

# Data management in CocoaSoils – ontology and links to EiA data

---

Arun Pratihast, Ekatherina  
Vasquez, Sander Janssen, Eloi  
Ribeiro and Hugo de Groot

---

Annual CocoaSoils Forum 2024  
Abidjan, Ivory Coast



# Data challenges in Cocoa Sector

- Many of Cocoa experimental datasets ended up in separate repositories in various formats often with **missing headings** and **contextual information**
- Lacks data **sharing protocols**, and collaborative re-use of data
- Creates barriers in data discovery, gathering, managing and harmonizing from different sources for better decision making
- Urgent need for FAIR data solutions to improve Cocoa economic, social and environmental sustainability

# Objective

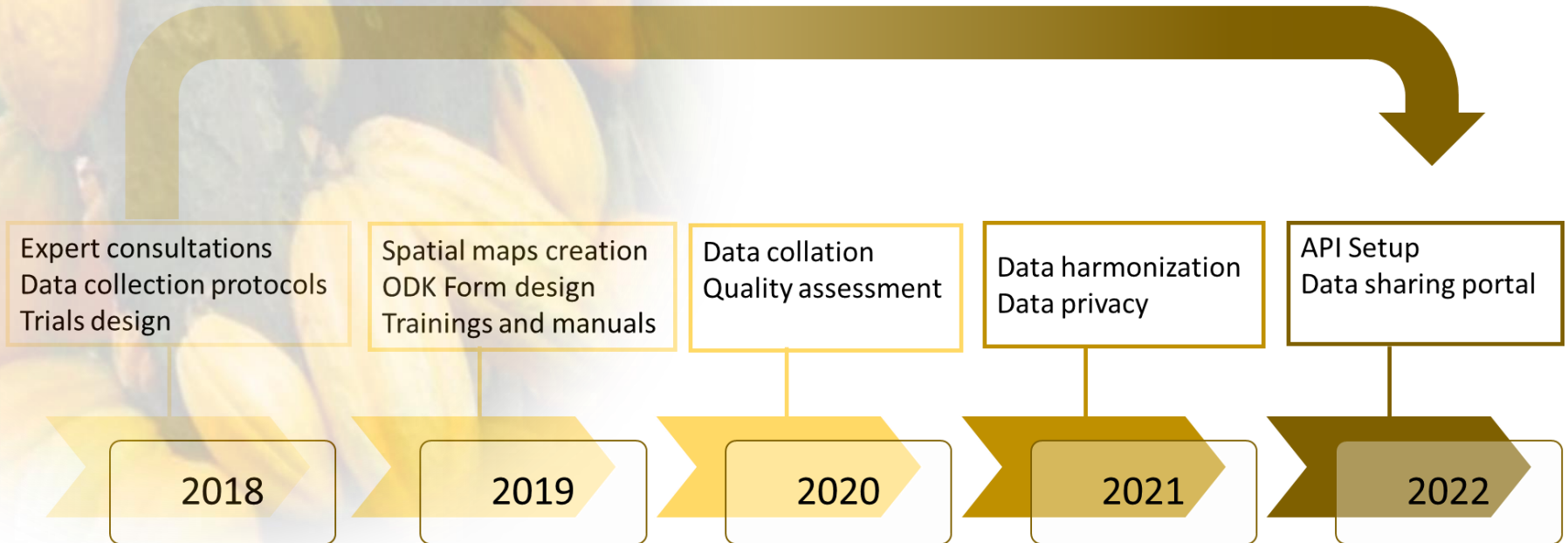
- The main objective is to develop digital solutions by combining local knowledge and mobile technologies, to support cocoa productivity and human well-being

## **Specific objectives:**

- To design an integrated mobile based data collection, transmission and visualization system
- To standardize vocabulary, concepts and relations between cocoa agronomy and production processes through cocoa ontology
- To implement and operationalize the developed solutions in cocoa production site of Cote d'Ivoire, Ghana, Nigeria, and Cameroon, Indonesia and Ecuador

