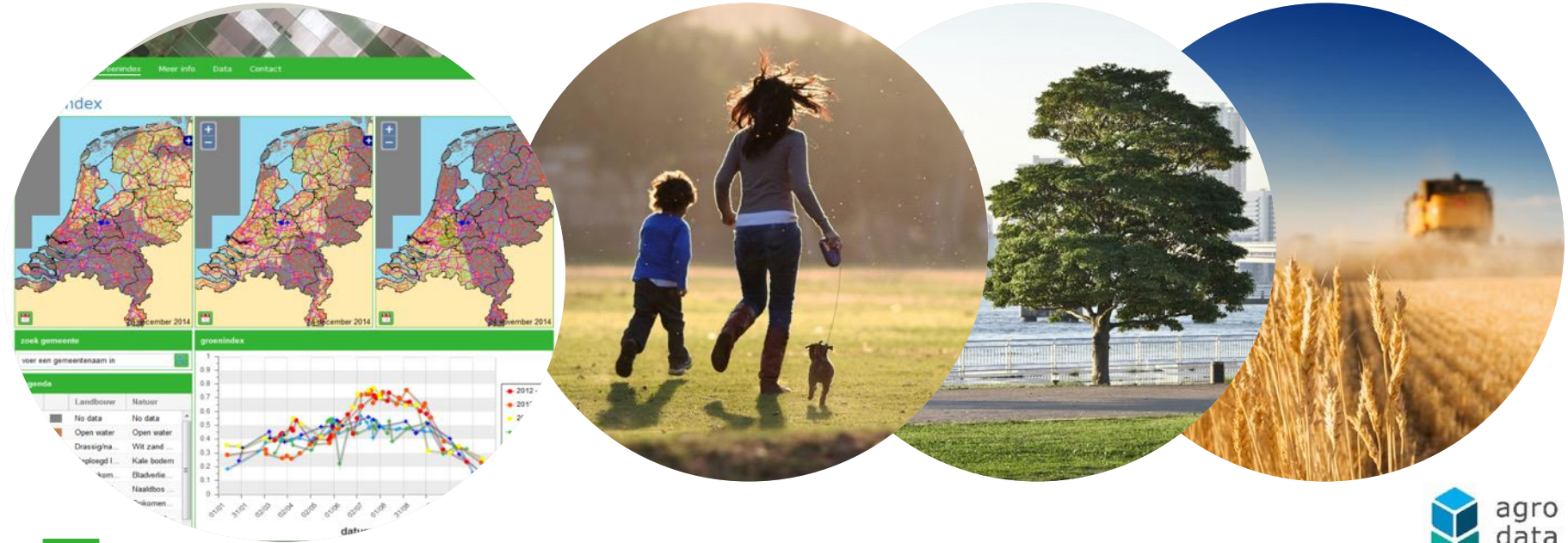
AgroDataCube: A Big Open Data Collection for Agri-Food Applications

