



Data Governance in Big Data

Uitdagingen & kansen omtrent data
governance in de life sciences

Symposium ter ere van de 98e Dies Natalis van Wageningen UR
Woensdag 9 maart 2016



WAGENINGEN UR
For quality of life



Data Governance in Big Data

Uitdagingen & kansen omtrent data governance in de life sciences

Big Data is een opkomend veld dat veel kansen biedt voor nieuwe ontwikkelingen in de Agrofood sector & life sciences. Dit brengt ook uitdagingen met zich mee, bijvoorbeeld . Ddata governance. Data Governance stelt , en de lastige vraag 'van wie is data en wie mag er wat mee?', zijn lastige vragen die cruciaal zijn om de kansen die Big Data biedt te verwaarden.

Big Data is een van de strategische thema's van Wageningen UR. Onderzoekers in verschillende domeinen buigen zich over de mogelijkheden van Big Data. De verwachtingen zijn hoog. De 98e Dies Natalis van Wageningen UR in maart 2016 stond dan ook in het teken van Digitalisation of Nature.

Ter ere van de 98^e Dies Natalis is voorafgaand aan de Dies viering een speciaal symposium georganiseerd over Data Governance in Big Data. Het symposium bracht 43 deelnemers vanuit het bedrijfsleven, overheid, maatschappelijke organisaties en onderzoekers bij elkaar om uitdagingen rondom data governance in big data applicaties met elkaar te bespreken, mogelijke oplossingen te delen, en een strategie te bediscussiëren om big data in de agrofood sector een succes te maken.

Take home messages

1. Smart Use of Data: Voor het ontwikkelen van nieuwe kennis op basis van big data toepassingen is het noodzakelijk om over sectoren en domeinen te werken en data te koppelen.

De grootste kans voor Big Data in de agro-food sector en life sciences is het koppelen van verschillende data bronnen, met name data uit verschillende sectoren en domeinen (bv. plant, dier, consument, keten, ruimtelijk). Bij het ontwikkelen van dit soort koppelingen is het veilig kunnen delen van data cruciaal. Het eigendom van de betreffende data moet hierbij geborgd worden. Data eigenaar en data gebruiker moeten een duidelijke overeenkomst met elkaar aangaan op vlak van toekenning van de data, aansprakelijkheid en licenties.

2. Access to Data: Een veilige plek om data te delen is van belang om op basis van bestaande en nieuwe data nieuwe kennisontwikkeling in de agro-food en life sciences te realiseren. Een data-platform zou deze veilige plek kunnen organiseren.

Data governance gaat voor een groot deel over het delen van data. Het delen van data is belangrijk om in de agro-food sector een competitief voordeel te houden en nieuwe data-gebaseerde toepassingen te ontwikkelen. Vertrouwen in de data en tussen actoren is cruciaal om data te delen. Een gebrek aan vertrouwen kan leiden tot een kip en ei situatie: als er geen data gedeeld wordt, is er geen mogelijkheid om toepassingen te ontwikkelen die de waarde van de data aantonen, en als er geen aansprekende toepassingen zijn is het voor actoren een groetere stap om hun data vrij te geven.

De vraag is of en welke actor(en) de leiding kan nemen voor het organiseren van een veilige plek om data te delen. In verschillende sectoren vinden relevante ontwikkelingen plaats, die stappen zijn op weg naar een dergelijke veilige plek. De vervolgstap is het ontwikkelen van een organisatie-model die een dergelijke veilige plaats vorm kan geven. Randvoorwaarde is dat deze zelf belangeloos is in het gebruik van data.

