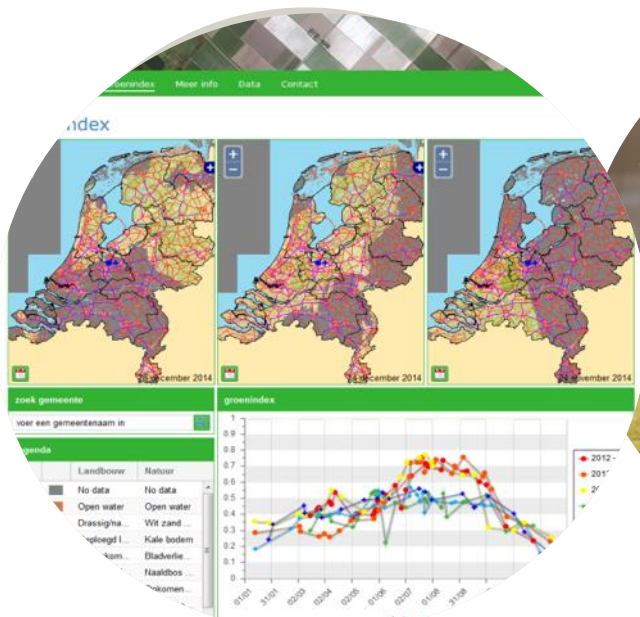


# AgroDataCube: A Big Open Data Collection for Agri-Food Applications

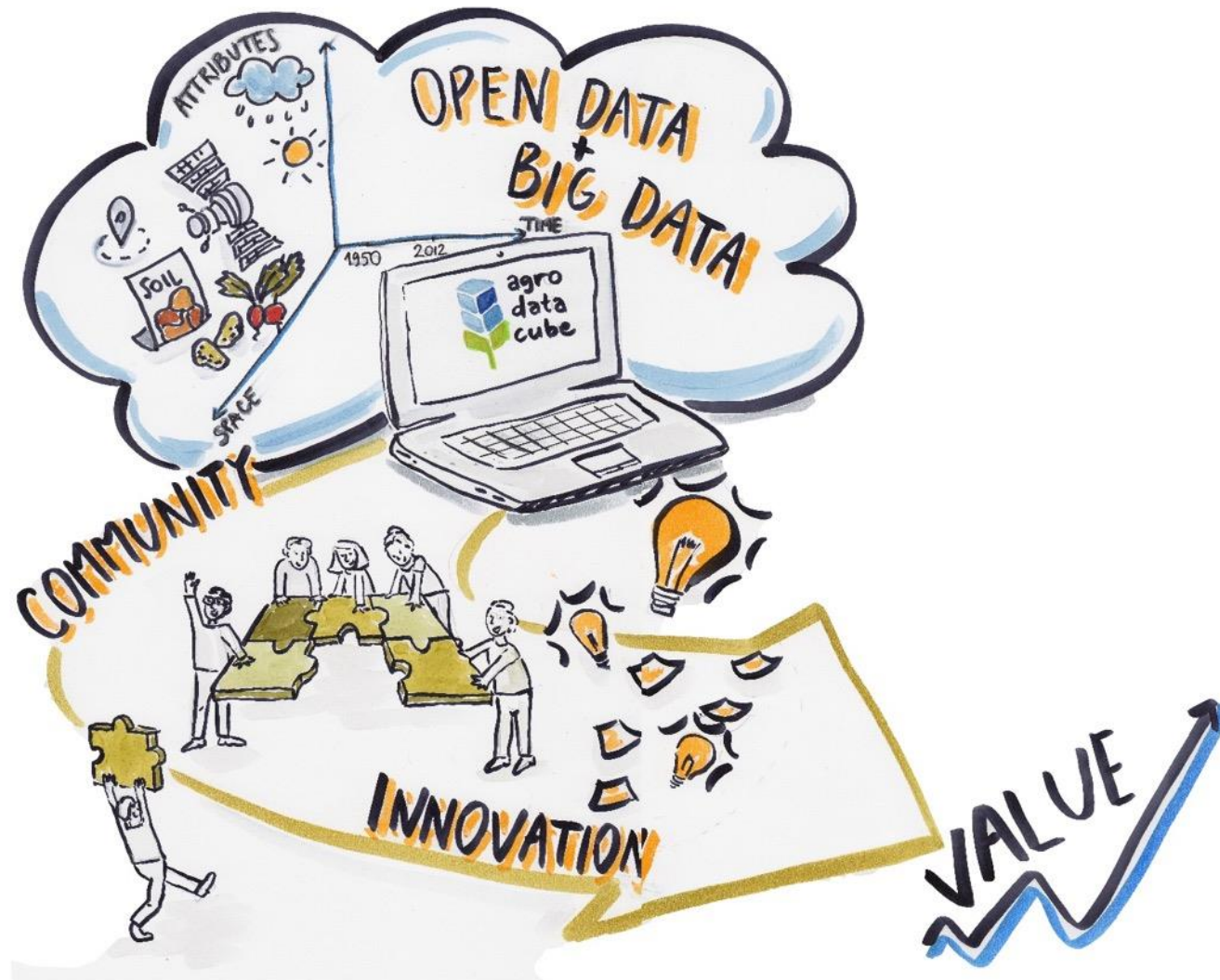
*RDNL, UKB werkgroep 28 November 2018*

Rob Knapen, Sander Janssen, Henk Janssen, e.a.,  
Wageningen Environmental Research, @wurcgi



# Contents

- What is AgroDataCube?
- WUR Data Management Support
- Lessons Learned / Recommendations