RDNL, UKB werkgroep 28 November 2018

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



WAGENINGEN
UNIVERSITY & RESEARCH

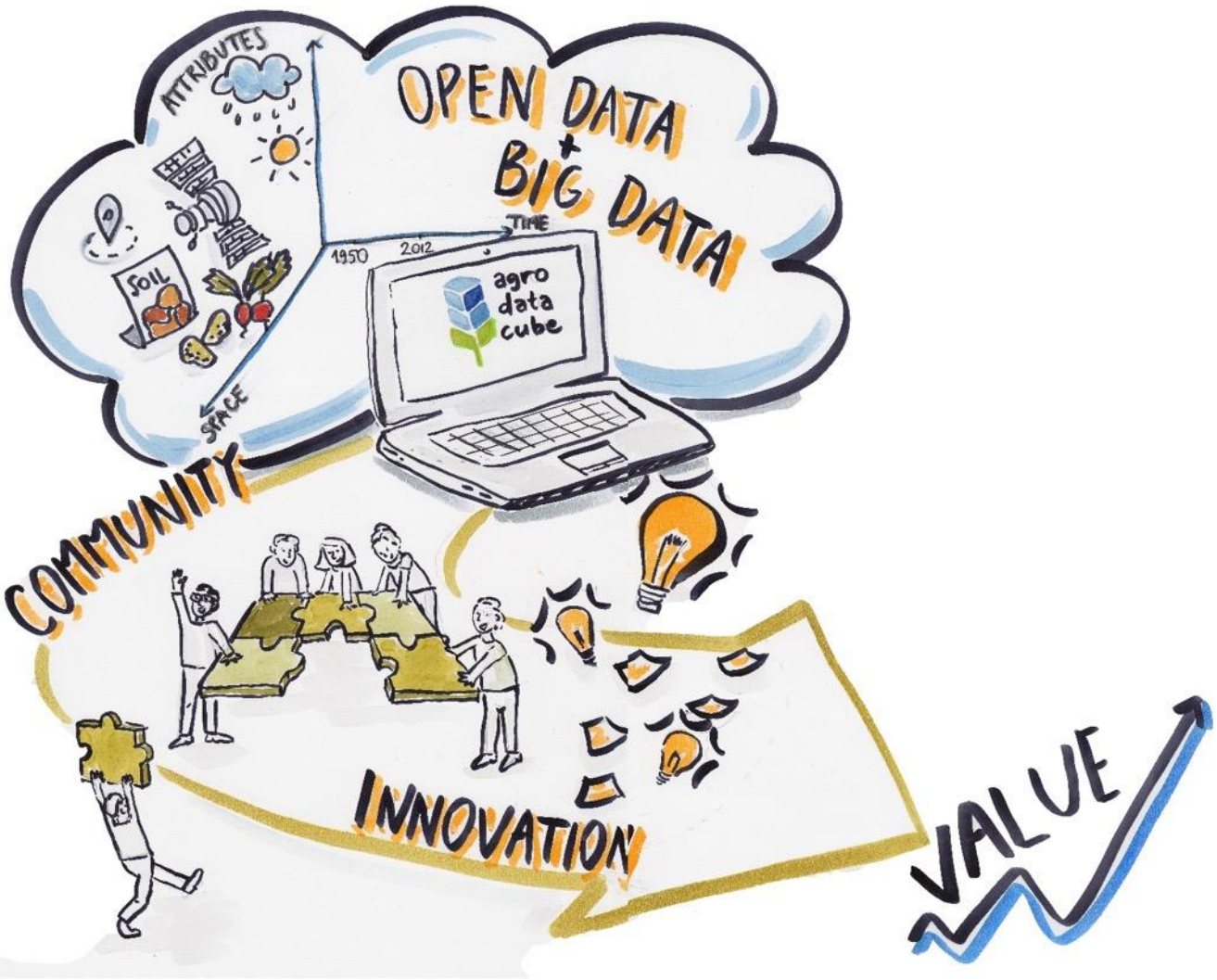
agro
data
cube

Contents

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

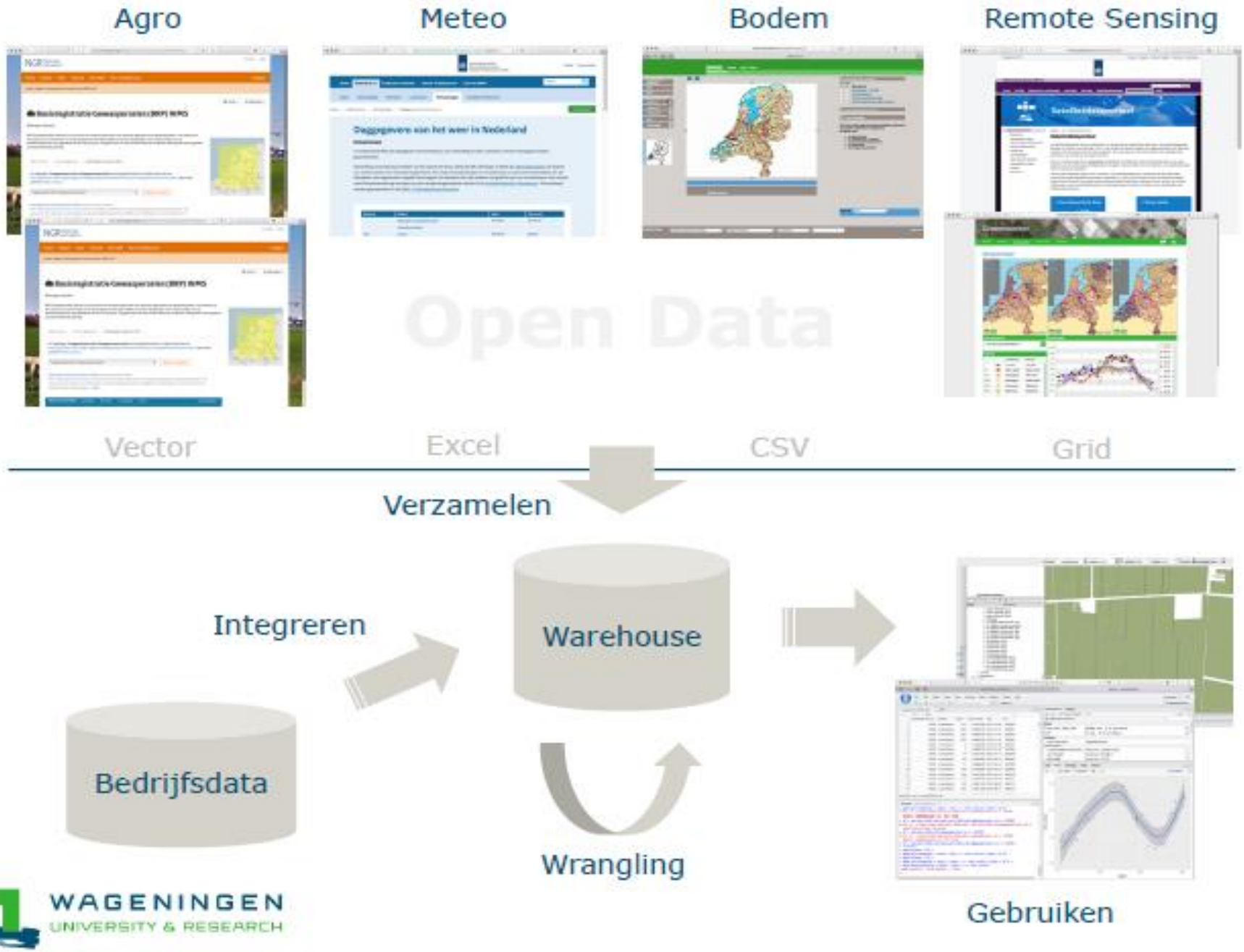
WAGENINGEN
UNIVERSITY & RESEARCH

agro
data
cube

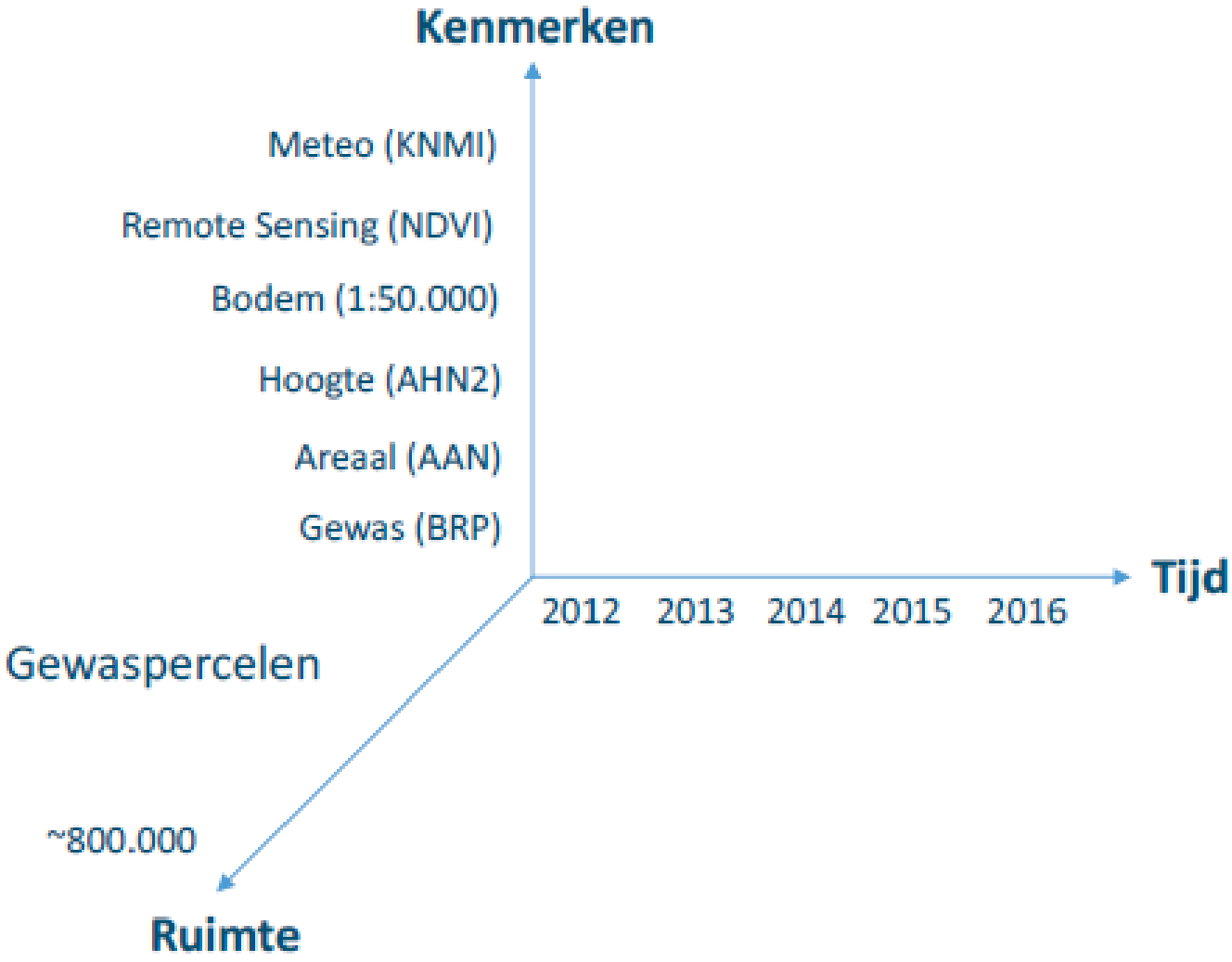


WAGENINGEN
UNIVERSITY & RESEARCH

agro
data
