studieadviseurs, monitoren en studietrajectbegeleiders. |
| SURF | SURF is de ICT-samenwerkingsorganisatie van het onderwijs en onderzoek in Nederland (https://www.surf.nl/) |
| Unobtrusive measures | Engelse term die in dit standpunt verwijst naar niet-invasieve metingen die dus het onderwijsleerproces niet verstoren. |
| Virtual Learning Environment | Een *Virtual Learning Environment* (VLE) is in de educatieve technologie een webgebaseerd platform voor de digitale aspecten van de studie, dat meestal gebruikt wordt binnen onderwijsinstellingen. Het platform geeft typisch virtuele toegang tot de vakken, leermateriaal, punten, tests en opdrachten, maar is dikwijls ook het portaal naar ander webgebaseerd materiaal in het kader van onderwijs. Voorbeelden zijn Toledo voor de KU Leuven en Minerva van de UGent of het opensource-alternatief Moodle. Hoewel een VLE heel gelijkaardig is aan LMS, zijn er toch subtiele verschillen die ontstaan door de wijze waarop de technologische omgevingen worden gebruikt. Dit standpunt gebruikt de termen op een uitwisselbare manier. |
| VLE | *Virtual Learning Environment* (zie hierboven) |

# VIII. Bronnen

## A. Wetgeving, Richtlijnen en Aanbevelingen

Verordening (EU) 2016/679 van het Europees Parlement en de Raad van 27 april 2016 betreffende de bescherming van natuurlijke personen in verband met de verwerking van persoonsgegevens en betreffende het vrije verkeer van die gegevens en tot intrekking van Richtlijn 95/46/EG (Algemene Verordening Gegevensbescherming), Publ, 4.5.2016, L 119/1-88, ook beschikbaar op https://eur-lex.europa.eu/legal-content/NL/TXT/HTML/?uri=CELEX:32016R0679&from=NL

Groep Gegevensbescherming Artikel 29, Richtsnoeren voor gegevensbeschermings-effectbeoordelingen en bepaling of een verwerking "waarschijnlijk een hoog risico inhoudt" in de zin van Verordening 2016/679, zoals laatst gewijzigd en vastgesteld op 4 oktober 2017, WP 248 rev.01, beschikbaar op http://ec.europa.eu/newsroom/just/item-detail.cfm?item_id=50083

Groep Gegevensbescherming Artikel 29, Richtlijnen over geautomatiseerde geïndividualiseerde besluitvorming en profilering voor de doeleinden van de Verordening 2016/679, zoals laatst gewijzigd en vastgesteld op 22 augustus 2018, WP251rev01, 34 p., beschikbaar op https://ec.europa.eu/newsroom/article29/item-detail.cfm?item_id=612053

CBPL, Aanbeveling 06/2017 betreffende het Register van de verwerkingsactiviteiten (artikel 30 van de AVG) (CO-AR-2017-011), 14 juni 2017, 36 p., beschikbaar op https://www.privacycommission.be/sites/privacycommission/files/documents/aanbeveling_06_2017_0.pdf

## B. Andere bronnen

Apereo foundation. (2017). Student Pulse: Visualizations for Learning Analytics Open Dashboard. Retrieved July 31, 2018, from https://www.apereo.org/content/student-pulse-visualizations-learning-analytics-open-dashboard

Arnold, K. E., & Pistilli, M. D. (2012). Course Signals at Purdue: Using Learning Analytics to Increase Student Success. In *Proceedings of the 2Nd International Conference on Learning Analytics and Knowledge* (pp. 267–270). New York, NY, USA: ACM. https://doi.org/10.1145/2330601.2330666

Barr, J., & Gunawardena, A. (2012). Classroom salon. In *Proceedings of the 43rd ACM technical symposium on Computer Science Education - SIGCSE '12* (p. 197). New York, New York, USA: ACM Press. https://doi.org/10.1145/2157136.2157196

Baylis, L. (2019). A Technical Look into Learning Analytics Data and Visualisations | Effective Learning Analytics. Retrieved August 1, 2018, from https://analytics.jiscinvolve.org/wp/2016/06/28/a-technical-look-into-learning-analytics-data-and-visualisations/

Broos, T., Peeters, L., Verbert, K., Van Soom, C., Langie, G., & De Laet, T. (2017). *Dashboard for actionable feedback on learning skills: Scalability and usefulness*. *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)* (Vol. 10296 LNCS). https://doi.org/10.1007/978-3-319-58515-4_18

Broos, T., Peeters, L., Verbert, K., Van Soom, C., Langie, G., & De Laet, T. (2017). Small Data as a Conversation Starter for Learning Analytics: Exam Results Dashboard for First-year Students in Higher Education. *Journal of Research in Innovative Teaching & Learning*, *minor revi*.