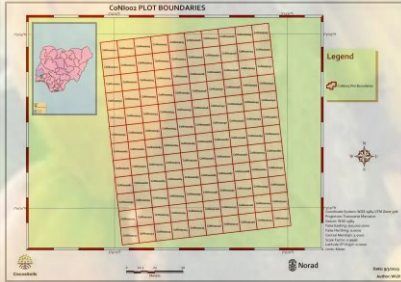
# CocoaSoils data Journey

## Cocoa ontology development



# Trials set up

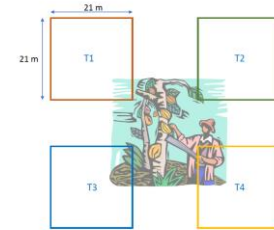
## CORE TRIALS



11 Long-term nutrient response experiments



## SATELLITE TRIALS



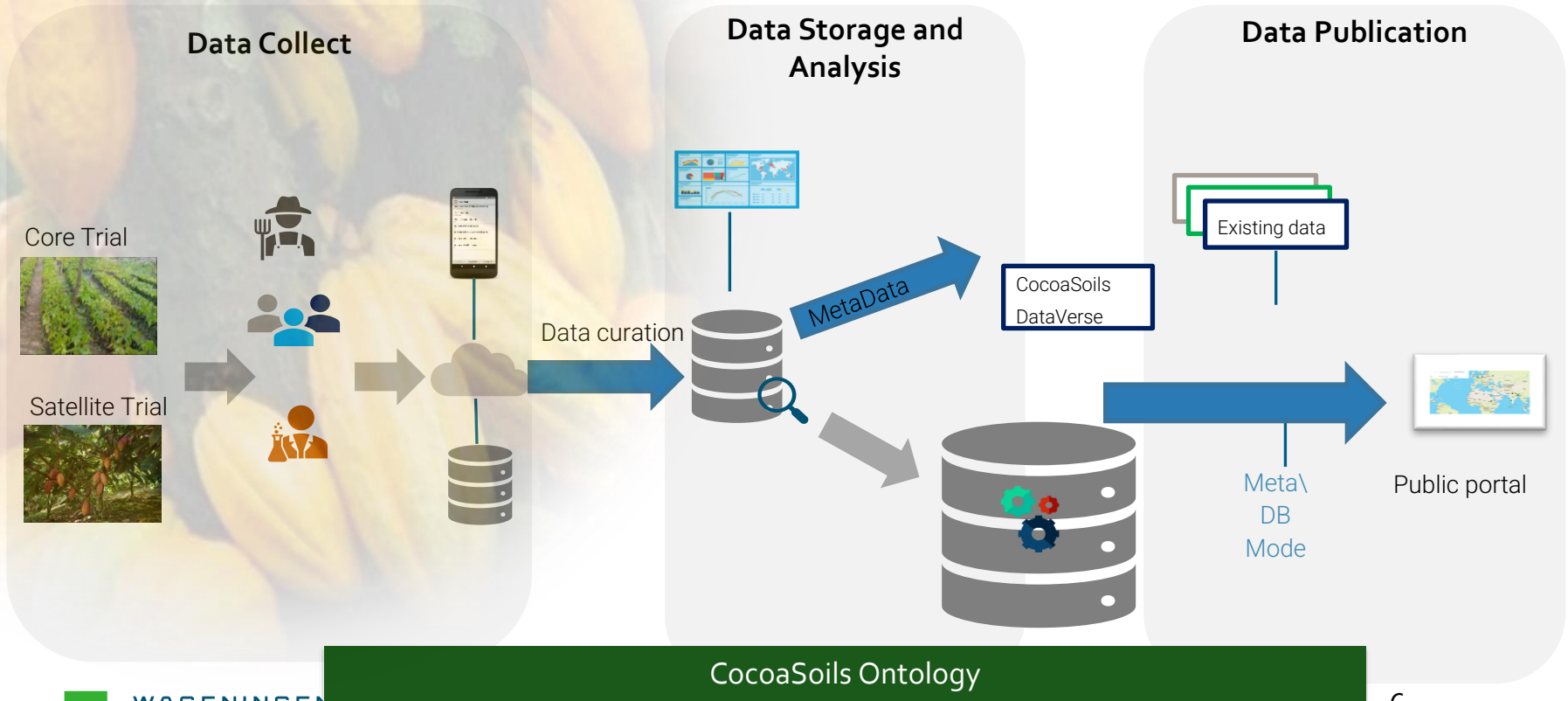
393 Farmer-managed sites  
4 countries: Ghana, Nigeria, Cameroon, and Ivory Coast  
T1: C+ I  
T2: BMP + no F  
T3: BMP + NF  
T4: BMP + OF

Effects of:

- Multifactorial fertiliser combinations (N-P-K) as part of integrated soil fertility management
- GAP: Shading, pruning, weeding, pest control.