3. Access to Data: Data standaardisatie is een vitale stap om data-driven applicaties te realiseren.

In verschillende sectoren zijn organisaties aan het werk om de data stromen te standaardiseren. Deze standaardisatie is



nationaal en internationaal cruciaal om efficiënte uitwisseling van data mogelijk te maken. In de FAIR data principes is standaardisatie ook een cruciale stap onder de I van 'interoperability'. Binnen de verschillende sectoren zijn deze standaarden geaccepteerd en in gebruik. De agro-food sector en life sciences kunnen hier van leren. Een volgende stap is de verdere ontwikkeling van deze standaarden en het koppelen van de standaarden over sectoren heen.

4. Change in Systems Thinking: De perceptie en bekwaamheid in data-driven toepassingen voor big data verschilt in de verschillende sectoren.

De kansen die Big Data biedt zijn niet bij iedereen bekend. Dit geldt ook voor actoren in de in de agro-food sector. Deze onbewuste onbekwaamheid over de kansen van big data en data-gebruik kan leiden tot weerstand bij individuele ondernemers om hun (bedrijfs-) data te delen. Er zijn echter ervaringen dat het delen van data voor individuele ondernemers geen probleem is als dit leidt tot zichtbaar betere dienstverlening, bijvoorbeeld de dienstverlening van coöperaties.

Het bewust bekwaam krijgen van alle belanghebbenden is een gezamenlijke uitdaging voor zowel overheid, onderzoek als bedrijfsleven. Het bewust bekwaam maken van actoren richt zich op de capaciteiten van organisaties en mensen. De technische infrastructuur om data te delen moet daarbij geen belemmering vormen. Dit vereist intensieve communicatie over de mogelijkheden van big data toepassingen en het tonen van goede voorbeelden.

Wageningen UR & Big Data

Big Data is een van de strategische thema's van Wageningen UR voor de jaren 2015-2018. Rondom dit thema is een programma opgezet dat de initiatieven en activiteiten binnen WUR met elkaar afstemt en verbindt. Het doel is data-driven research: big data als middel om betekenisvol onderzoek te doen in de agro-food en life sciences. De verschillende onderzoeksgroepen van Wageningen UR zijn actief in het toepassen en onderzoeken van Big Data. Een greep uit de activiteiten:

- Ontwikkeling van FAIR principes. FAIR is Findable, Accessible, Interoperability, R= Reusable (zie rechts).
- Het ontwikkelen van een Big data infrastructuur voor smart farming (precisielandbouw).
- Het verbeteren van resource efficiency van dierlijke productie met behulp van Big Data.
- Data inzetten voor het checken van allergieën, gezondheid en informatie over voeding optimaliseren.
- Informatie systemen om Big Data mogelijk te maken.
- Het voorspellen van oogst van gewassen met behulp van verschillende data bronnen.

In de verscheidenheid aan Big Data activiteiten binnen WUR zijn 3 pijlers geïdentificeerd waar het programma Big Data zich op richt:

- Access to Big Data
- Smart use of Big Data
- Change in Systems Thinking

Data governance is een belangrijk aspect van alle drie deze pijlers. Het symposium over data governance moet meer inzicht geven in de huidige ontwikkelingen en uitdagingen op het gebied van data governance in de sector.

Visies op data governance van bedrijven & overheid

Tijdens het symposium gaven verschillende sprekers vanuit bedrijfsleven en overheid hun visie op big data en data governance.

Rene Penning de Vries, boegbeeld van de topsector ICT gaf aan dat 35% van de groei in NL is te koppelen aan een directe ICT toepassing. De huidige ontwikkelingen in ICT en Big Data gaan razendsnel. Dat geldt ook voor de agro-food en life sciences. De kwaliteit van ICT ontwikkelingen in Nederland scoren internationaal hoog (vergelijkbaar met de VS). De ICT gemeenschap in Nederland staat echter vrij los van ontwikkelingen in andere sectoren. Het samenbrengen van deze sectoren is een aandachtspunt voor de Topsector ICT.

"Big Data can become Big Value"

De grote bedrijven (bijv. Apple, Google, Facebook) doen grote investeringen in Big Data ontwikkelingen en zijn in staat



De FAIR principes. www.nature.com/sdata

daaruit ook veel waarde te creëren. De uitdagingen voor de agro-sector is hetzelfde te doen: waarde te maken van de data die in de sector wordt gegenereerd.