Broos, T., Verbert, K., Vansoom, C., Langie, G., & De Laet, T. (2017). Dashboard for Actionable Feedback on Learning Skills: How Learner Profile Affects Use. In *Springer Lecture Notes in Computer Science (LNCS) series. (Proceedings of the ECTEL 2017 conference; ARTEL workshop)* (p. to be published).

Card, S. K., Mackinlay, J. D., & Shneiderman, B. (1999). *Readings in information visualization : using vision to think*. Morgan Kaufmann Publishers. Retrieved from https://dl.acm.org/citation.cfm?id=300679

Charleer, S., Klerkx, J., Duval, E., De Laet, T., & Verbert, K. (2017). Towards balanced discussions in the classroom using ambient information visualisations. *International Journal of Technology Enhanced Learning*, *9*(2–3). https://doi.org/10.1504/IJTEL.2017.084501

Charleer, S., Vande Moere, A., Klerkx, J., Verbert, K., & De Laet, T. (2017). Learning Analytics Dashboards to Support Adviser-Student Dialogue. *IEEE Transactions on Learning Technologies*, 1–1. https://doi.org/10.1109/TLT.2017.2720670

Dawson, S. (2010). 'Seeing' the learning community: An exploration of the development of a resource for monitoring online student networking. *British Journal of Educational Technology*, *41*(5), 736–752. https://doi.org/10.1111/j.1467-8535.2009.00970.x

De Boeck, P., & Wilson, M. (2004). *Explanatory item response models : a generalized linear and nonlinear approach*. Springer.

De Laet, T., Broos, T., Langie, G., Pinxten, M., Van Soom, C., & Verbert, K. (2017). Nieuwe dashboards - Hoe je studenten kunt laten profiteren van hun eigen data. *Th&ma Hoger Onderwijs*, 6–9. Retrieved from http://www.themahogeronderwijs.org/archief/

Deeva, G., De Smedt, J., De Koninck, P., & De Weerdt, J. (2018). Dropout Prediction in MOOCs: A Comparison Between Process and Sequence Mining (pp. 243–255). Springer, Cham. https://doi.org/10.1007/978-3-319-74030-0_18

del Blanco, A., Serrano, A., Freire, M., Martinez-Ortiz, I., & Fernandez-Manjon, B. (2013). E-Learning standards and learning analytics. Can data collection be improved by using standard data models? In *2013 IEEE Global Engineering*

*Education Conference (EDUCON)* (pp. 1255–1261). IEEE. https://doi.org/10.1109/EduCon.2013.6530268

Duval, E. (2012). Learning Analytics and Educational Data Mining | Erik Duval's Weblog. Retrieved October 30, 2017, from https://erikduval.wordpress.com/2012/01/30/learning-analytics-and-educational-data-mining/

Elias, T. (2011). *Learning Analytics: Definitions, Processes and Potential*. Retrieved from http://www.learninganalytics.net/LearningAnalyticsDefinitionsProcessesPotential.pdf

Embretson, S. E., & Reise, S. P. (2000). *Item response theory for psychologists*. L. Erlbaum Associates.

Engelfriet, A., Manderveld, J., & Jeunink, E. (2017). *Learning Analytics onder de wet bescherming persoonsgeevens*. Retrieved from https://www.surf.nl/binaries/content/assets/surf/nl/kennisbank/2017/learning-analytics_wbp-editie2017_def.pdf

Ferguson, R. (2012). Learning analytics: drivers, developments and challenges. *International Journal of Technology Enhanced Learning*, *4*(5/6), 304–317. https://doi.org/10.1504/IJTEL.2012.051816

Few, S. (2006). *Information dashboard design : the effective visual communication of data*. O'Reilly.

