

<b>Algemene gegevens</b>	
PPS-nummer	1406-01
Titel	BIGt&u: Luisteren naar consument en keten
Topsector en innovatiethema	T&U
Projectleider (onderzoek)	Tim Verwaart (LEI Wageningen UR)
PPS-coördinator (namens private partij)	Wilco van den Berg (GroentenFruit Huis)
Contactpersoon overheid	Kees van Drunen (Ministerie van EZ)
Status (lopend of afgerond)	Lopend
Type onderzoek (F, T of V)	Gedeeltelijk F, gedeeltelijk T
Werkelijke startdatum	1 januari 2015
Werkelijke einddatum	Verwacht 31 december 2018
Korte omschrijving inhoud	Onderzoek naar mogelijkheden voor ontwikkeling van gezamenlijke IT-infrastructuur (HortiCube), governance en businessmodel voor het verzamelen en integreren van markt-data uit een grote diversiteit aan bronnen en Social Media.

<b>Highlights</b>
<p>In 2015 is er gewerkt aan diverse resultaten:</p> <ul style="list-style-type: none"> <li>• De afdeling Wiskunde van de Vrije Universiteit Amsterdam heeft methoden en algoritmen ontwikkeld voor het ontsluiten van Twitter data ten behoeve van de topsector. Er wordt real-time statistieken verzameld van, voor de sector, relevante twitter berichten. Dit heeft geleid tot de ontwikkeling van een real-time Top 10 Groenten en Fruit en Top 10 Bloemen en planten overzicht waardoor men in staat is om te ontdekken waar de consument over praat. Een dergelijk overzicht over heel 2015 onderverdeeld naar seizoenen is ook visueel gemaakt.</li> <li>• In samenwerking met de private partners zijn pilots in verzameling van data op Social Media uitgevoerd en in diverse gremia gepresenteerd:       <ul style="list-style-type: none"> <li>○ Nationale tulpendag januari 2015</li> <li>○ Twuijnbijt campagne</li> <li>○ Lievelingsbloem campagne</li> <li>○ Herkenbaarheid van kleur van sierteeltproducten in Twitterberichten</li> </ul> </li> <li>• Door de private partners zijn lijsten van relevante trefwoorden voor verzameling van data op Social Media opgesteld en is over de use cases overleg met de achterban van de organisaties georganiseerd.</li> <li>• In samenspraak met de Business Council Markt &amp; Statistiek binnen GroentenFruit Huis is een plan van aanpak opgesteld met betrekking tot analyse Twitter. Er zal in 2016 worden ingezet op: top-10 actueel en lange termijn, alert functie, themadetectie en effect/eventmeting. Als thema's om te onderzoeken zijn de volgende vastgesteld : maaltijdboxen, retail in NL, eten/koken (wat gebeurt er achter de voordeur), biologische groente &amp; fruit en tweets koks/horeca. In samenspraak met de sierteeltsector zijn verzameling van Twitterberichten gerelateerd aan marketingcampagnes en verzameling van data over trends onder consumenten gepland.</li> <li>• In het fundamentele ICT onderzoek is er een update gemaakt van het bestaande linked data stappenplan voor het linked maken van databronnen. Daarin is gebruik gemaakt van diverse tools die het in combinatie mogelijk maken om ruwe data (in csv of PDF vorm) te cleanen, om te zetten naar RDF linked data format en toegankelijk te maken dmv een linked data mechanisme.</li> </ul>

