
EBEO2.0 Living Labs: verslag workshop 4



Jip de Vries¹, Marijn Gülpen¹, Ralf Verdonchot¹

¹ Wageningen Environmental Research

Deze workshop is georganiseerd door Wageningen Environmental Research in opdracht van STOWA

Wageningen Environmental Research
Wageningen, maart 2025

Inhoud

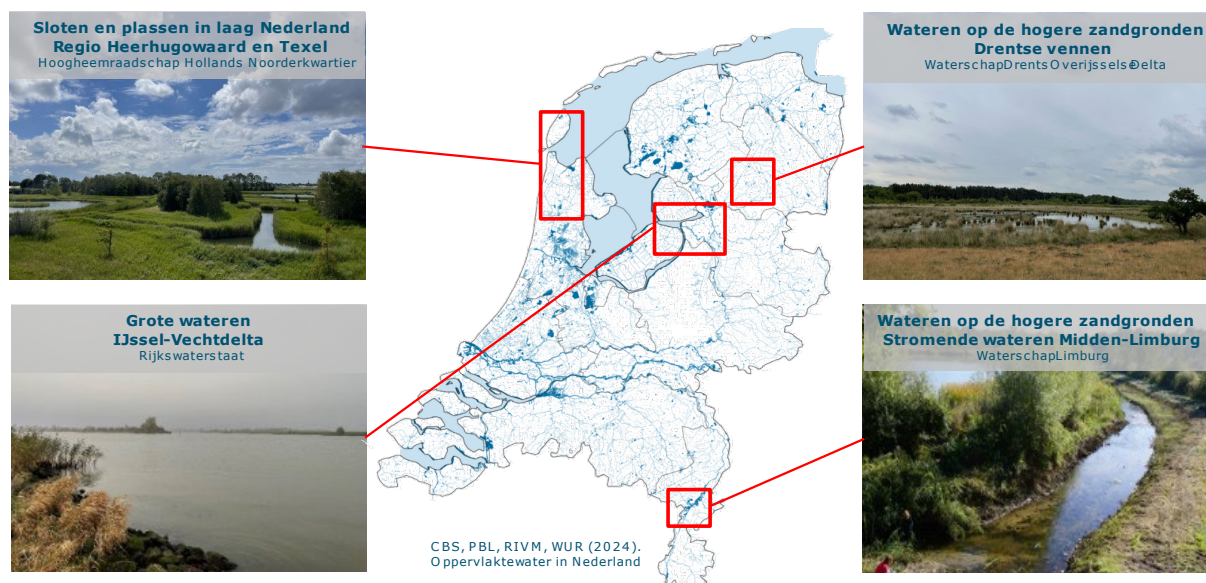
1	Inleiding	2
2	Ervaringen, discussiepunten en potentiële oplossingsrichtingen met data- doorstroom	5
3	Inventarisatie van de behoeften in dataopslag van nieuwe technieken	8
	Bijlage 1: Deelnemerslijst	9
	Bijlage 2: Agenda	10

1 Inleiding

Hoewel de waterkwaliteit de afgelopen jaren is verbeterd, is de toestand van veel waterlichamen in Nederland momenteel nog ontoereikend. Voor deze waterlichamen is herstel noodzakelijk om aan de doelen van de Kaderrichtlijn water (KRW) te voldoen. Om effectieve maatregel(en) te kunnen nemen om het gewenste herstel te realiseren, is het belangrijk per waterlichaam de oorzaken van de ontoereikende toestand vast te stellen. De huidige beoordelingssystematiek heeft echter (te) weinig diagnostische waarde. We kunnen ermee op hoofdlijnen zien hoe de toestand op enig moment is, maar we weten vaak onvoldoende waarom de toestand is zoals die is.

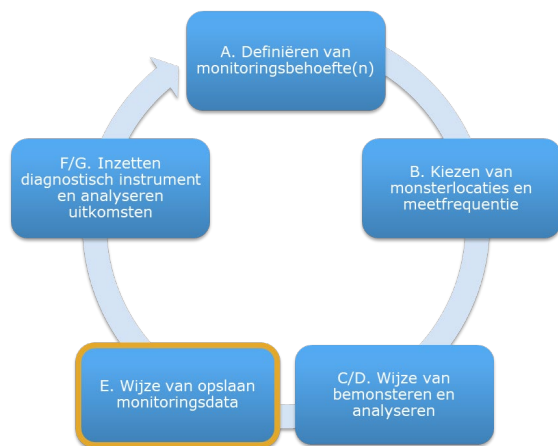
Het doel van het EBEO2.0 Living Labs project is om een nieuwe methode te ontwikkelen voor een diagnostische beoordeling van de ecologische waterkwaliteit. Voor het ontwikkelen en in de praktijk toetsen van de nieuwe EBEO2.0 methode zijn drie 'levende laboratoria ingericht' die recht doen aan de verschillende watertypen die we in Nederland hebben (Figuur 1.1). Het gaat om:

- **Grote wateren, IJssel-Vechtdelta**, i.s.m. Rijkswaterstaat;
- **Sloten en plassen in polders van laag Nederland**, regio Heerhugowaard en Texel, i.s.m. Hoogheemraadschap Hollands Noorderkwartier;
- Wateren op de hogere zandgronden, met twee subtypen:
 - a) **Stromende wateren, Midden-Limburg** i.s.m. Waterschap Limburg
 - b) **Drentse vennen**, i.s.m. Waterschap Drents Overijsselse Delta.



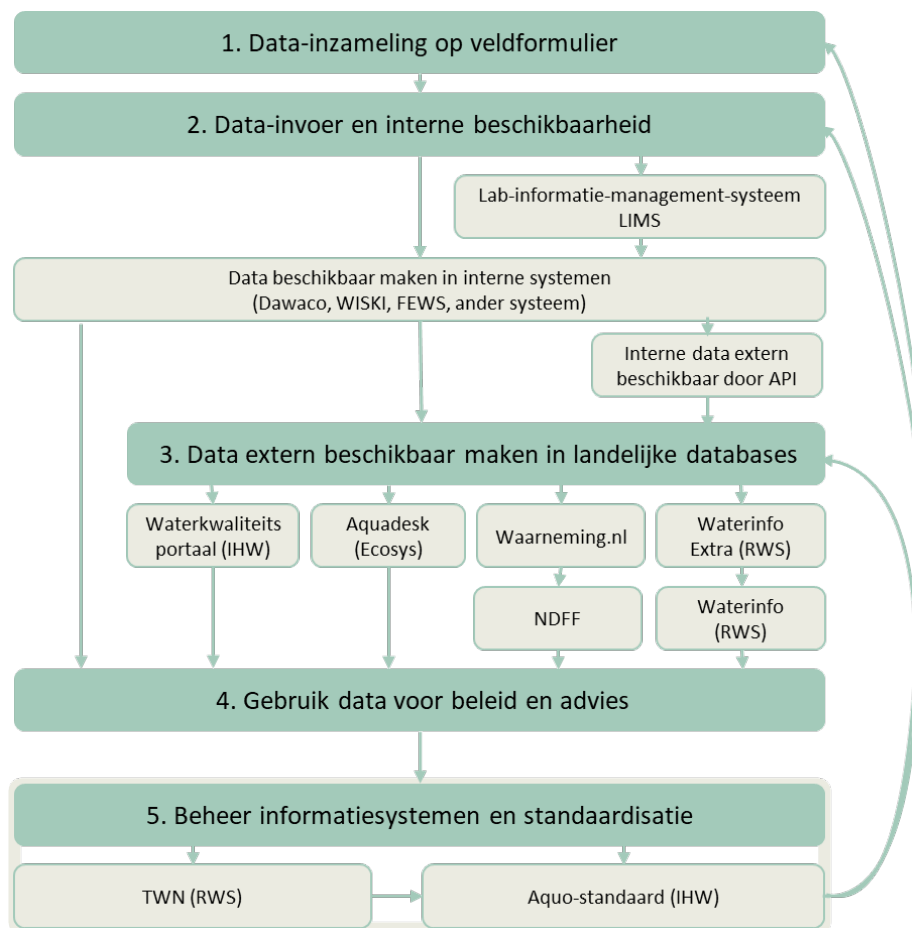
Figuur 1.1 *Overzicht van de ligging van de Living Labs van EBEO2.0.*

In iedere projectfase zal een workshop worden georganiseerd om te zorgen dat er zo veel mogelijk aansluiting is met de praktijk (Figuur 1.2). Dit is het verslag van de vierde workshop van het EBEO2.0 Living Labs project van 17 maart 2025 in Amersfoort met als thema 'Wijze van opslaan van monitoringsdata'.