De uitdagingen op dit gebied hangen niet alleen af van technologische ontwikkeling. Ook ethische, organisatorisch en culturele uitdagingen spelen een rol. Een belangrijke vraag is: hoe ga je samenwerken. Dat is vaak moeilijker dan men verwacht.

Conclusies van Penning de Vries over data governance:

- Big data can be big value
- Value creation is more than R&D
- Data Governance is a multi-disciplinary process
- Commit2Data is the national BD Platform

Volgens **Harry Schmeitz** (FruG1.Com) is de AGF branche vergelijkbaar met ICT: het produceren van AGF producten is grotendeels gebaseerd op data. Data geeft tuinbouwers tevens veel informatie aan tuinbouwers over de markt. Om de exponentieel groeiende hoeveelheid data te benutten is een aspect heel belangrijk: verbinden. Data uit verschillende systemen die bij tuinbouw betrokken zijn moet met elkaar worden verbonden om waarde te creëren. Op dit moment is die data gestructureerd en ongestructureerd aanwezig.

"Tuinbouwers moeten meer bewust bekwaam worden van de kansen die big data biedt voor hun business en de sector"

Er zijn verschillende uitdagingen om de data-ontwikkeling in de AGF sector verder te brengen: Enerzijds zijn veel tuinbouwers in Nederland onbewust onbekwaam op het gebied van big data en de mogelijkheden die het biedt. Daardoor ontstaat ook een terughoudendheid om de data te delen. Dit is een gemiste kans op het verbeteren van de marktpositie. Het bewust bekwaam maken van tuinbouwers is een volgende stap. Daarvoor moet duidelijk worden gemaakt wat de meerwaarde is voor de tuinbouw aan de hand van praktische voorbeelden.

Daarnaast speelt Keten-data-mining een rol: data in de keten zit bij zowel overheid, sector en bedrijven. Die data moet beschikbaar worden en worden gedeeld. Data open stellen zou veel kansen geven voor de sector. Een mogelijkheid zou zijn om een platform voor data op te richten waarin alle data samenkomt. Daarbij moeten goede afspraken worden gemaakt. De openstelling de data is afhankelijk van de gebruikers en de toepassingen.

Harry Smit van Food & Agri Research van de [Rabobank](#) ziet dat door de grotere hoeveelheden data de keuzes die boeren moeten maken veel complexer worden. Van '0-1 beslissingen' (het wel of niet sproeien) wordt nu op basis van veel data beslissingen genomen over de hoeveel sproeien op wel perceel en hoeveel kunstmest gebaseerd op het ras dat je hebt gezaaid op welk deel van het perceel op welk moment. Dit komt de efficiëntie van productie ten goede.

Veel keuzes op het boerenbedrijf zijn grotendeels intuïtief en minder gebaseerd op feiten. Er zijn nog veel verschillen tussen de beste en de minste boer. De 'groene vingers' van de boer spelen hierin een grote rol. Die verschillen zijn intuïtief, maar kunnen met behulp van slimme data toepassingen verbeteren. Hierdoor kan je gedetailleerder gaan sturen. Dit vraagt wel een andere manier van handelen van de boer.

"Big Data zet de groene vingers van de boer om in fact-based beslissingen"

Momenteel ontstaat er echter een wildgroei van apps voor het boerenbedrijf. Boeren moeten data invoeren in verschillende systemen voor verschillende toepassingen. Dit is tijdsintensief en zou veel efficiënter kunnen.

Smit ziet het in het belang van de boer dat er een level playing field komt van aanbieders van data diensten, zodat data diensten, diegenen die apps ontwikkelen hetzelfde vertrekpunt hebben en concurreren op basis van kwaliteit van hun apps. Daarvoor is het nodig dat apps ontwikkelaars data beschikbaar hebben en goede apps kunnen maken.