Agro



Meteo



Bodem



Remote Sensing



Open Data

Vector

Excel

CSV

Grid

Verzamelen



Integreren



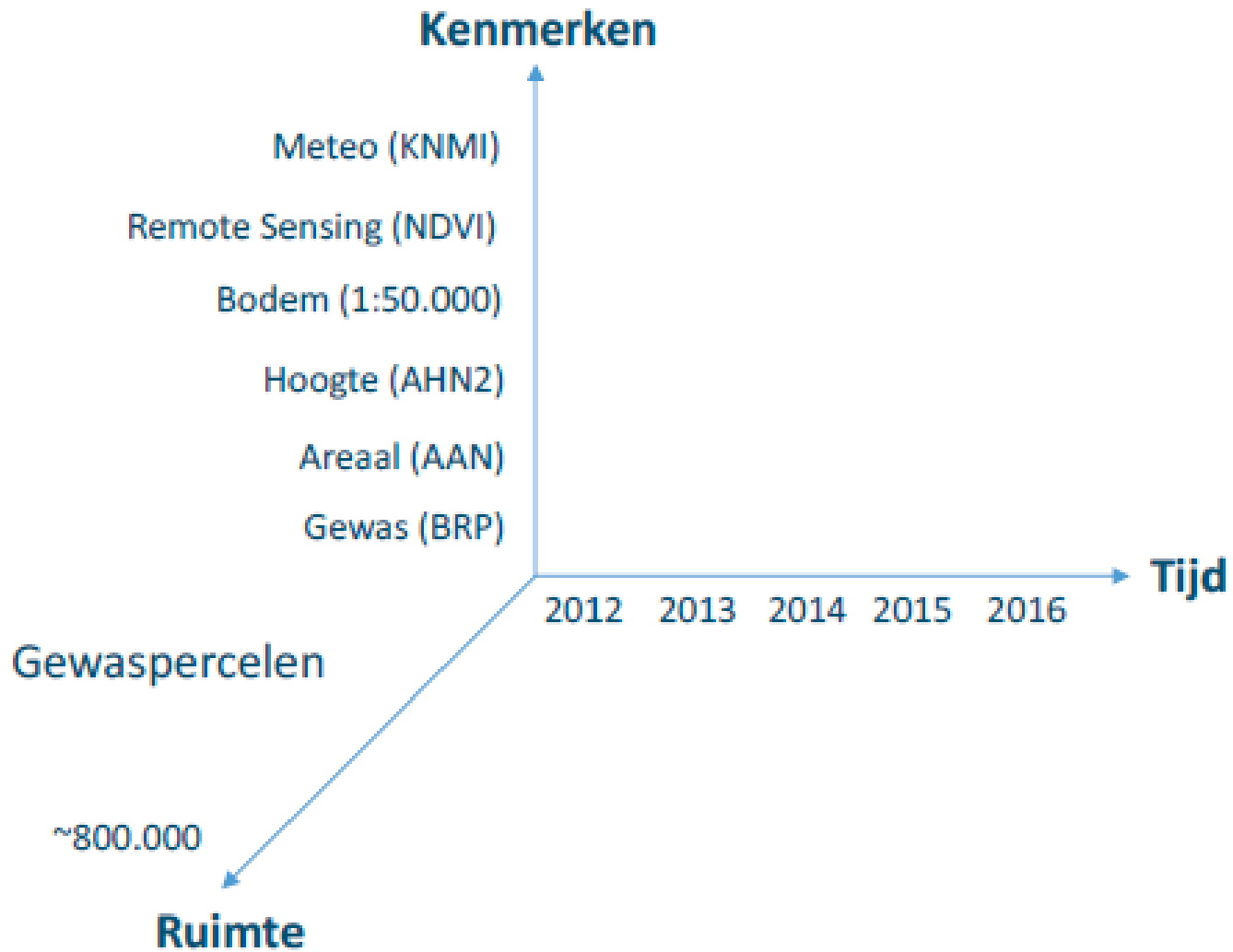
Warehouse



Wrangling

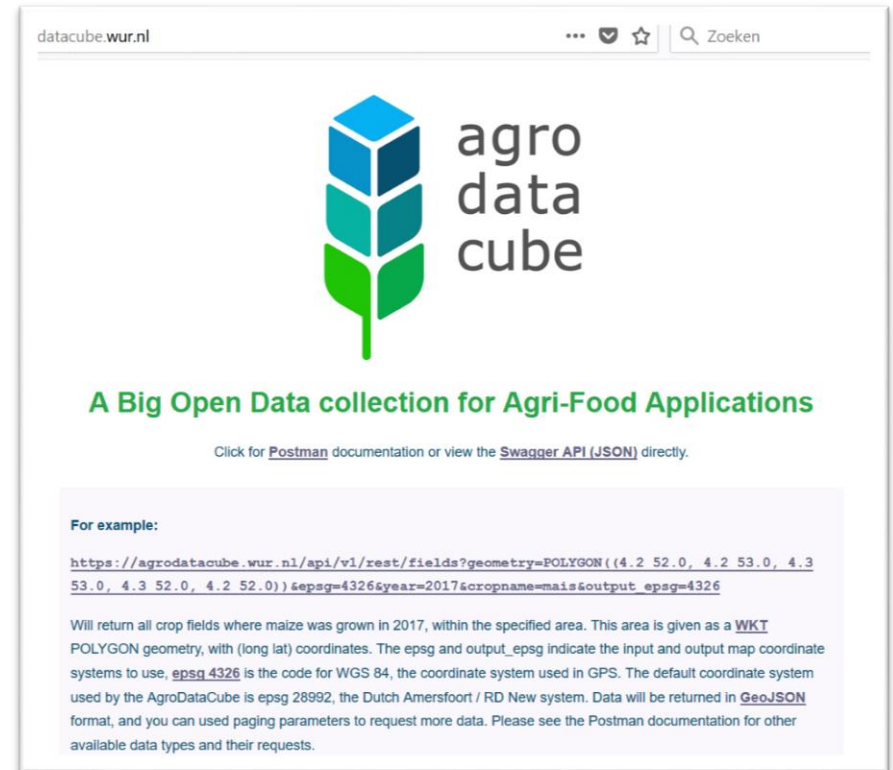


Gebruiken



# Technology Perspective AgroDataCube

- Brings together many open data sources available in the Netherlands;
- has consistent time series and geographical units from 2012 to the current year;
- is available through an open source API;
- with Postman documentation, meta-data and shared coding examples on GitHub.



