



# Versnellingsplan Onderwijsinnovatie met ICT

veilig en betrouwbaar  
benutten van studiedata

## Simulatiedataset LMS data

### Waarom een simulatiedataset met LMS data?

Onderwijsinstellingen gebruiken leermanagementsystemen (LMS) zoals Canvas, Blackboard en Brightspace om onderwijs aan te bieden. Deze systemen verzamelen grote hoeveelheden studiedata die gebruikt kunnen worden om het onderwijs te verbeteren. Echter, Nederlandse onderwijsinstellingen hebben veelal slechts beperkt toegang tot data in hun LMS. Dit kan om technische redenen zijn maar bijvoorbeeld ook omdat instellingen (nog) geen beleid hebben met betrekking tot het verwerken van deze data. Door het hebben van fictieve maar realistische data kunnen onderwijsinstellingen alvast ervaring opdoen zonder dat zij eerst data uit de eigen systemen hoeven te ontsluiten. Een simulatiedataset kan hier uitkomst bieden. Het voordeel van simulatie ten opzichte van het anonimiseren of pseudonimiseren van data is dat bij simulatie zowel de bruikbaarheid als de privacy beiden gewaarborgd blijven. Een verkenning naar de behoeftes en randvoorwaarden van een simulatiedataset met LMS data is nu afgerond.

### Waarvoor willen instellingen de simulatiedataset met LMS data gebruiken?

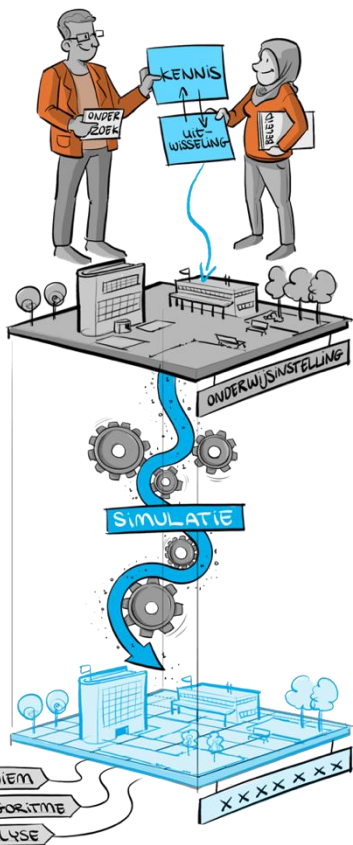
Instellingen willen de simulatiedataset voor verschillende activiteiten inzetten, bijvoorbeeld het creëren van awareness, gebruikers alvast laten starten met analyses van studiedata, het testen en vergelijken van algoritmes en IT-infrastructuur, gebruik binnen het eigen onderwijs, analyseren en verbeteren van cursusontwerpen of het bepalen van de effectiviteit van data-ondersteunde interventies. We kunnen hierin verschillende niveaus onderscheiden. Waar een beperkte, kleine dataset al genoeg is om gebruikers de mogelijkheden van studiedata te laten ervaren, is voor complexere analyses en toepassingen een diepgaandere, uitgebreidere dataset nodig.

### Waarom moet de simulatiedataset met LMS data voldoen?

De analyse van alleen LMS data leidt vaak tot weinig inzichten. Door LMS data te combineren met data over studenten en cursussen kunnen betere analyses worden gedaan. De simulatiedataset moet dus meer dan alleen LMS data bevatten. Tevens moet de simulatiedataset de werkelijkheid weerspiegelen. Dit betekent dat de karakteristieken, onderlinge correlaties en patronen in de dataset moeten worden gebaseerd op data van echte studenten die studeren aan een echte instelling. De simulatiedataset met LMS data wordt daarmee een synthetische dataset die is gebaseerd op reële studentdata. De simulatie moet zo uitgevoerd worden dat deze data op geen enkele manier meer terug te herleiden is tot de oorspronkelijke gegevens. Tijdens het bouwen van de simulatiedataset voor universiteiten met instellingsdata heeft de zone hier ervaring in opgedaan.

### Wat is de volgende stap?

Momenteel onderzoeken we welke onderwijsinstelling als leverancier voor de LMS data kan optreden. Daarna kunnen we de data ontsluiten en opschonen. Hierbij betrekken we domeinexperts zoals docenten en opleidingscoördinatoren. Vervolgens schrijven we de benodigde scripts en kan een daadwerkelijke simulatiedataset worden gegenereerd.



Waar kan ik meer  
informatie vinden?

[Doe-meer-met-  
studiedata.nl](https://doe-meer-met-studiedata.nl)

Hier vind je de  
simulatiedataset voor  
universiteiten met  
instellingsdata.