cube



WAGENINGEN
UNIVERSITY & RESEARCH

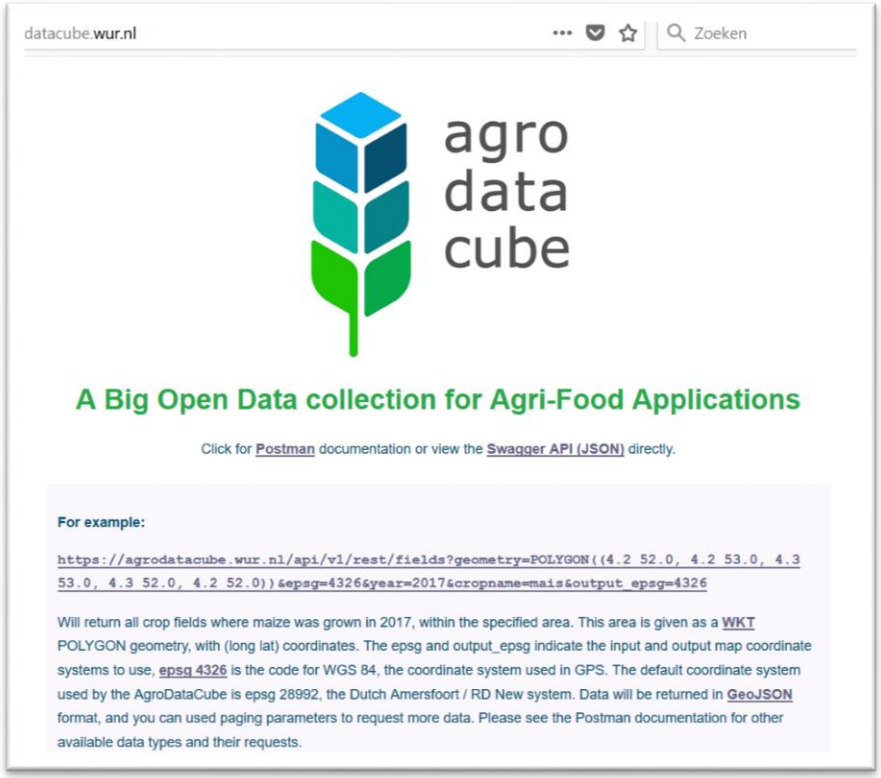


WAGENINGEN
UNIVERSITY & RESEARCH

agro
data
cube

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.



WAGENINGEN
UNIVERSITY & RESEARCH

agro
data
cube

Community Perspective AgroDataCube

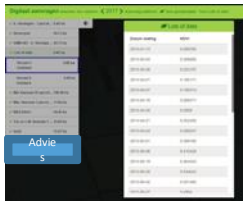
- Activates the community of developers to further improve it as a collective public good;
- has been used in several hackathons and development sprints as common reference point on data;
- has achieved over 900 users in 10 months;
- is used by private sector, governmental bodies and research projects.

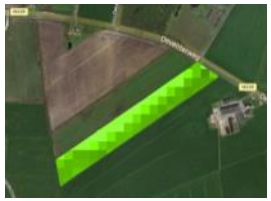
WAGENINGEN
UNIVERSITY & RESEARCH

agro
data
cube

lots-of-data

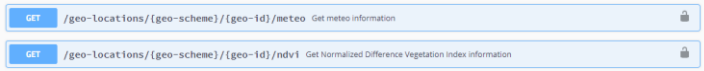
JoinData
The Future of Smart Farming



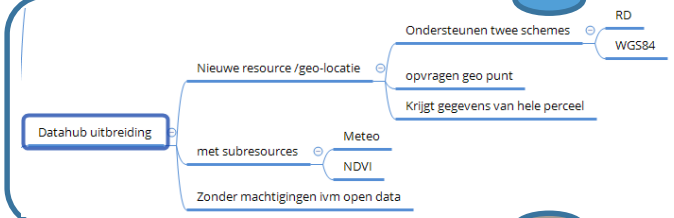


akkerweb.