Figuur 1.2 Overzicht van de fases binnen het EBEO2.0 Living Labs project.

Het doel van de derde workshop was: het evalueren van de wijze van opslaan van monitoringsdata, in het bijzonder meta-data, aanvullende data voor diagnose en data uit nieuwe technieken. Gedurende de middag zijn ervaringen opgehaald met bottlenecks en oplossingsrichtingen voor een goede data-doorstroom en juist gebruik van de data-standaarden door inzamelaars, verwerkers en gebruikers van data.



Figuur 1.3 Overzicht van de 'data-keten' van monitoringsdata.

Er waren bij de deze workshop 12 deelnemers aanwezig vanuit het waterbeheer (waterschappen en Rijkswaterstaat), databeheer (Aqun en Ecosys) en beheer van informatiesystemen en data-standaard (InformatieHuis Water) (voor de volledige deelnemerslijst, zie bijlage 1). De deelnemers zijn uitgenodigd vanwege hun expertise in verschillende delen van de 'data-keten'. Deze keten is in een voorstelronde in beeld gebracht, en bestaat uit het verzamelen van data, het invoeren en intern beschikbaar maken van data,

het extern beschikbaar maken van data in landelijke databases, het beheer van informatiesystemen en data-standaarden en het gebruik van data voor beleid en advies (Figuur 1.3). Voor het ophalen van ervaringen en mogelijke verbeteringen rond data-doorstroom is gebruik gemaakt van een invulbare poster. Er is gediscussieerd in twee parallelle sessies.

Dit verslag is als volgt opgebouwd:

- Hoofdstuk 2: Ervaringen, discussiepunten en potentiële oplossingsrichtingen met data-doorstroom
- Hoofdstuk 3: Inventarisatie van de behoeften in dataopslag van nieuwe technieken

2 Ervaringen, discussiepunten en potentiële oplossingsrichtingen met data-doorstroom

Sessie A.

Sessiebegeleider: Ralf Verdonshot (Wageningen Environmental Research)

Deelnemers: Arthur Bennis (Waterschap Drents Overijssels Delta), Barend van Maanen (Waterschap Limburg), Geri Wolters (Ecosys), Hinne Reitsma (InformatieHuis Water), Lidewij Servatius (Aquon), Wouter Abels (Rijkswaterstaat)

1. Verwerking ontbrekende parameters, nieuwe soorten of gewijzigde taxonomische inzichten in naamgeving en andere nieuwe informatie die in databases gestandaardiseerd moet worden vastgelegd

De eerste bottleneck die werd genoemd ging over ontbrekende parameters. Soms is een parameter, soort of methode niet aanwezig in de Aquo-standaard of in de standaard Taxa Waterbeheer Nederland (TWN). Aanpassing van de Aquo-standaard via aanvraag duurt soms heel lang of soms wordt een aangekaart probleem in de lijst (bijv. verkeerde soortnaam) niet doorgevoerd. De reden hiervoor is dat aanpassingen een complex proces zijn, door vele koppelingen in het model. Ook moet er consensus zijn bij discussiepunten of een wijziging wel valide is. Overigens, als de aanvraag - voor een nieuwe parameter of domeinwaarden - duidelijk een aanvulling op de bestaande lijst betreft, kan deze binnen enkele dagen worden doorgevoerd.

De oplossingsrichting die werd aangedragen is het proces van de verwerking van een aanvraag transparanter te maken (wat gebeurt er met een melding, welke stappen worden doorlopen) en beter en sneller te communiceren naar de aanvrager tijdens de verwerking van een aanvraag over de status.

2. Hoe te voorkomen dat foutieve data in de database terecht komt

Een tweede bottleneck die werd genoemd was dat fouten in ingevoerde gegevens vaak pas in een laat stadium worden ontdekt. Er bestaat nog geen automatische validatie van ecologische gegevens, zoals bij chemie (realistische waarden). De opdrachtgever controleert de ingevoerde data niet goed, of is hier zelfs niet toe in staat omdat deze onvoldoende in de materie ingevoerd is. Het niet invullen van benodigde data komt veel minder voor door het gebruik van verplichte velden in digitale veldformulieren.