datacube.wur.nl

... 🔔 ☆ 🔍 Zoeken



agro  
data  
cube

**A Big Open Data collection for Agri-Food Applications**

Click for [Postman](#) documentation or view the [Swagger API \(JSON\)](#) directly.

**For example:**

```
https://agrodatacube.wur.nl/api/v1/rest/fields?geometry=POLYGON((4.2 52.0, 4.2 53.0, 4.3 53.0, 4.3 52.0, 4.2 52.0))&epsg=4326&year=2017&cropname=mais&output_epsg=4326
```

Will return all crop fields where maize was grown in 2017, within the specified area. This area is given as a [WKT](#) POLYGON geometry, with (long lat) coordinates. The epsg and output\_epsg indicate the input and output map coordinate systems to use, [epsg 4326](#) is the code for WGS 84, the coordinate system used in GPS. The default coordinate system used by the AgroDataCube is epsg 28992, the Dutch Amersfoort / RD New system. Data will be returned in [GeoJSON](#) format, and you can use paging parameters to request more data. Please see the Postman documentation for other available data types and their requests.

# Community Perspective AgroDataCube

- Activates the community of developers to further improve it as a collective public good;
- has been used in several hackathons and development sprints as common reference point on data;
- has achieved over 900 users in 10 months;
- is used by private sector, governmental bodies and research projects.

lots-of-data

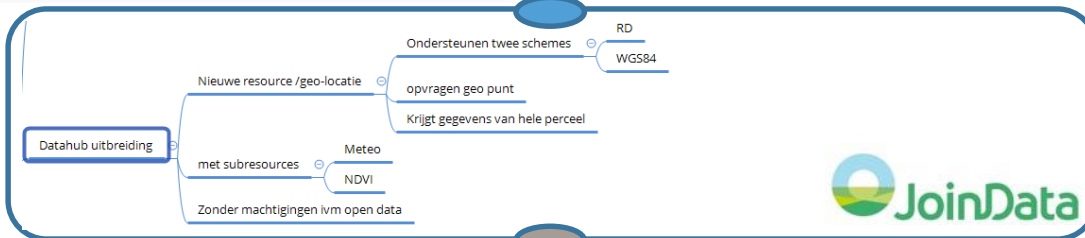


akkerweb

Uitbreiding User Interface Akkerweb

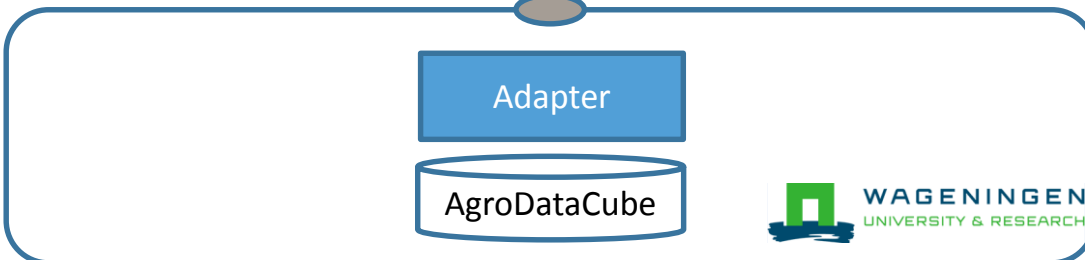
```
GET /geo-locations/{geo-scheme}/{geo-ld}/meteo Get meteo information
GET /geo-locations/{geo-scheme}/{geo-ld}/ndvi Get Normalized Difference Vegetation Index information
```