**Aim: To evaluate the impact of optimized cocoa tree management and nutrient supply.**

# Ecosystem of data services

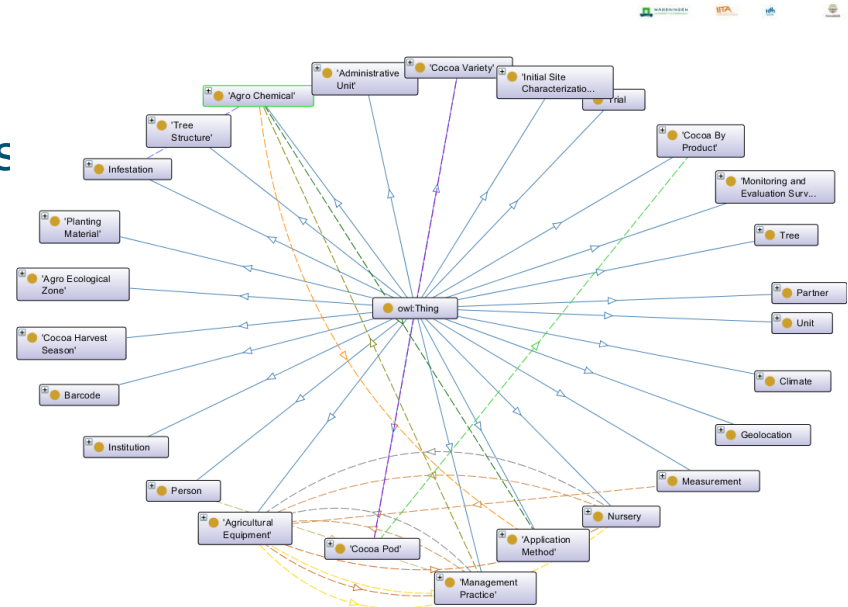


# Semantic enrichment- Cocoa Ontology



January 2022

- Cocoa Ontology designed by all and accepted by all and used by all.
- Define, describe, and standardize vocabulary, concepts and relations between cocoa farming system and components;
- Interoperability of CocoaSoils datasets, in order to enable value addition and data-driven innovation in cocoa research.

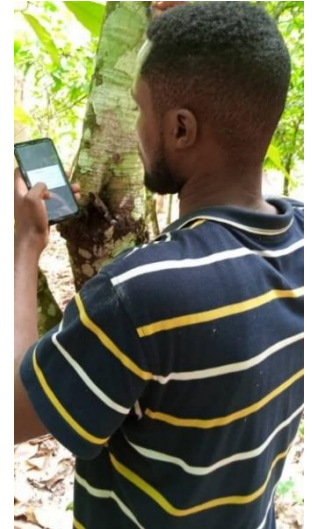
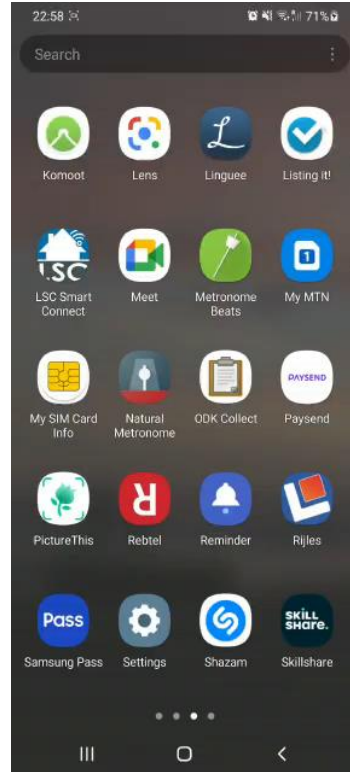


# Data Collection



XLS

	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X	Y	Z
1																										
2																										
3																										
4																										
5																										
6																										
7																										
8																										
9																										
10																										
11																										
12																										
13																										
14																										
15																										
16																										
17																										
18																										
19																										
20																										
21																										
22																										
23																										
24																										
25																										
26																										
27																										
28																										
29																										
30																										
31																										
32																										
33																										
34																										
35																										
36																										
37																										
38																										
39																										
40																										
41																										
42																										
43																										
44																										
45																										
46																										
47																										
48																										
49																										
50																										






# ODK data server

← → ↻ 🏠 [cocoasoils.wur.nl/ODKAggregate/Aggregate.html#submissions/filter/](https://cocoasoils.wur.nl/ODKAggregate/Aggregate.html#submissions/filter/) ⌵ ☆ ⚙️ 🖨️ 🌐

Apps → Editorial Manager®... SCB - Sports overvi... Free English Vocab... FAORStudio Open Data Kit > XL... AKVO FileSender Beta Remote Sensing an... Jobseeker | Mailbox... NGI Elephant riding and... Intranet Wageninge... IELTS >> 📁 Other bookma