Een simpele oplossing is de data goed controleren voordat deze opgenomen wordt in de database, maar dit is lastig met veel data en vraagt vaak specialistische kennis van soortgroepen of gebieden, die niet altijd voorhanden is. Een oplossing voor biologische data kan in *datascience* gevonden worden, bijvoorbeeld door het ontwikkelen van een automatisch validatiesysteem voor soorten op basis van bekende verspreiding en regionale en nationale zeldzaamheid. Een alternatief is checken of er in de monitoringsdata van een bepaalde locatie en moment soorten voorkomen die hier eerder niet voorkwamen, of dat er soorten missen die hier eerder wel voorkwamen. Foto's toevoegen als bewijs in de database zou ook een optie kunnen zijn. Verder is het vermelden van de determinandus aan te raden, zodat altijd navraag kan worden gedaan bij twijfel. Echter, dit vastleggen van namen kan mogelijk problemen geven met AGV.

3. Opslag van ruwe gegevens in plaats van bewerkte gegevens

Het is belangrijk de oorspronkelijke ruwe data op te slaan, en niet alleen bewerkte gegevens. Bijvoorbeeld: niet enkel abundanties van soorten omgerekend naar een standaard monsterlengte, maar ook de werkelijk getelde aantallen. Dit vraagt dus ook altijd de beschikbaarheid van metadata. Als er wel wijzigingen doorgevoerd moeten worden, bijvoorbeeld voor een herdeterminatie, dan is het belangrijk dat alle wijzigingen die in het verleden gedaan zijn terug te vinden zijn in de database. Hiervoor moet een methode worden ontwikkeld, waarbij alle wijzigingen in het verleden in het bestand terug te zien zijn.

4. Zorgvuldig vastleggen van metadata

Het is belangrijk metadata altijd vast te leggen. Het komt voor dat laboratoria/personen bepaalde handelingen standaard altijd uitvoeren en dit daarom niet registreren (bijvoorbeeld het gebruik van bepaalde maaswijdte zeven). Het is belangrijk dit altijd wel te doen, want bij veranderingen, bijvoorbeeld een wisseling in personeel, kan een afwijkende werkwijze opeens tot andere resultaten leiden. Deze werkwijze wordt daarom het best gedocumenteerd in de metadata. Ook is het goed links op te nemen bij de data naar voorschriften en naar de rapportage die de resultaten beschrijft, zodat indien nodig altijd nog meer informatie opgehaald kan worden. Dit werkt met terugwerkende kracht, de rapportage is er immers later dan de dataset.

Sessie B.

Sessiebegeleider: Marijn Gülpen (Wageningen Environmental Research)

Deelnemers: Albert Dees (Aquon), Anneke ter Schure (HHNK), Arne Klink (Ecosys), Martin Soesbergen (Rijkswaterstaat), Peter van Rooij (InformatieHuis Water)

1. Benodigde voorkennis

De eerste bottleneck die werd genoemd tijdens de discussie is dat er een veel voorkennis nodig is om met de data te kunnen werken. Deze ligt vooral bij de datagebruikers, die door een gebrek aan kennis van de data verkeerde conclusies kunnen trekken. Niet ontdekte fouten kunnen zorgen voor verkeerd beleid, of voor extra werk wanneer fouten op een laat moment wel worden gezien. Daarnaast kan dit ook voor de databeheerders lastig zijn omdat zij soms erg veel vragen krijgen van de gebruikers.

Een voorbeeld dat werd aangehaald ging over ecologische kennis die nodig is rondom taxonomie. Men is er vaak niet van op de hoogte dat er door de tijd herindelingen plaatsvinden, en dat daardoor oude biologische monitoringsdata niet zomaar vergeleken kan worden met recente data. De trend is op dit moment dat de mensen met deze taxonomische kennis afnemen. Een van de oplossingsrichtingen voor dit probleem specifiek zou zijn dat er een grote vertaaltabel zou komen die duidelijk maakt wanneer welke taxa nieuw zijn gedefinieerd, samengevoegd of juist afgesplitst. Daarnaast is er meer bewustwording nodig.

2. Data-invoer

Door het gebruik van open velden in plaats van een pull-down menu worden veel spellingsfouten gemaakt. Hierdoor is data-analyse erg moeilijk en kunnen deze velden soms niet worden meegenomen. Aan de andere kant is het wel belangrijk om vrij invulbare velden te hebben om belangrijke contextuele informatie te kunnen opnemen. Dit maakt dat de opdrachtgever en de uitvoerende partij de formulieren extra goed zullen moeten controleren voordat ze een database ingaan. Het is daarbij essentieel om duidelijk te hebben wie deze verantwoordelijkheid draagt. Daarnaast kan een goede validatie van de data ook inspelen op deze fouten.

