

# PERFECTE AARDAPPELS KWEKEN MET BIG DATA

- **Laxmi Parida van IBM Research sprak op Dies**
- **Gezamenlijk plan voor computerprogramma Watson Potato**

De universiteit was weer jarig op 9 maart. De Diesrede werd dit jaar verzorgd door Laxmi Parida, hoofd van de Computational Genomics Group van IBM Research. Zij heeft verteld hoe je met het slim combineren van veel databestanden kanker effectiever kunt behandelen of betere aardappels kunt telen. Wageningen UR wil samenwerken met haar instituut op het gebied van precisielandbouw en -veredeling.

Laxmi Parida werkt bij IBM Research in de VS aan het combineren van veel informatie om specifieke vragen te beantwoorden. Dat gebeurt met het computerprogramma Watson. Het onderzoeksinstituut heeft bijvoorbeeld IBM Cancer ontwikkeld. Samen met vijf academische ziekenhuizen in de VS heeft het instituut veel informatie over de ontwikkeling van kanker verzameld van de ziekenhuizen, uit de wetenschappelijke literatuur en de sociale media. Hiermee kan IBM een analyse van hun patiënten invoeren in het programma, die vervolgens een lijst geeft van de meest waarschijnlijke aandoeningen van de patiënt. Dat helpt de artsen om sneller een goede diagnose te stellen en een adequaat behandelplan te maken.

Richard Visser, hoofd van Wageningen UR Plant Breeding, is afgelopen jaar met zijn Wageningse collega's Thomas Been en Richard Finkers op bezoek geweest bij IBM Research. Wageningen UR gaat samen met het Amerikaanse instituut Watson Potato ontwikkelen, een computerprogramma met alle wetenschappelijke en praktijkinformatie over aardappels. Vis-

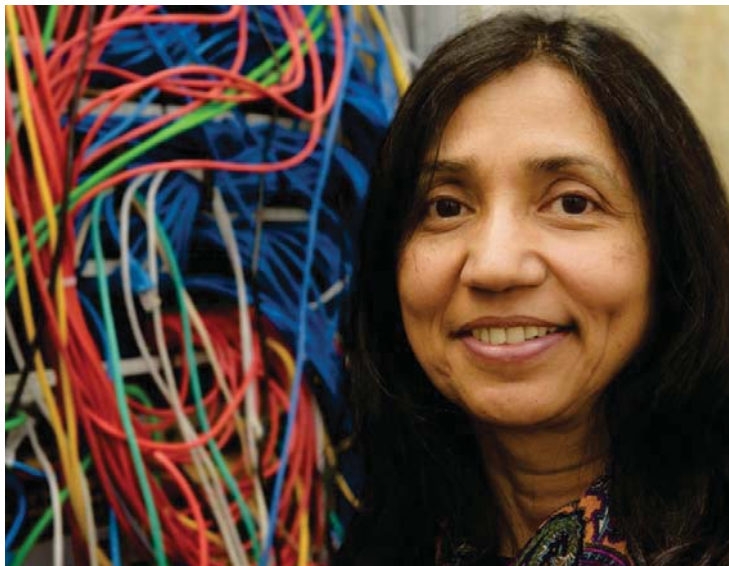


FOTO: GUY ACHERMAN

ser: 'Wij hebben heel veel informatie en datasets over aardappels en zij weten hoe je de databases moet koppelen.'

Het gebruik van grote hoeveelheden data vereist een diepgaande kennis van databeheer en -integratie, zegt Parida desgevraagd. 'Veel data zijn onzichtbaar voor de computer, dus moet je computers bouwen die functioneren als het menselijk brein, zodat je deze ongestructureerde data kunt begrijpen en interpreteren. Dit heet *cognitive computing*. Dit vereist onder meer taalverwerking en leermiddelen voor de computer. Bovendien gebruiken we een topografische data-analyse, om zo verborgen structuren in de data op te sporen.'

De samenwerking met IBM Research moet ertoe leiden dat de Wageningse aardappelonderzoekers sneller en gericht kunnen

veredelen. Visser: 'We kennen alle genen in het aardappelgenoom, maar we weten nog niet precies wat ze doen. Stel: we willen een droogteresistente aardappel ontwikkelen. Dan zijn er bijvoorbeeld vijfhonderd genen die correleren met droogteresistentie, maar welke zijn de belangrijkste?' Door de datasets te koppelen, kan Watson aangeven welke tien genen veelbelovend zijn om verder onderzoek aan te doen en eventueel al de kruisingsprogramma's mee te begijnen.'

Visser ziet veel toepassingen voor het computerprogramma in Wageningen. 'In de biologie heb je veel ongestructureerde datasets die je moet combineren om complexe vragen te kunnen beantwoorden. Denk aan het persoonlijk voedingsadvies en aan milieuvraagstukken. Dan heb je veel rekenkracht en slimme programma's nodig.' Cru-

ciaal in de benutting van big data is uiteindelijk niet de hoeveelheid informatie – die neemt nog steeds exponentieel toe – maar het stellen van de goede vragen, aldus Visser.

Parida wil graag samenwerken met Wageningen UR, omdat die wereldleider is op het gebied van landbouw- en milieuwetenschappen. 'Om computerprogramma's te ontwikkelen die kunnen redeneren, begrijpen en leren om complexe vraagstukken op te lossen, heb je de gecombineerde kennis nodig van computerdeskundigen en vakspecialisten. Die twee partijen moeten voortdurend samenwerken om programma's als Watson te 'leren' hoe die het vraagstuk moet oplossen.' Daarmee wordt Data Sciences een nieuwe wetenschappelijke discipline, aldus Parida, waarbij onderzoekers en bedrijven uit meerdere disciplines betrokken zijn. **AS**