Uitbreiding User Interface Akkerweb



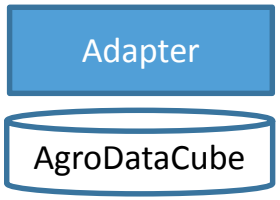
Uitbreiding koppelvlak en geïntegreerd



JoinData

Uitbreiding DataHub

Uitbreiding koppelvlak en geïntegreerd



WAGENINGEN
UNIVERSITY & RESEARCH

Adapter geschreven op AgroDataCube

WAGENINGEN
UNIVERSITY & RESEARCH

agro
data
cube

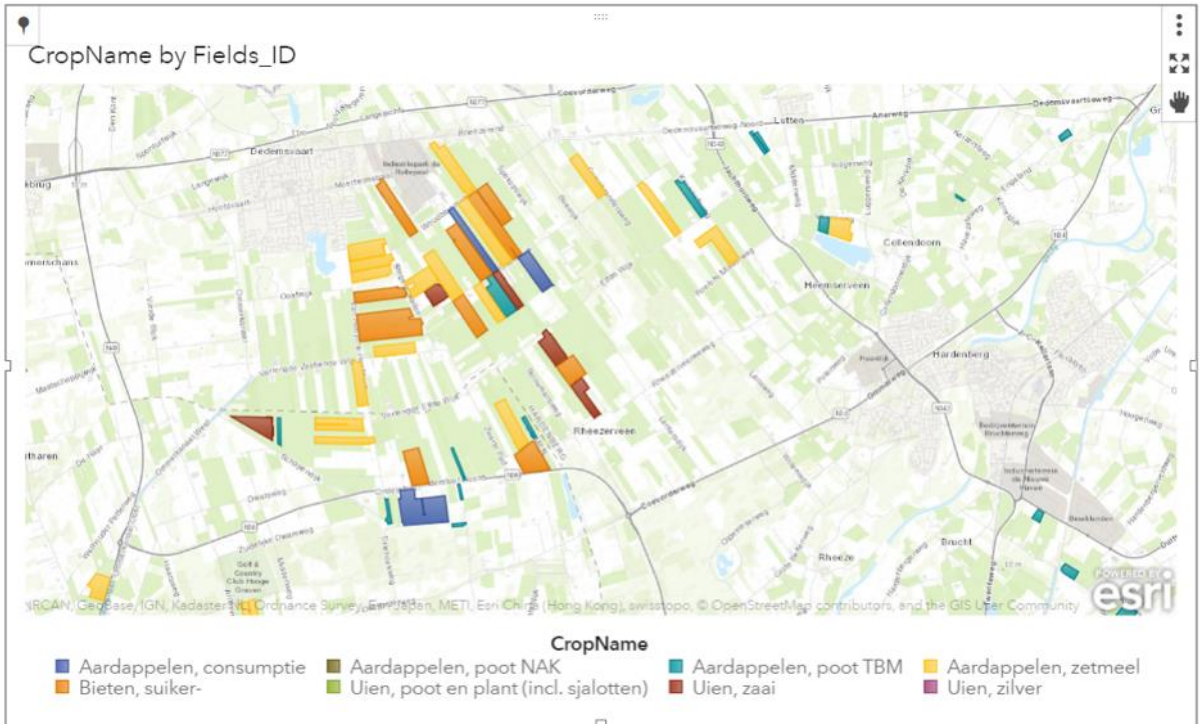
Crop Emergence Analysis with SAS

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.