Uitbreiding koppelvlak en geïntegreerd



Uitbreiding DataHub

Uitbreiding koppelvlak en geïntegreerd



Adapter geschreven op AgroDataCube



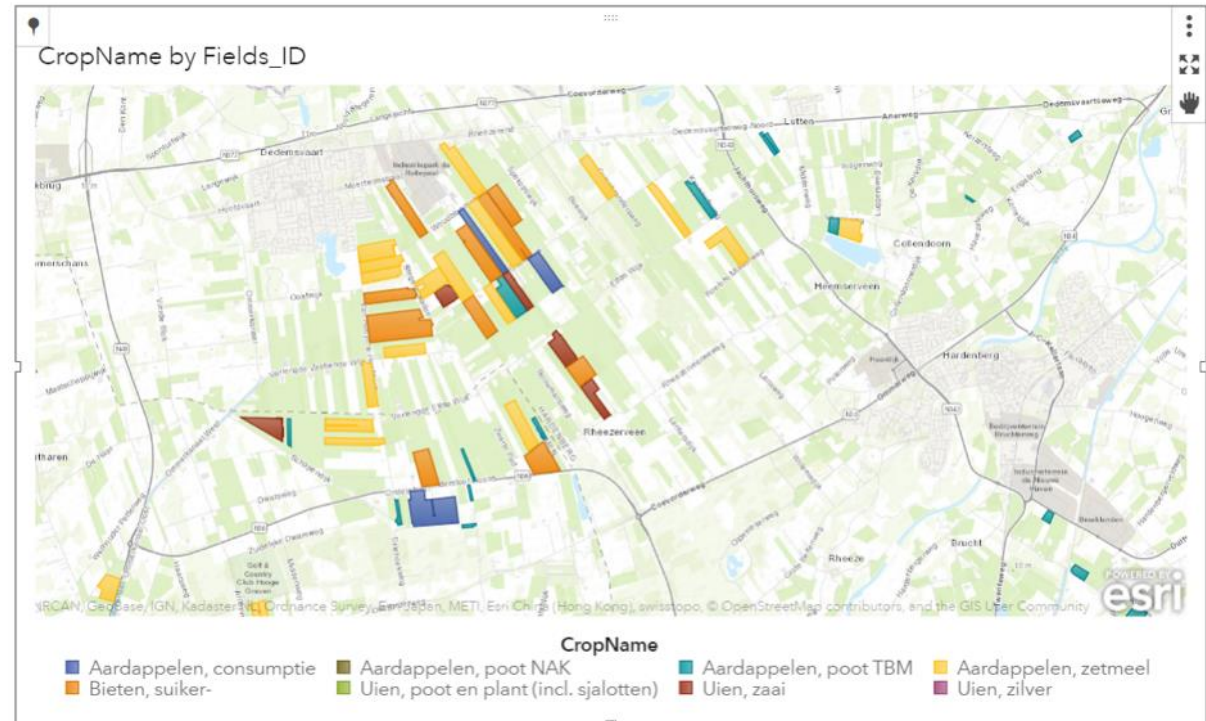
# Crop Emergence Analysis with SAS

In this project, we aim to find relevant factors effecting crop emergence for onions, sugar-beets and potatoes in the Netherlands covering multiple years using advanced analytics techniques such as Random Forest and Gradient Boosting.

With AgroDataCube, relevant and rich information about multiple ten thousands of fields can be accessed and crossed with:

- Soil information
- Meteorology data
- Information from Actueel Hoogtebestand data

to infer non linear and linear patterns effecting crop emergence for these crops.