Gašević, D., Kovanović, V., & Joksimović, S. (2017). Piecing the learning analytics puzzle: a consolidated model of a field of research and practice. *Learning: Research and Practice*, *3*(1), 63–78. https://doi.org/10.1080/23735082.2017.1286142

Gelan, A., Fastré, G., Verjans, M., Martin, N., Janssenswillen, G., Creemers, M., … Thomas, M. (2018). Affordances and limitations of learning analytics for computer-assisted language learning: a case study of the VITAL project. *Computer Assisted Language Learning*, *31*(3), 294–319. https://doi.org/10.1080/09588221.2017.1418382

Greller, W., & Drachsler, H. (2012). Translating Learning into Numbers: A Generic Framework for Learning Analytics. *Educational Technology & Society*, *15*(3), 42–57. Retrieved from http://www.ifets.info/journals/15_3/4.pdf

Johnson, L., Adams Becker, S., Cummins, M., Cummins, M., Estrada, V., Freeman, A., & Hall, C. (2016). *The NMC Horizon Report: 2016 Higher Education Edition*. Retrieved from https://www.learntechlib.org/p/171478/

Johnson, L., Adams Becker, S., Cummins, M., Freeman, A., Ifenthaler, D., & Vardaxis, N. (2013). *Technology Outlook for Australian Tertiary Education 2013-2018 An NMC Horizon Project Regional Analysis Time-to-Adoption Horizon: Two to Three Years*. Austin, Texas: The New Media Consortium. Retrieved from https://www.nmc.org/pdf/2013-Technology-Outlook-for-Australian-Tertiary-Education.pdf

Johnson, L., Adams, S., & Cummings, M. (2012). *The NMC Horizon Report: 2012 Higher Education Edition*. Austin, Texas: The New Media Consortium. Retrieved from https://www.nmc.org/pdf/2012-horizon-report-HE.pdf

Kadengye, D. T., Ceulemans, E., & Van den Noortgate, W. (2015). Modeling Growth in Electronic Learning Environments Using a Longitudinal Random Item Response Model. *The Journal of Experimental Education*, *83*(2), 175–202. https://doi.org/10.1080/00220973.2014.907226

Kaplan, D. (2009). *Structural equation modeling : foundations and extensions*. SAGE. Retrieved from https://uk.sagepub.com/en-gb/eur/structural-equation-modeling/book227519

Keim, D. A., Mansmann, F., Schneidewind, J., Thomas, J., & Ziegler, H. (2008). Visual Analytics: Scope and Challenges. In *Visual Data Mining* (pp. 76–90). Berlin, Heidelberg: Springer Berlin Heidelberg. https://doi.org/10.1007/978-3-540-71080-6_6

Kindt, E. (2017). Grondige analyse en beoordeling vereist. Gevolgen van de nieuwe Europese verordening voor leerplatformen. *Thema: Tijdschrift Voor Hoger Onderwijs En Management*, (3), 34–38. Retrieved from https://limo.libis.be/primo-explore/fulldisplay?docid=LIRIAS1711818&context=L&vid=Lirias&search_scope=Lirias&tab=default_tab&lang=en_US&fromSitemap=1

Klerkx, J., Verbert, K., & Duval, E. (2017). Learning Analytics Dashboards. In *Handbook of Learning Analytics* (1st ed., pp. 143–150). Society for Learning Analytics Research. https://doi.org/10.18608/hla17.012

Klinkenberg, S., Straatemeier, M., & van der Maas, H. L. J. (2011). Computer adaptive practice of Maths ability using a new item response model for on the fly ability and difficulty estimation. *Computers & Education*, *57*(2), 1813–1824. https://doi.org/10.1016/J.COMPEDU.2011.02.003

Leitner, P., Khalil, M., & Ebner, M. (2017). Learning Analytics in Higher Education—A Literature Review (pp. 1–23). Springer, Cham. https://doi.org/10.1007/978-3-319-52977-6_1

Long, P., & Siemens, G. (2011). Penetrating the Fog: Analytics in Learning and Education | EDUCAUSE. *EDUCAUSE Review*, *46*(5), 30. Retrieved from https://er.educause.edu/articles/2011/9/penetrating-the-fog-analytics-in-learning-and-education

