

Hoe WUR een nieuw vakgebied integreert in het onderwijs

# Operatie data science

De toekomst draait om data. Veel universiteiten ontwikkelen daarom aparte masteropleidingen datawetenschappen. WUR kiest een andere weg. Studenten kunnen zich hier binnen de bestaande studies verdiepen in bigdataverzameling en -analyse. 'Data science in context. Dat past bij Wageningen.'

*tekst* Luuk Zegers *foto* Marte Hofsteenge *illustratie* Pascal Tieman



**H**ightech innovaties, robots en kunstmatige intelligentie gaan een steeds grotere speler spelen in voedselsystemen, landbouw, gezondheid en milieu. Om studenten hier beter op voor te bereiden, is Wageningen University begonnen aan een forse uitbreiding van het onderwijs in datawetenschappen. Zo is er sinds vorig collegejaar een data-science minor, zijn er tien nieuwe data-science vakken ontwikkeld en bieden steeds meer opleidingen een mastertrack data science aan.

#### **Nieuwe vaardigheden**

Datawetenschap? Maar alle wetenschap maakt toch gebruik van data? Klopt, zegt Dean of Education Arnold Bregt. 'Data zijn zo oud als de wetenschap. Je hebt een vraag, je verzamelt data, die analyseer je en dan kom je tot een conclusie. Maar dankzij toenemende digitalisering is het makkelijker geworden

om steeds grotere hoeveelheden data te verzamelen en te analyseren. Zo worden datasets steeds omvangrijker en ingewikkelder. Dat biedt prachtige nieuwe onderzoeksmogelijkheden, maar betekent ook dat er nieuwe vaardigheden nodig zijn om dat onderzoek op de juiste manier uit te voeren.'

Data worden op steeds grotere schaal gebruikt om processen te optimaliseren, voorspellingen te doen, objecten te herkennen in foto's met hulp van kunstmatige intelligentie (zie kader) en meer. Zo ontwikkelden Arjen Daane en zijn collega's van Wageningen Economic Research de Global Detector. Daarin worden verschillende databronnen en algoritmes gecombineerd om vragen te beantwoorden zoals: welke locatie in de wereld is het meest geschikt voor tomatenteelt of wat is de vraag naar melk in Azië in 2022.

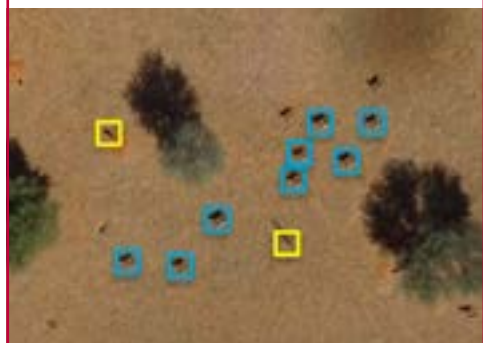
#### **Digital twins**

Een ander voorbeeld van datawetenschap

## Dieren tellen met data

Hoe voer je een dierentelling uit in een Afrikaans nationaal park van honderden vierkante kilometer? 'Vroeger vloog je met een team in een helikopter over het park en ging je met z'n allen tellen', zegt Lukasz Grus, onderwijscoördinator data science bij WUR. 'Voordeel: gaaf om te doen. Nadeel: duur en niet heel accuraat. Devis Tuia en Benjamin Kellenberger van Geo-information Science and Remote Sensing hebben een manier bedacht die accurater en goedkoper is. Een drone vliegt over het land en schiet ruim honderdvijftig luchtfoto's per vierkante kilometer. Vervolgens gebruiken ze een vorm van kunstmatige intelligentie die bekendstaat als *deep learning*. Simpel gezegd leerden ze een computerprogramma om dieren te herkennen op luchtfoto's.'

Het computerprogramma kreeg daarbij hulp van de mens. Honderden vrijwilligers identificeerden eerst dieren op de luchtfoto's. Die foto's 'voerden' de wetenschappers aan het programma om het te laten zien welke pixels dieren zijn en welke bijvoorbeeld rotsblokken of struiken. Het programma analyseerde die input en ging vervolgens zelf dieren identificeren. De output werd daarna weer gecontroleerd door mensen. Als het programma een struikje voor een dier had aangezien, kreeg het één strafpunt. Als het programma dieren had gemist op een foto, kreeg het tachtig strafpunten. Zo leerde het algoritme om dieren en struiken van elkaar te onderscheiden, maar vooral om geen dieren te missen. 'Nu zegt het programma bijvoorbeeld: deze pixels zijn 90 procent zeker een gnoe of die pixels daar zijn 80 procent zeker een antilope', zegt Grus. Met behulp van het programma kan één mens nu in een week tijd alle dieren in een park tellen.