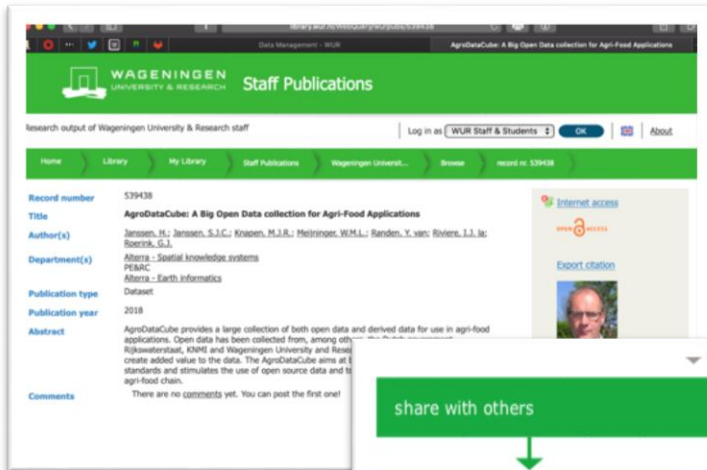


# WUR Data Management Support

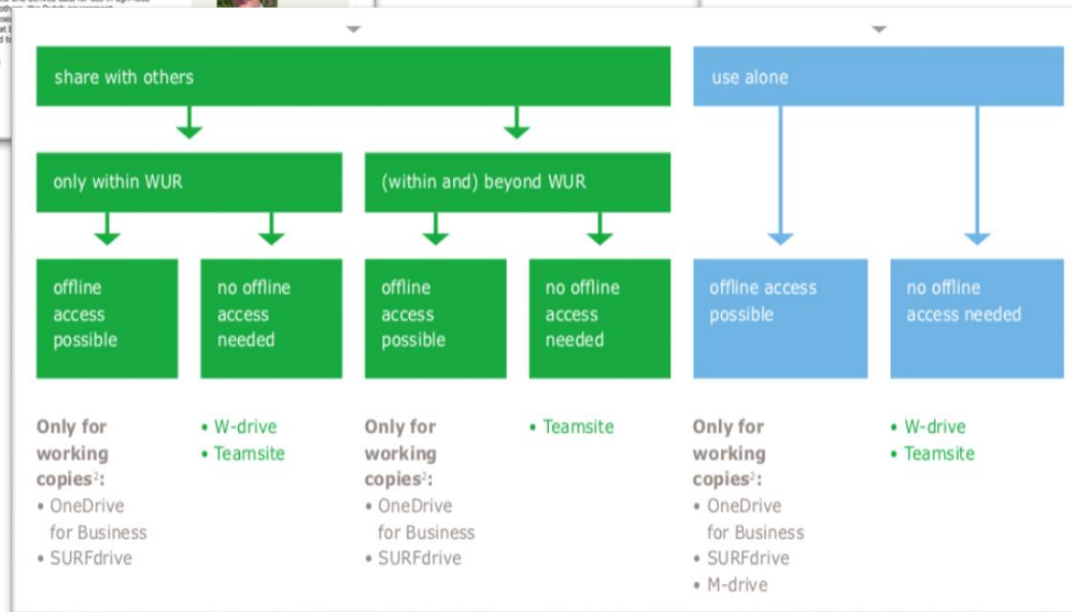
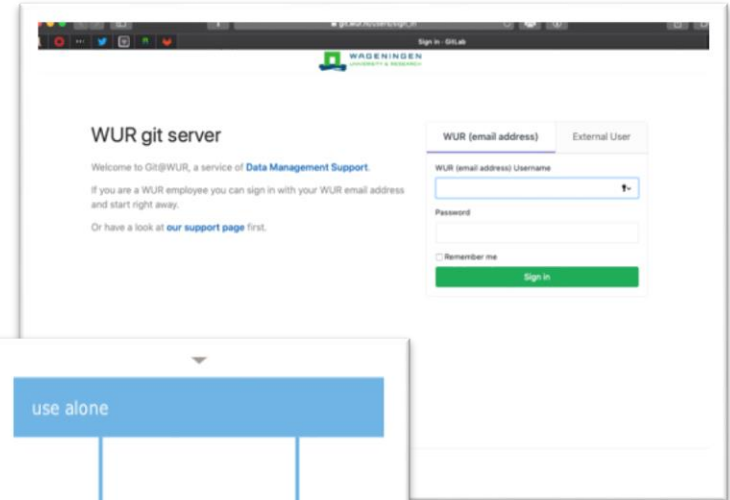
The screenshot displays the WUR Data Management website. The top navigation bar includes 'Home', 'Value Creation & Cooperation', 'WCCC', and 'Data Management'. The main content area features a 'Data Management' heading, a brief introduction, and a 'What is Research Data Management?' section. A sidebar on the right contains a search bar and a list of links under the heading 'Data Management'. The list includes: 'formats', 'Funders and data management costs', 'Exchanging data', 'File formats', 'High Performance Computing', 'Manage your source code with Git@WUR', 'Organising files and folders', 'Storage solutions', 'Finishing your research', 'Data licenses', 'Journal requirements', 'When and why to publish your dataset', 'Publishing your dataset in a repository', and 'Data policy at WUR', which includes 'Data Management Plans', 'Data storage during research', 'Data archiving after research', and 'Registering archived datasets'. The WUR logo and 'WUR is serious about Data' tagline are also visible.