Lust, G. (2012, December 21). *Opening the Black Box. Students' Tool-use within a Technology-Enhanced learning environment: An Ecological-Valid Approach*. Retrieved from https://lirias.kuleuven.be/handle/123456789/369830

Mazza, R., & Milani, C. (2005). Exploring Usage Analysis in Learning Systems: Gaining Insights From Visualisations. In: Workshop on Usage analysis in learning systems. In *12th International Conference on Artificial Intelligence in Education*

*(AIED* (pp. 65-72). Retrieved from http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.594.4040

Merceron, A., & Yacef, K. (2005). Educational Data Mining: a Case Study. In C.-K. Looi, G. I. McCalla, B. Bredeweg, & J. Breuker (Eds.), *Artificial Intelligence in Education - Supporting Learning through Intelligent and Socially Informed Technology, Proceedings of the 12th International Conference on Artificial Intelligence in Education, {AIED} 2005, July 18-22, 2005, Amsterdam, The Nether* (Vol. 125, pp. 467–474). {IOS} Press. Retrieved from http://www.booksonline.iospress.nl/Content/View.aspx?piid=1343

Mitchell, T. M. (Tom M. (1997). *Machine Learning*. McGraw-Hill.

Mohamad, S. K., & Tasir, Z. (2013). Educational Data Mining: A Review. *Procedia - Social and Behavioral Sciences*, *97*, 320–324. https://doi.org/10.1016/J.SBSPRO.2013.10.240

Raudenbush, S. W., & Bryk, A. S. (2002). *Hierarchical linear models : applications and data analysis methods*. Sage Publications. Retrieved from https://uk.sagepub.com/en-gb/eur/hierarchical-linear-models/book9230

Rogers, T., Dawson, S., & Gaevi, D. (2014). Learning Analytics and the Imperative for Theory-Driven Research. In *The SAGE Handbook of E-learning Research* (pp. 232–250). London: SAGE Publications Ltd. https://doi.org/10.4135/9781473955011.n12

Romero, C., Romero, J. R., & Ventura, S. (2014). A Survey on Pre-Processing Educational Data (pp. 29–64). Springer, Cham. https://doi.org/10.1007/978-3-319-02738-8_2

Romero, C., & Ventura, S. (2013). Data mining in education. *Wiley Interdisciplinary Reviews: Data Mining and Knowledge Discovery*, *3*(1), 12–27. https://doi.org/10.1002/widm.1075

Sclater, N. (2015). A taxonomy of ethical, legal and logistical issues of learning analytics v1.0 | Effective Learning Analytics. Retrieved January 5, 2018, from https://analytics.jiscinvolve.org/wp/2015/03/03/a-taxonomy-of-ethical-legal-and-logistical-issues-of-learning-analytics-v1-0/

Sclater, N. (2017). *Learning Analytics Explained*. Routledge. https://doi.org/10.1177/0022146512469014

Sedrakyan, G., De Weerdt, J., & Snoeck, M. (2016). Process-mining enabled feedback: "Tell me what I did wrong" vs. "tell me how to do it right." *Computers in Human Behavior*, *57*, 352–376. https://doi.org/10.1016/J.CHB.2015.12.040

Siemens, G. (2011). Learning and Knowledge Analytics. Retrieved August 27, 2018, from http://www.learninganalytics.net/uncategorized/learning-and-academic-analytics/

Slade, S., & Prinsloo, P. (2013). Learning Analytics. *American Behavioral Scientist*, *57*(10), 1510–1529. https://doi.org/10.1177/0002764213479366

Tabak, I. (2004). Synergy: A Complement to Emerging Patterns of Distributed Scaffolding. *Journal of the Learning Sciences*, *13*(3), 305–335. https://doi.org/10.1207/s15327809jls1303_3

The University of Edinburgh Information Services. (2017). *Learning Analytics Explained*. Routledge. https://doi.org/10.1177/0022146512469014

Van der Perre, G., & Van Campenhout, J. (2015). *Hoger onderwijs voor de digitale eeuw: KVAB-reflectiegroep Blended Learning* (Standpunten 2015, 34). Brussel: KVAB Press (Koninklijke Vlaamse Academie van België voor Wetenschappen en Kunsten). Retrieved from http://www.kvab.be/nl/standpunten/blended-learning