3. Parameterinterpretaties

Parameterinterpretaties kunnen verschillen tussen veldwaarnemers. Zo is er bijvoorbeeld het verschillend geïnterpreteerde onderscheid in flab/draadalg of de begrenzing van de oeverzone. Een oplossing hiervoor is om veldopnames te harmoniseren, door samen het veld in te gaan en samen te kijken waarom er verschillen bestaan in waarnemingen. Alternatief kan er ook een veld-ringtest worden gedaan, waar niet alleen de meest ervaren veldwerkers, maar ook beginners van elkaar kunnen leren. Als de meetlocatie niet verandert zou het ook goed zijn om zo veel mogelijk standaardinformatie zoals de oevergrens al op te nemen en dit niet door de veldonderzoeker te laten invullen.

4. Beschikbaarheid van oude data

Een tweede aandachtspunt betrof historische data die niet (meer) beschikbaar is, terwijl dit wel zou moeten. Er is onvoldoende overzicht van welke historische gegevens er zijn. Dit gaat voornamelijk over oude data, die niet zijn gedigitaliseerd. Verlies van databeschikbaarheid komt ook voor wanneer een organisatie overstapt naar een andere datasysteem of een folderstructuur. Een mogelijke oplossingsrichting zou zijn dat er meer

mankracht wordt ingezet om oude data te archiveren. Prioritering in wat wel of niet relevant zou zijn, is dan wel nodig.

5. Gebruik van data-standaarden

Er werd geconstateerd dat uitvoerders van projecten soms niet de juiste standaarden gebruiken. Het gebeurt nog te vaak dat opdrachtgevers de eisen niet op tijd delen. Men komt er dan pas achter dat er iets niet klopt als het onderzoek al gedaan is. Om de opgehaalde data dan toch nog in de goede standaard te krijgen kost vaak meer tijd en dus ook meer geld. De oplossing is hiervoor dat uitvoerder altijd de standaarden moeten noemen en mee moeten sturen. Communicatie verloopt het liefst niet via een tussenpersoon, maar direct met de monsternemer.

Veranderingen in data-standaarden kunnen ervoor zorgen dat verschillende systemen niet meer op elkaar aansluiten (bijv. TWN en niet-bijgewerkte maatlatten waarop EKR-berekening wordt gebaseerd). Over het algemeen leek het een goed idee om standaarden op vaste momenten te updaten en niet al te vaak zodat de mensen die ermee moeten werken hier ook van op de hoogte zijn. Dit gebeurt bijvoorbeeld eens per jaar betreffende de mogelijke verandering in de maatlatten en wat dit betekent voor de AQUO-standaard. Door dit jaarlijks te doen blijft het overzichtelijk.

6. Validatie en communicatie rondom het aanpassen van data

Verder werden er verschillende punten genoemd rondom de validatie van data. Het is bijvoorbeeld niet altijd duidelijk of er een validatiestap is uitgevoerd, en op welke manier. Daarnaast zijn er onduidelijkheden over wie de verantwoordelijkheid draagt.

Er is een wens om gegevensvalidatie centraal, op landelijk niveau, aan te pakken. Aanleiding is dat monitoringsdata regelmatig nog verandert nadat deze is gerapporteerd (bijv. door het rechtzetten van fouten). Bij het opslaan van data uit bronsystemen worden veranderingen/mutaties in de data niet altijd duidelijk gecommuniceerd met iedere betrokken partij. Communicatie over aanpassingen loopt soms direct tussen gebruikers van de data en waterbeheerders, waarbij echter een aanpassing van de data in landelijke databases mist. Zo kunnen er discrepanties in datasets ontstaan.

Verandering van validatie-eisen van bijvoorbeeld de EU rondom de KRW wordt ook gezien als een knelpunt. Wanneer de eisen voor gegevensvalidatie worden veranderd kan het voorkomen dat het IHW met reeds aangeleverde data terug moet naar de waterbeheerder met extra vragen, en moeten waterbeheerders data heraanleveren.

3 Inventarisatie van de behoeften in dataopslag van nieuwe technieken

Tijdens het afsluitende deel van de workshop is er nagegaan welke nieuwe data er mogelijk beschikbaar kan komen in het werkveld, en waar behoeften liggen in de opslag van deze data.