## 'Bij reguliere wetenschap is data het medium, bij datawetenschap bestudeer je dit medium'

Lukasz Grus, onderwijscoördinator data science

schap waarin WUR investeert is *digital twins*: het maken van een digitale kopie van een systeem of organisme, zoals een koe. Moderne koeienstallen meten precies hoeveel een koe eet, hoeveel melk ze geeft, de temperatuur in de stal, enzovoort. Door met die data een digitale kopie te maken van de koe, kun je proberen de realiteit beter te begrijpen én modellen maken om toekomstvoorspellingen mee te doen, bijvoorbeeld over de invloed van een dieetverandering op de melkproductie. Wetenschappers kunnen *digital twins* maken van cellen, planten, dieren, mensen en ecosystemen, maar ook van abstractere zaken zoals voedselketens.

'Kort gezegd is data science het halen van kennis uit data', zegt Lukasz Grus, onderwijscoördinator data science bij WUR. 'Een datawetenschapper onderzoekt de verschillende methodes om data te verzamelen, te analyseren, te verwerken en te visualiseren. Bij reguliere wetenschap is data het medium waarmee je probeert een onderzoeksvraag te beantwoorden. Bij datawetenschap bestudeer je dit medium, verken je wat je uit data kunt halen en probeer je vraagstukken op te lossen door patronen in data te ontdekken.'

### De rol van Wageningen

In zijn rol als onderwijscoördinator maakte Grus vorig jaar een inventarisatie van wat Wageningen University al aanbood op het gebied van datawetenschappen. 'Ik zag dat we al best veel deden, maar allemaal toegepast binnen onze domeinen. We hebben hier geen aparte data-science master, maar we leren studenten toegepaste data science.' Hij wijst onder andere naar de opleidingen Geo-Information Science, Biosystems Engineering en Bioinformatics. 'Dat zijn een soort data-science masters binnen bepaalde domeinen.'

Die benadering past bij Wageningen, zegt Bregt. 'De masters data science schieten nu overall in Nederland bij wijze van spreken uit de grond, maar wij laten het opleiden van de hardcore datawetenschappers en het ontwikkelen

van nieuwe algoritmes graag over aan universiteiten als Delft en Eindhoven. Door datawetenschappers op te leiden binnen onze domeinen vervult Wageningen een brugfunctie. Data science in context. Want je kunt wel ontzettend veel data over elkaar heen leggen en iets uitrekenen, maar je moet wél kunnen beoordelen of de verbanden die je vindt ook hout snijden. Ik denk dat dat een heel mooie rol is, waarmee je ook een meerwaarde hebt in de markt.'

### Nieuwe vakken

Om die rol op een goede manier te vervullen, moest het aanbod van onderwijs in data science worden uitgebreid en verbreed. Bij data science komen diverse skills kijken, vertelt Bregt. 'Je moet iets weten van statistiek, kunstmatige intelligentie, dataverzameling. Maar ook visualisatie is belangrijk: hoe laat je de resultaten zien? Verder is ethiek een belangrijk onderdeel. Op het gebied van privacy bijvoorbeeld, maar ook bij de keuze van je werkwijze. Bepaalde methodes dwingen bepaalde antwoorden af. Daar moet je over nadenken.'

Diverse skills dus, die vragen om een divers aanbod van vakken. Een deel van die vakken bestond al, een deel moest nog worden ontwikkeld. Bregt: 'In een jaar tijd hebben we tien nieuwe vakken ontwikkeld die dit jaar al worden gegeven. Technische en methodische vakken over hoe je data moet analyseren, een aantal advanced vakken over data science in context, zoals *Data Science for Health*, *Data Science for Ecology* en *Smart Environment*, en een vak over data en ethiek.' Masteropleidingen kunnen een combinatie van die vakken aanbieden als mastertrack. Bregt: 'Studenten die voor zo'n track kiezen, worden datawetenschapper binnen hun domein. Zo leiden we studenten op met dataskills plus de domeinkennis van hun master, die ook nog eens nadenken over hun ethische verantwoordelijkheden als wetenschappers.'