Van Laer, S., & Elen, J. (2018). Adults' Self-Regulatory Behaviour Profiles in Blended Learning Environments and Their Implications for Design. *Technology, Knowledge and Learning*, 1–31. https://doi.org/10.1007/s10758-017-9351-y

Vandewaetere, M., Desmet, P., & Clarebout, G. (2011). The contribution of learner characteristics in the development of computer-based adaptive learning environments. *Computers in Human Behavior*, *27*(1), 118–130. https://doi.org/10.1016/J.CHB.2010.07.038

Veenman, M. V. J. (2015). Metacognition. In *Handbook of Individual Differences in Reading*. Routledge. https://doi.org/10.4324/9780203075562.ch3

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. https://doi.org/10.1007/s00779-013-0751-2

Ware, C. (2012). *Information visualization : perception for design*. Morgan Kaufmann.

Wauters, K., Desmet, P., & Van den Noortgate, W. (2010). Adaptive item-based learning environments based on the item response theory: possibilities and challenges. *Journal of Computer Assisted Learning*, *26*(6), 549–562. https://doi.org/10.1111/j.1365-2729.2010.00368.x

## Samenstelling van de werkgroep

Jan Aerts (UHasselt, KU Leuven)

Dries Benoit (UGent)

Stefaan Cottenier (UGent)

Frederik Cornillie (KU Leuven, imec)

Tinne De Laet (KU Leuven)

Piet Desmet (KU Leuven, imec)

Olga De Troyer (VUB)

Jochen De Weerdt (KU Leuven)

Jan Elen (KU Leuven)

Anouk Gelan (UHasselt, EIT-InnoEnergy)

Katie Goeman (KU Leuven)

Els Kindt (KU Leuven, Universiteit Leiden)

Erik Mannens (UGent)

Dimitri Mortelmans (UAntwerpen)

Wim Van den Noortgate (KU Leuven, imec)

Joos Vandewalle (KU Leuven, KTW)

Jan Velghe (UGent)

Katrien Verbert (KU Leuven)

Pieter Vos (Vlaams Ministerie van Onderwijs en Vorming)

KTW = Klasse van de Technische Wetenschappen

## **RECENTE STANDPUNTEN** (vanaf 2015)

35. Hugo Hens e.a. – *Energiezuinig (ver)bouwen: geen rechttoe rechtaan verhaal*, KVAB/Klasse Technische wetenschappen, 2015.

36. Marnix Van Damme – *Financiële vorming*, KVAB/Klasse Menswetenschappen, 2015.

37. Els Witte – *Het debat rond de federale culturele en wetenschappelijke instellingen (2010-2015)*, KVAB/Klasse Menswetenschappen, 2015.

38. Irina Veretennicoff, Joos Vandewalle e.a. – *De STEM-leerkracht*, KVAB/Klasse Natuurwetenschappen en Klasse Technische wetenschappen, 2015.

39. Johan Martens e.a. – *De chemische weg naar een $CO_2$-neutrale wereld*, KVAB/Klasse Natuurwetenschappen, 2015.

40. Herman De Dijn, Irina Veretennicoff, Dominique Willems e.a. – *Het professoraat anno 2016*, KVAB/Klasse Natuurwetenschappen, Klasse Menswetenschappen, Klasse Kunsten en Klasse Technische wetenschappen, 2016.

41. Anne-Mie Van Kerckhoven, Francis Strauven – *Een bloementapijt voor Antwerpen*, KVAB/Klasse Kunsten, 2016.

42. Erik Mathijs, Willy Verstraete (e.a.), *Vlaanderen wijs met water: waterbeleid in transitie*, KVAB/Klasse Technische wetenschappen, 2016.

43. Erik Schokkaert - *De gezondheidszorg in evolutie: uitdagingen en keuzes*, KVAB/Klasse Menswetenschappen, 2016.

44. Ronnie Belmans, Pieter Vingerhoets, Ivo Van Vaerenbergh e.a. – *De eindgebruiker centraal in de energietransitie*, KVAB/Klasse Technische Wetenschappen, 2016.