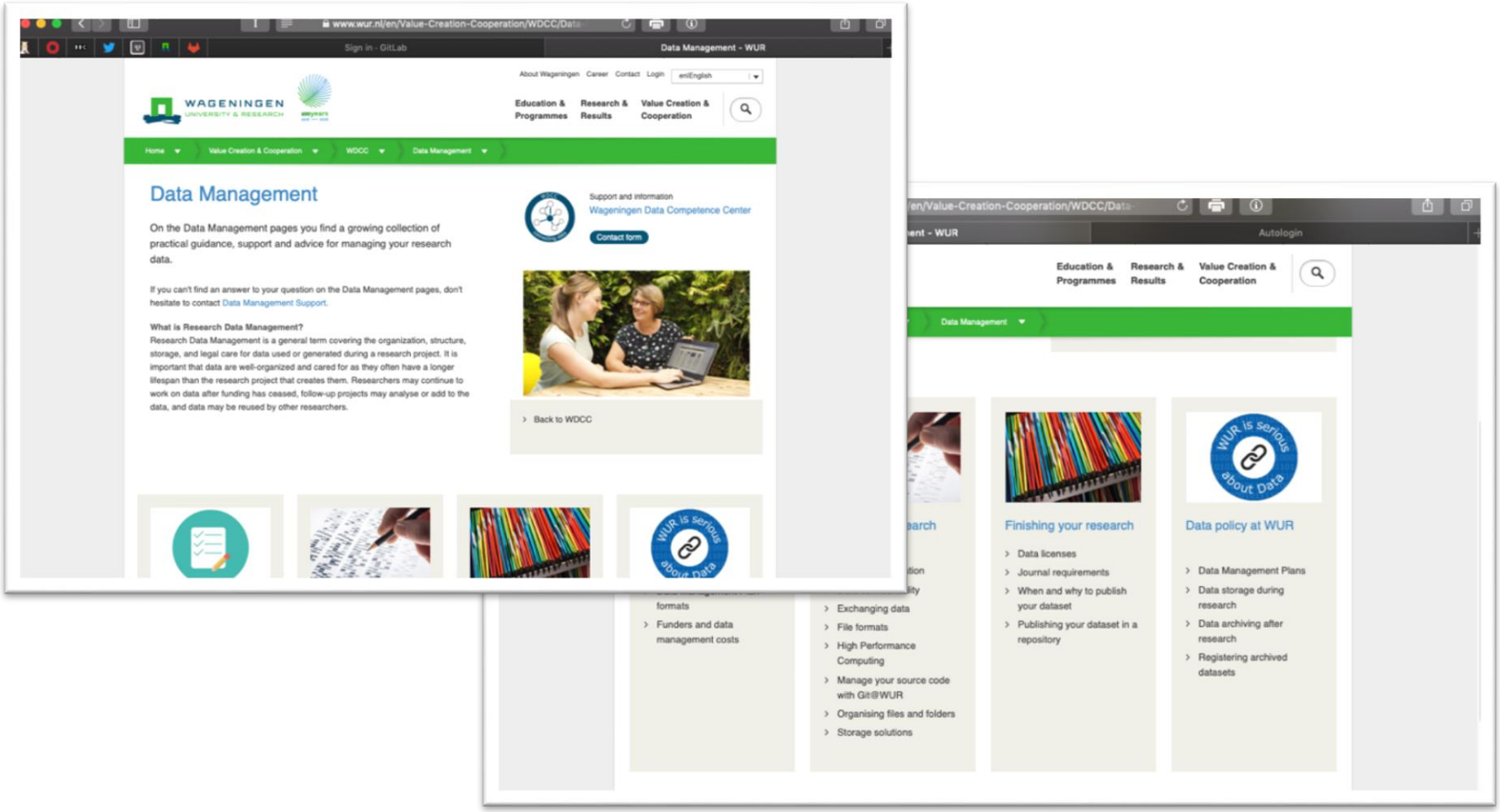
With AgroDataCube, relevant and rich information about multiple ten thousands of fields can be accessed and crossed with:

- 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

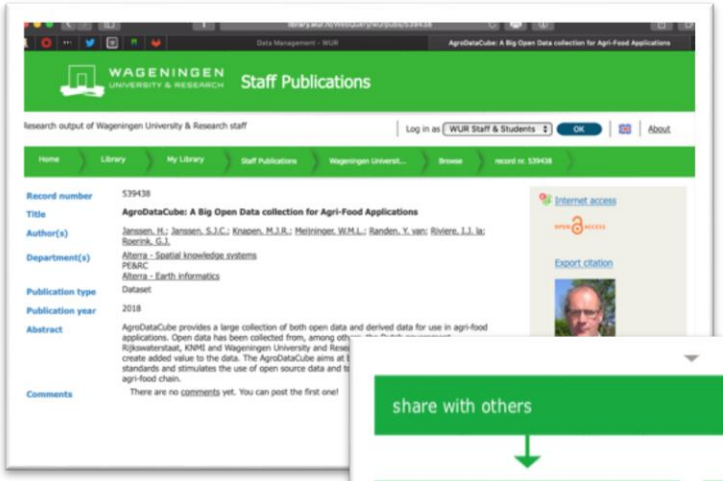


WAGENINGEN
UNIVERSITY & RESEARCH

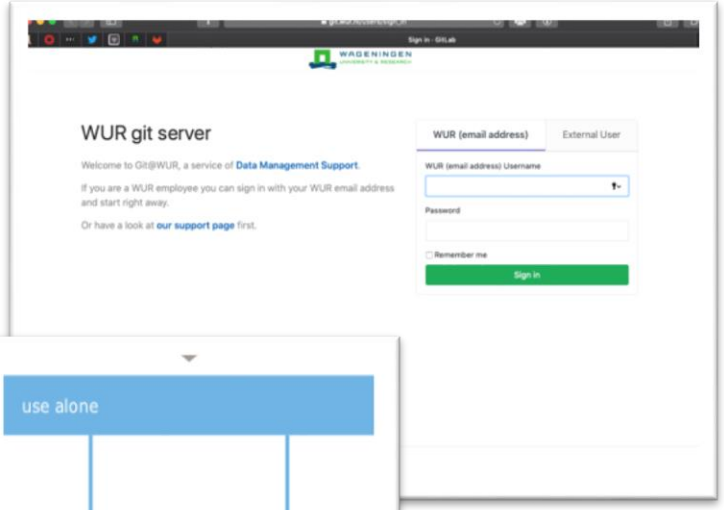
agro data cube

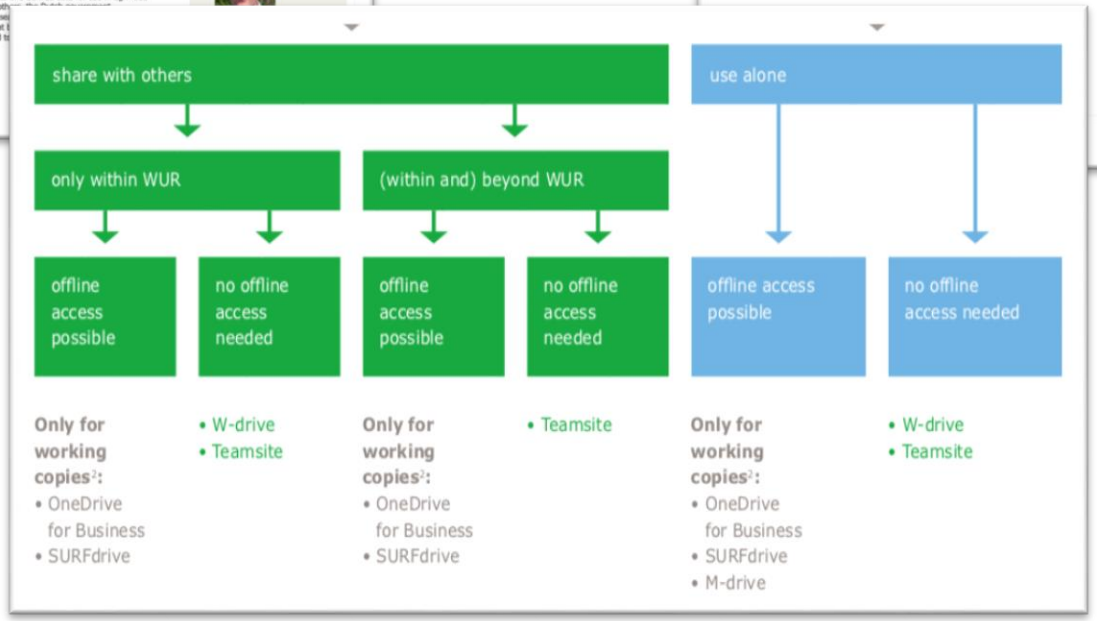
WUR Data Management Support

Digital Object Identifier (DOI)