Bijvoorbeeld door op een plek, een platform of data coöperatie, waarin boeren al hun data uploaden. Zo ontstaat een ultieme database op basis waarvan apps en toepassingen worden gemaakt. Volgens Smit is het creëren van een platform in collectief verband noodzakelijk. Dit zou een speciale verantwoordelijkheid van de coöperaties kunnen zijn om deze ontwikkeling op gang te brengen.

Sijne van der Beek van [CRV](#) geeft aan dat de fokkerijsector al decennia bezig is met Big Data. In 1874 is gestart met de aanleg van een stamboek, waarmee data over runderen is vastgelegd. In de fokkerijwereld is big data business-as-usual. Wat er bij komt is de exponentiele groei van data door sensoren en meetinstrumenten, en het ontsluiten van de data door middel van internet. Deze nieuwe ontwikkelingen kunnen worden ingezet om dienstverlening te verbeteren. Hierbij komen ook nieuwe governance vraagstukken kijken.

Ven der Beek ziet het grootste probleem in de huidige ontwikkeling dat er een idee is ontstaan dat data het nieuwe olie is. Dit hindert het delen van data. Data is als koken. Data zijn de ingrediënten. Pas door de juiste ingrediënten bij elkaar te zetten in het juiste recept en met de beste kok, krijg je nieuwe waarde.

De oplossing voor het beter delen van data begint bij de boer/veehouder. Boeren zien nu vaak te weinig terug van het delen van hun data. De data is van de veehouder en de voordelen moeten ook bij hem terecht komen. Dit moet goed geregeld worden via platform of databases zo dat de data beschermt is. Data kan benut worden door derden die gemachtigd worden. Het uitgangspunt daarbij moet echter zijn dat de boer centraal staat. Waarde toevoegen voor de veehouder is de kern.



Arie Baak van [Euretos](#), een kennisplatform voor bioinformatics, stelt ook het aspect van data eigendom aan de orde. Op het internet zijn er veel collecties met data beschikbaar, een van de vragen is of die wel voldoen aan IPR. In het geval van coöperatief data delen, zou het duidelijk moet zijn van wie de data is, zodat diegene bepaald wat er mee gebeurt. Dat geeft mensen en bedrijven meer zekerheid. Data wordt dan gekoppeld met ownership gedeeld. Ook het is mogelijk om nano-publicatie te bouwen op basis van het valideren van unieke statements op basis van een synthese uit wetenschappelijke artikelen. Het delen van deze nano publicaties kan dan samen een grote data set opbouwen, waar degene die data delen allemaal profijt van hebben.

"Waarde toevoegen voor de boer staat centraal"

Ook **Rob van Winden** van [Agrifirm](#) stelt dat de coöperatieve sfeer waarin data wordt gedeeld van belang is. De ambitie van Agrifirm is het creëren van waarde voor veehouders en akkerbouwers om het bedrijfssaldo te verbeteren. Dat doen ze door: werken aan uniformiteit van data, het creëren van duidelijkheid over eigenaarschap, het realiseren van betrouwbare data snelweg. Dat doet Agrifirm voor o.a. [Smart Dairy Farming](#) en [Akkerweb](#)

Akkerweb biedt telers de gelegenheid om bepaalde data voor een bepaalde tijd en doeleinde beschikbaar te stellen. Dat doen ze door die data en voorkeuren aan te vinken. Zo worden afspraken gemaakt tussen data eigenaar en gebruiker. Cruciaal hierin is ook de technische applicatie. Op dit moment zijn er veel loketten waar de boer zijn data moet delen en vinden. Als er een loket zou zijn is dan praktischer.

Maarten Kool van het Ministerie van Economische Zaken stelt dat ook het vertrouwen in de kwaliteit van data van het grootste belang is, en hoe deze kwaliteit gewaarborgd kan worden.