45. Willem Elias, Tom De Mette – *Doctoraat in de kunsten*, KVAB/Klasse Kunsten, 2016.

46. Hendrik Van Brussel, Joris De Schutter e.a., *Naar een inclusieve robotsamenleving*, KVAB/Klasse Technische Wetenschappen, 2016.

47. Bart Verschaffel, Marc Ruyters e.a., *Elementen van een duurzaam kunstenbeleid*, KVAB/Klasse Kunsten, 2016.

48. Pascal Verdonck, Marc Van Hulle (e.a.) - *Datawetenschappen en gezondheidszorg*, KVAB/Klasse Technische wetenschappen, 2017.

49. Yolande Berbers, Mireille Hildebrandt, Joos Vandewalle (e.a.) - *Privacy in tijden van internet, sociale netwerken en big data*, KVAB/Klasse Technische wetenschappen, 2017.

50. Barbara Baert (e.a.), *Iconologie of 'La science sans nom'*, KVAB/Klasse Kunsten, 2017.

51. Tariq Modood, Frank Bovenkerk – *Multiculturalism. How can Society deal with it?* KVAB/Klasse Menswetenschappen, 2017.

52. Mark Eyskens – *Europa in de problemen.* KVAB/Klasse Menswetenschappen, 2017.

53. Luc Steels – *Artificiële intelligentie. Naar een vierde industriële revolutie?.* KVAB/Klasse Natuurwetenschappen, 2017.

54. Godelieve Gheysen, René Custers, Dominique Van Der Straeten, Dirk Inzé*, Ggo's anno 2018. Tijd voor een grondige herziening*. KVAB/Klasse Natuurwetenschappen, 2017.

55. Christoffel Waelkens (e.a.) – *Deelname van Vlaanderen aan grote internationale onderzoeksinfrastructuren: uitdagingen en aanbevelingen*, KVAB/Klasse Natuurwetenschappen, 2017.

55. Addendum. Jean-Pierre Henriet. – *Mijlpalen in internationale wetenschappelijke samenwerking*, KVAB/Klassen Natuurwetenschappen, 2017.

56. Piet Swerts, Piet Chielens, Lucien Posman – *A Symphony of Trees. Wereldcreatie naar aanleiding van de herdenking van de Derde Slag bij Ieper, 1917*, KVAB/Klasse Kunsten, 2017.

57. Willy Van Overschée e.a. – *De mobiliteit van morgen: zijn we klaar voor een paradigmawissel*?, KVAB/Klasse Technische Wetenschappen, 2018.

De volledige lijst met standpunten en alle pdf's kunnen worden geraadpleegd op
www.kvab.be/standpunten



De populariteit van Learning Analytics, dat gegevens over studenten en hun omgeving wil gebruiken om het onderwijs te verbeteren, is aan een steile opmars bezig. Dankzij de groeiende digitalisering komen er meer en meer gegevens beschikbaar van de activiteiten van studenten, docenten en studieloopbaan-begeleiders. Maar kan Learning Analytics wel de hooggespannen verwachtingen inlossen? Zal het persoonlijke feedback en begeleiding faciliteren? Zal het de drop-out verminderen en studieprestaties verhogen?

Dit KVAB Standpunt legt uit wat Learning Analytics is, hoe het kan ingezet worden binnen het Vlaams hoger onderwijs, welke technieken en technologieën aan de basis liggen, maar ook welke uitdagingen er zijn op vlak van ethiek en privacy. Het standpunt illustreert alle aspecten met concrete voorbeelden uit binnen- en buitenland. Zestien aanbevelingen begeleiden de beleidsmakers bij de efficiënte en effectieve implementatie van Learning Analytics in het Vlaams hoger onderwijs.

De reeks Standpunten van de Academie is een bijdrage tot het wetenschappelijk onderbouwd debat over actuele maatschappelijke en artistieke thema's. De auteurs, leden en werkgroepen van de Academie schrijven in eigen naam, onafhankelijk en met volledige intellectuele vrijheid. De goedkeuring voor publicatie door een of meerdere Klassen van de Academie waarborgt de kwaliteit van de gepubliceerde studies.

Vlaamse overheid

RICHTING MORGEN