- Het nieuwe linked data stappenplan is gebruikt om diverse databronnen linked te maken en in een primaire versie van de HortiCube te ontwikkelen. Daarbij gaat het om data van RvO, WAPA, KCB en KNMI.
- Door de private partner Floricode is de GPC als toekomstige standaard voor productclassificatie in de sierteeltketen ontwikkeld. Op basis hiervan is in samenwerking met private partners VGB en Flora Holland een conceptvoorstel opgesteld voor toepassing van deze classificatie in de internationale handelsstatistieken, met het oog op betere aansluiting bij informatiestromen in de keten en vermindering van lastendruk door uniformering van gebruikte coderingen.
- Er is een start gemaakt met een algemene ontologie voor het mappen/afbeelden van soortgelijke termen in verschillende databronnen. Hierin speelt de GPC een centrale rol. Een specifiek invulling van het mappen van de codes is gemaakt voor de appel/peren namen/codes die worden gebruikt binnen de WAPA en de KCB bronnen.
- Met de mappings is een pilot in uitvoering bij Fruitmasters (lid van GroentenFruit Huis), in samenwerking met de private partners GroentenFruit Huis en Frug I Com. Op basis van deze pilot is door Frug I Com een nieuwe versie van de groenten- en fruitcodering in de internationale GPC-standaard voorgesteld. Deze wordt naar verwachting medio 2016 van kracht.
- Er is een eerste versie beschikbaar van een demo applicatie van de appel/peren use case. Daarin kunnen productie en export gegevens van appels en peren in EU landen worden getoond in grafieken en staaftabellen. Dit geeft beter inzicht in de trends en ontwikkelingen in deze markt dan mogelijk is op basis van de ruwe data zelf.
- Onderzoek naar aspecten van gegevensbescherming en data governance op basis van de BOMOD-methode is gestart en wordt in 2016 voortgezet. TNO en DLO werken samen met private partner Nederland ICT in onderzoek om naar passende governance-modellen voor aspecten zoals privacy, concurrentie etc. die in het bedrijfsleven vragen oproepen.
- Door private partner Q-ray is onderzoek verricht naar toepassingsmogelijkheden van NoSQL data bases. De resultaten daarvan worden in 2016 toegepast in de HortiRadar, de IT-component voor verzamelen en terugzoeken van uitingen op Social Media.
- Overkoepelend informatiemodel voor metadata, informatiearchitectuur en requirements voor de eerste versie van de HortiCube zijn opgesteld door de kennisinstellingen in samenwerking met private partner Q-ray.
- Open Data over productielocaties van opengrondsteelten uit de Basisregistratie Percelen van RVO zijn onderzocht op toepasbaarheid, en goed bruikbaar bevonden en voor opname in het MIT dashboard.
- Ook voorlopige (vroeg beschikbare) data over glastuinbouwarealen uit de landbouwtelling zijn onderzocht op toepasbaarheid, maar de conclusie is dat gebruik van deze data in de HortiCube voor marktinformatie geen toegevoegde waarde biedt boven de cijfers die het CBS publiceert op basis van de Landbouwtelling.
- RIKILT heeft voor de deelnemende partijen relevante data bronnen geïdentificeerd en besproken (bijvoorbeeld voedselveiligheid: RASFF (voor EU), GEMSFOD (wereldwijd; WHO), voorspellingen van gewasgroei, ziektes etc (wereldwijd, WHO: GIEWS), media rapporten ziektes, voedselveiligheid, en fraude: EU: EMM).
- RASFF en EMM zijn verder geanalyseerd als "early warning" systeem. Een Bayesians Netwerk (BN) model voor fraude waarmee ook voorspellingen gemaakt kunnen worden van de type fraude in groente en fruit notificaties in RASFF is gemaakt en gedemonstreerd. Verdere besluitvorming over toepassing van dit model vindt in 2016 plaats.
- Door private partners Flora Holland en GroentenFruithuis is het project besproken en zijn requirements verzameld met respectievelijk ICT klankbordgroep en

Business Council.

- Met private partner VGB is onderzoek gestart hoe de daar ontwikkelde voorziening van marktinformatie voor leden ook in andere sectoren kan worden toegepast en welke data uit de horticulture daarop aangesloten kunnen worden.
- Door private partner Nederland ICT is de website [www.bigtu.nl](http://www.bigtu.nl) ontwikkeld en is gestart met onderzoek naar governance en in het bijzonder convenanten voor datagebruik, waarbij de resultaten van het Smart Industrie PPS als voorbeeld dienen.

## **Opgeleverde producten in 2015**

### Wetenschappelijke artikelen:

- "Modeling trend progression through an extension of the Polya Urn Process" (submitted)
- "HortiCube: A Platform for Transparent, Trusted Data Sharing in the Food Supply Chain" (accepted for presentation in "System Dynamics and Food Networks", Iglis, Austria, februari 2016)

### Rapporten:

- GPC standard voor productclassificatie in de sierteelt  
<http://www.floricode.com/nl-nl/coderen/classificatiecoderinggpc.aspx>
- Informatiearchitectuur voor HortiCube, HortiConnector en HortiRadar. (Actuele versie van modellen en toelichting wordt beschikbaar gemaakt op [www.bigtu.nl](http://www.bigtu.nl))

### Inleidingen/ workshops/ invited lectures (presentaties zijn beschikbaar als downloads van [www.bigtu.nl](http://www.bigtu.nl)):

- Floricode relatiemiddag "Sierteelt Digitaal" 22 januari 2015: Presentaties over en discussie met ondernemers uit de sector over mogelijkheden van Big Data en analyse van Social Media door [Frans Feldberg](#) en [Sandjai Bhulai en Tim Verwaart](#).
- KAVB-Ledendag 22 januari 2015: Presentatie over analyse van Social Media door [Sandjai Bhulai](#)
- Tuinbouw Digitaal Community Meeting 14 april 2015: presentatie door [Wilco van den Berg en Tim Verwaart](#)
- Europese workshop "Big data for food, agriculture and forestry: opportunities and challenges" 22 september 2015: presentatie van het project door [Tim Verwaart](#)
- Kennisdag voor de Fruitteelt 2015: De fruitketen van teelt tot consument, 20 november 2015: presentatie door [Wilco van den Berg](#)
- EU Fresh Info 2 december 2015: Presentaties door [Wilco van den Berg](#) en [Silja Eckartz en Jack Verhoosel](#)