Als datawetenschap zo belangrijk is, moet dat dan niet een verplicht onderdeel worden van

iedere opleiding? Bregt: 'De kerntaak van Wageningen University is studenten leren om goed onderzoek te doen binnen een discipline.'

#### Essentiële vaardigheid

Goed met data kunnen omgaan is een essentiële vaardigheid voor goed onderzoek; iedereen die in Wageningen afstudeert, moet dat kunnen. Dat wil niet zeggen dat elke student een datawetenschapper moet worden.

Maar als je dat wél wilt worden, is het onze taak als universiteit om dat te faciliteren.' Met tien nieuwe vakken en negen mastertracks is daarin nu een flinke stap gezet, aldus Bregt. 'De bouwstenen hebben we. Dus nu kunnen we gaan verbreden en kijken hoe we voor de opleidingen die nog geen tracks datawetenschappen aanbieden, ook mastertracks kunnen ontwikkelen. Die optimalisatieslag begint nu.' **B**



**'Je moet wél kunnen beoordelen of de verbanden die je vindt in de data ook hout snijden'**

Arnold Bregt, Dean of Education

## Domeinexpert met dataskills

Maartje Holtslag (25) deed de master Geo-Information Science aan WUR. 'Dat is eigenlijk een data-sciencemaster die focust op ruimtelijke data.' Kortom: een opleiding tot domeinexpert met dataskills.

Tijdens haar masterstage bij het Environmental Systems Research Institute (Esri) voerde Holtslag een opdracht uit voor de blindenstichting. 'Zij willen routes uitzetten voor blinde mensen, dus moeten ze weten waar zebrapaden liggen. Het klinkt misschien gek, maar daar is nergens in Nederland een overzicht van. Daarom heb ik een computermodel getraind om zebrapaden automatisch te detecteren in luchtfoto's.' Daarvoor moest Holtslag eerst trainingsdata maken. 'Ik tekende vlakken om zebrapaden op foto's, die het model kon gebruiken als voorbeeld. Vervolgens heb ik nieuwe foto's in het model gestopt om te kijken of het daarop ook zebrapaden herkent.'

Nadat ze haar diploma had gehaald, kon Holtslag aan de slag bij Esri als programmeur. 'Ik

zoek bijvoorbeeld uit op welke daken zonnepanelen liggen. Of ik gebruik data uit camera's op kruispunten om te kijken hoeveel auto's en fietsers er langskomen. Zo kun je uitzoeken of het fietsgedeelte

veilig is of hoe stoplichten beter op elkaar kunnen worden afgestemd. Als programmeur focus je op de technische kant. Mijn werk draait dus niet om het analyseren van de resultaten, maar meer om de weg naar de resultaten toe.' Holtslag vindt WUR's strategie – het opleiden van domeinexperts met dataskills – een goede keuze. 'Met data science kun je op veel plekken terecht, maar vaak wél binnen een specifiek domein. Dus je wilt zowel het werkveld kennen als met data kunnen omgaan.'



**'Ik heb een computermodel getraind om zebrapaden te detecteren in luchtfoto's'**

Maartje Holtslag, alumna Geo-Information Science

## Interesse in data science? Dit zijn de mogelijkheden

**Bachelor:** In je vrijekeuzeruimte kun je vakken data science volgen, zoals Data Management, Programming in Python, Big Data en Biological Data Analysis and Visualization. Ook kun je een minor data science doen waarin meerdere data-sciencevakken worden gecombineerd.

**Master:** De masteropleidingen Geo-Information Science, Biosystems Engineering en Bioinformatics draaien zó veel om data dat het in feite domeingebonden data-sciencemasters zijn. Zes andere masteropleidingen bieden mastertracks data science of clusters van data-sciencevakken aan: Biology, Earth and Environment, Forest and Nature Conservation, Nutrition and Health, Plant Biotechnology en Plant Sciences. Binnen elke master kun je in je vrijekeuzeruimte ook vakken data science volgen.

**Een compleet overzicht van de vakken data science vind je op [www.wur.eu/data/courses](http://www.wur.eu/data/courses)**