Hoogfrequente metingen sensoren

Door de ontwikkeling in sensoren zijn er steeds vaker hoogfrequent gemeten tijdsreeksen te verwerken. Er wordt geconstateerd dat het momenteel moeilijk is om data afkomstig van sensoren te verwerken wanneer het gaat om data van continumetingen (bijv. hoogfrequente zuurstofmetingen). Hierbij is niet zozeer de opslag zelf een probleem, maar een gebrek aan standaarden die bijvoorbeeld uitwisseling van data zouden vergemakkelijken. Een suggestie die wordt gedaan is het gebruik van opslagsystemen afkomstig uit de hydrologie, zoals FEWS.

eDNA

Daarnaast is het gebruik van eDNA in opkomst als monitoringsmiddel. Hiervoor is een datastandaard in ontwikkeling door het Informatiehuis Water. Deze heeft als primair doel het uitwisselen van data te stroomlijnen; dit gaat boven het vastleggen van alle data. Er wordt door enkele deelnemers echter ook belang gehecht aan de ruwe *sequence*-data. De reden hiervoor is dat met de ontwikkeling van het werkveld, en in het bijzonder het steeds completer worden van de *libraries*, het waarschijnlijk in de toekomst mogelijk zal zijn om met een herhaling van de analyse meer informatie te halen uit de ruwe data. Vaak staat ruwe *sequence*-data ook niet elders in een database, maar wordt onderling projectmatig in pdf's gedeeld. Deze informatie zou wel kunnen worden gepubliceerd in openbare databases, maar dit is niet de standaard aanpak. Er is geen consensus onder de deelnemers over het belang van het opslaan van deze data.

Overig

Andere vormen van data die in de toekomst meer gebruikt kunnen gaan worden zijn videomateriaal (bijvoorbeeld dronebeelden), data afkomstig uit bio-assays en data uit uitgebreide morfologische opnames van watergangen (bijvoorbeeld een Gewasserstrukturgütekartierung). Nu wordt dit type informatie alleen projectmatig opgeslagen en gedeeld. In het algemeen is er een behoefte om de aanpak rond de opslag en uitwisseling van grote datasets, ongeacht de informatiebron, te stroomlijnen.

Bijlage 1: Deelnemerslijst

Naam	Organisatie
Albert Dees	Aquon
Anneke ter Schure	Hoogheemraadschap Hollands Noorderkwartier
Arne Klink	Ecosys
Arthur Bennis	Waterschap Drents Overijssels Delta
Barend van Maanen	Waterschap Limburg
Geri Wolters	Ecosys
Hinne Reitsma	InformatieHuis Water
Jip de Vries	Wageningen Environmental Research
Lidewij Servatius	Aquon
Marijn Gülpen	Wageningen Environmental Research
Martin Soesbergen	Rijkswaterstaat
Peter van Rooij	InformatieHuis Water
Pui Mee Chan	STOWA
Ralf Verdonschot	Wageningen Environmental Research
Wouter Abels	Rijkswaterstaat

Bijlage 2: Agenda

12.00-13.00 Inloop en lunch

13.00-13.10 Welkom en introductie EBEO2.0 – Pui Mee Chan

13.10-13.40 Inventarisatie data-keten en terugblik workshop 3

13.40-14.45 Ervaringen met data-doorstroom delen – Deel 1

14.45-15.00 Koffie

15.00-15.40 Ervaringen met data-doorstroom delen – Deel 2

15.40-16.20 Nieuwe technieken

16.20-16.30 Afsluiting

Wageningen Environmental Research
Postbus 47
6700 AA Wageningen
T 0317 48 07 00
wur.nl/environmental-research

Wageningen Environmental Research
Rapport
ISSN 1566-7197



De missie van Wageningen University & Research is 'To explore the potential of nature to improve the quality of life'. Binnen Wageningen University & Research bundelen Wageningen University en gespecialiseerde onderzoeksinstituten van Stichting Wageningen Research hun krachten om bij te dragen aan de oplossing van belangrijke vragen in het domein van gezonde voeding en leefomgeving. Met ongeveer 30 vestigingen, 7.600 medewerkers (6.700 fte) en 13.100 studenten en ruim 150.000 Leven Lang Leren-deelnemers behoort Wageningen University & Research wereldwijd tot de aansprekende kennisinstellingen binnen haar domein. De integrale benadering van de vraagstukken en de samenwerking tussen verschillende disciplines vormen het hart van de unieke Wageningen aanpak.