GitLab



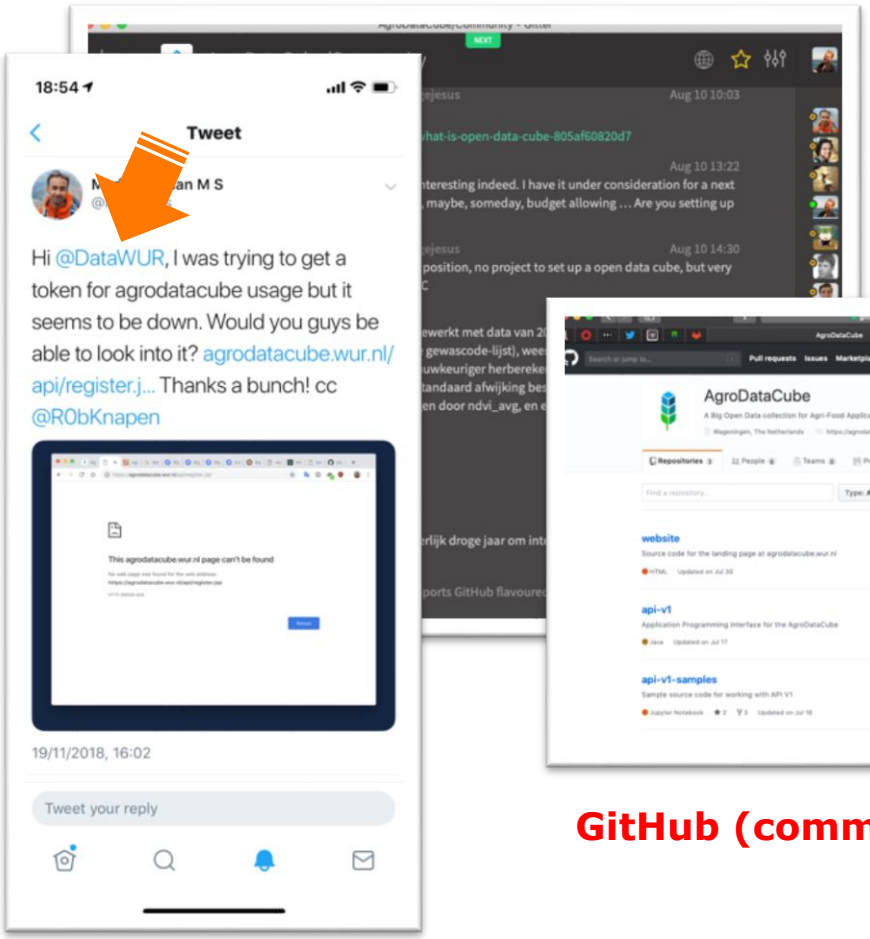


(Typical) Research Data Storage

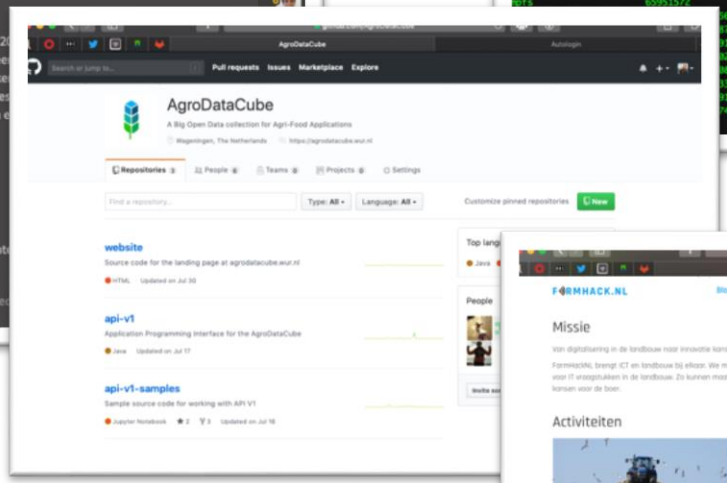
WAGENINGEN
UNIVERSITY & RESEARCH

agro data cube

Do-It-Yourself (DIY)

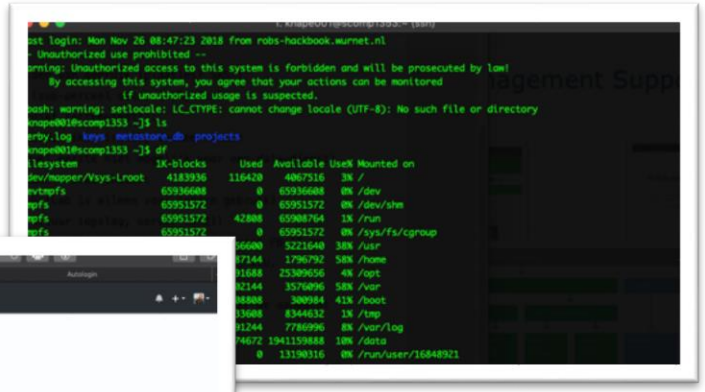


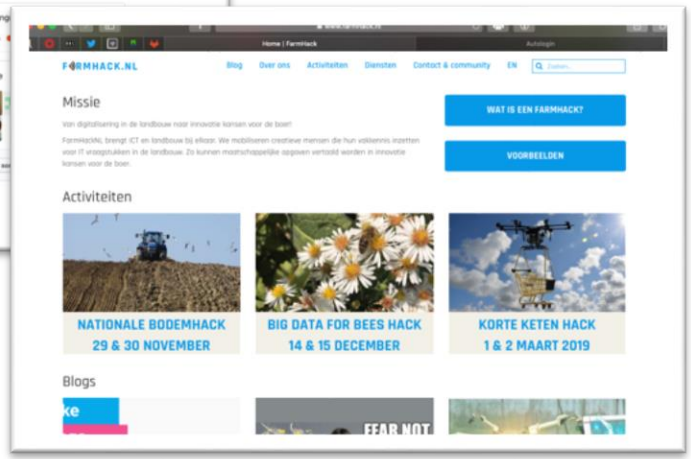
Support



GitHub (community)

Server, storage management (FB-IT)





Hackathons

WAGENINGEN
UNIVERSITY & RESEARCH

agro data cube

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

Sander.janssen@wur.nl

@wurcgi

WAGENINGEN
UNIVERSITY & RESEARCH

agro
data
cube