# WUR Data Management Support

## Digital Object Identifier (DOI)



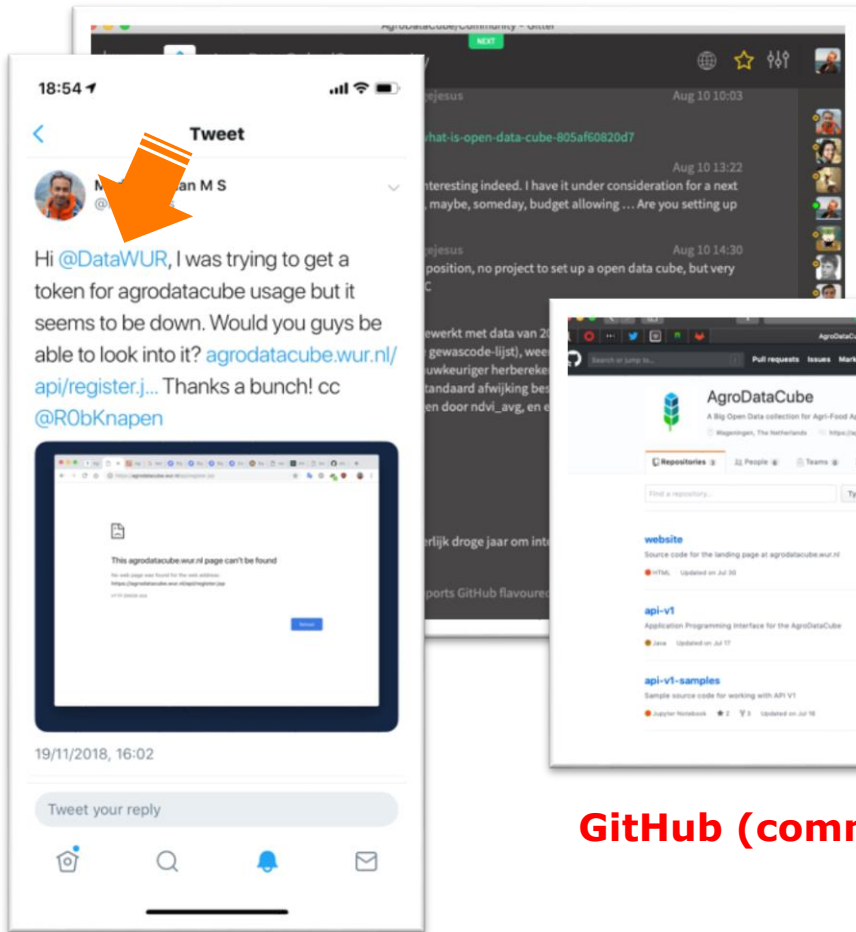
## GitLab



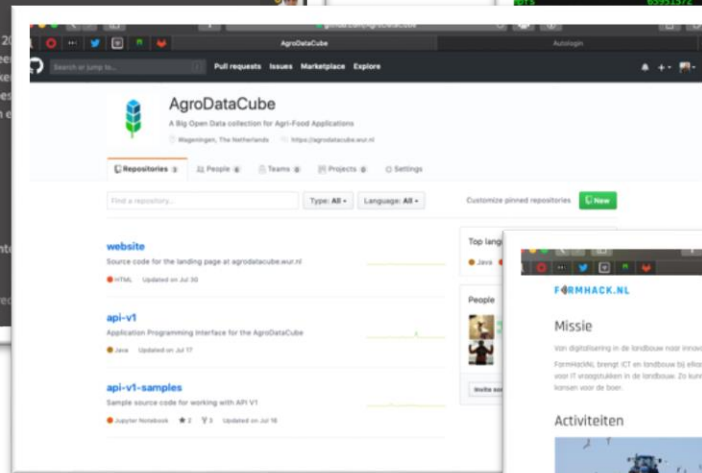
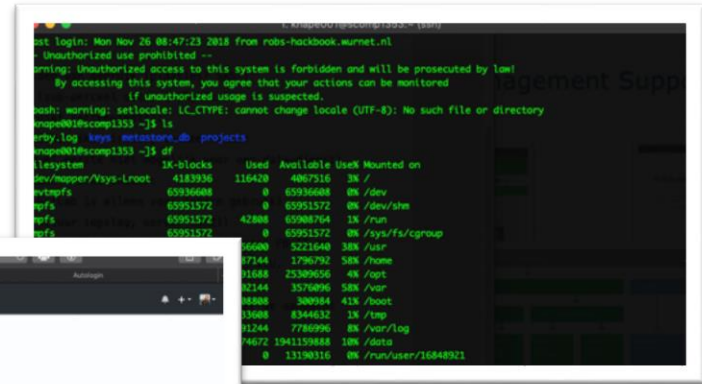
**(Typical) Research Data Storage**

# Do-It-Yourself (DIY)

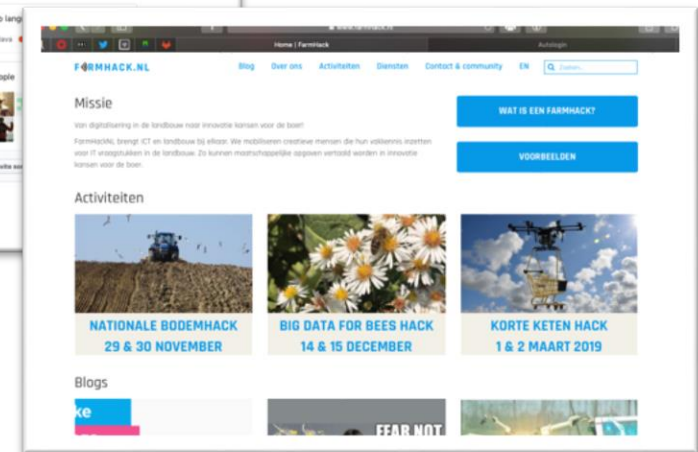
## Server, storage management (FB-IT)



## Support



## GitHub (community)



## Hackathons

# Lessons Learned / Recommendations

- Be practical about FAIR, consider it beyond the scientific domain. Ontology engineering can take forever.
- Focus on publishing data (including algorithms), even in a less perfect way, so that reuse can start. Only then the practical problems will become visible and can be addressed (iteratively).
- The benefits of FAIR are only achieved when data is actually reused by others. That is when innovation can start. It requires more than data storage and publishing. For example: updates, APIs, communities, licensing.

# Thank you!

[rob.knapen@wur.nl](mailto:rob.knapen@wur.nl)

[Sander.janssen@wur.nl](mailto:Sander.janssen@wur.nl)



@wurcgi