 Submissions Form Management Site Admin [Log Out pratihasta](#) ? 📄 📧


































Filter Submissions Exported Submissions

Form  Filter

Submissions per page

Display Metadata

**Images\_of\_st\_plot\_validation**

Introduction intro	farmer_plot_info farmers_name	farmer_plot_info plot_name	farmer_plot_info country	farmer_plot_info name community	farmer_plot_info name partner	farmer_plot_info name enumerator	farmer_plot_info start time user	structural_pruning_on_plot structural_pruning	structural_pruning_on_plot image struct prune	structural_pruning_on_plot image struct prune 2	struct...
	Nde Francis	Kolon	Cameroon	Makenene	OLAM	Nyingchia yvette	February 26, 2020 at 10:15:00 AM +01:00	Ok			
	Yossa Blaise	Kolon	Cameroon	Makenene	OLAM	Nyingchia yvette	February 26, 2020 at 1:14:00 PM +01:00	Partially done			
	Bekala marc	Nkoloutou	Cameroon	Ntui	Olam	Nyingchia yvette	February 28, 2020 at 9:23:00 AM +01:00	Regular			
	Bekala Victor Magloire	Ikopo	Cameroon	Ntui / Nkoloutou	Gic espoir	Nitcha Yvette	February 28, 2020 at 11:38:00 AM +01:00	Fairly done			
	Bilai Etoga Pascal	Natichigal	Cameroon	Ntui	Olam	Nyingchia yvette	February 28, 2020 at 2:38:00 PM +01:00	Ok			
	Noah Mbende Idrissou	Ngutte	Cameroon	Ntui	Olam	Nyingchia yvette	February 29, 2020 at 9:15:00 AM +01:00	Irregular			
	Boko Thomas	Nguite	Cameroon	Ntui	Olam	Nyingchia yvette	February 29, 2020 at 11:54:00 AM +01:00	Partially done			
	BIKEK Blaise	Atout Pajja	Cameroon	Atout	Olam	Etoundi Florent	March 9, 2020 at 3:19:00 PM +01:00	Not done			
	Ndzengue Owona Samuel Arnold	Atout/PAIJA	Cameroon	Atout	Olam	Etoundi Florent	March 9, 2020 at 4:13:00 PM +01:00	Bien fait			
	Akono Abada Landry	Atout/PAIJA	Cameroon	Atout	Olam	Etoundi Florent	March 10, 2020 at 9:20:00 AM +01:00	Partielle			
	EBA Pierre Guy	Atout/Pajja	Cameroon	Atout	Olam	Etoundi Florent	March 10, 2020 at 10:27:00 AM +01:00	Taille partielle			

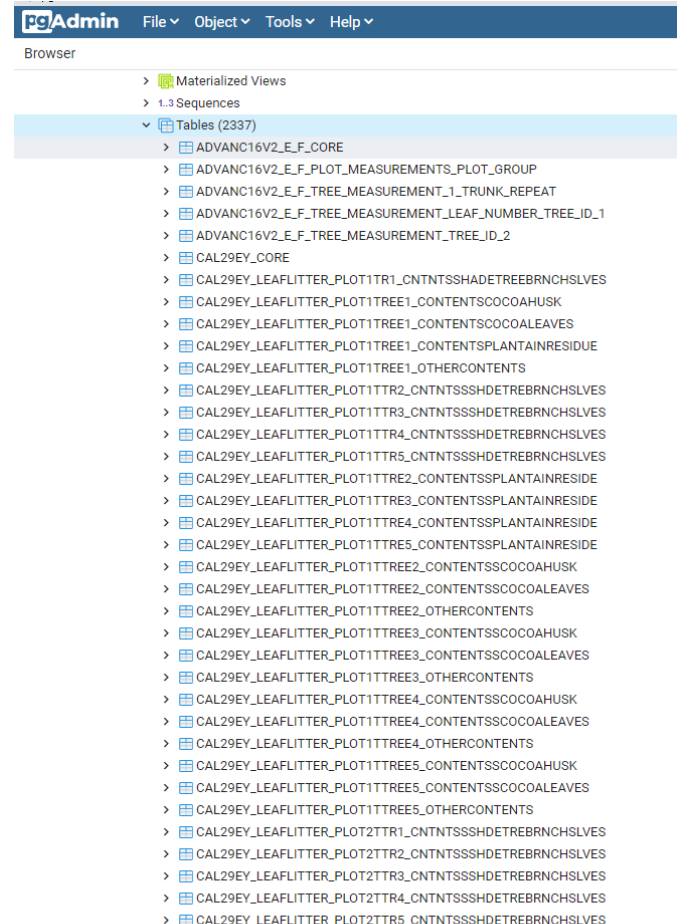
v2.0.3 - Update available

# ODK data harmonization

Multiple data  
collection forms

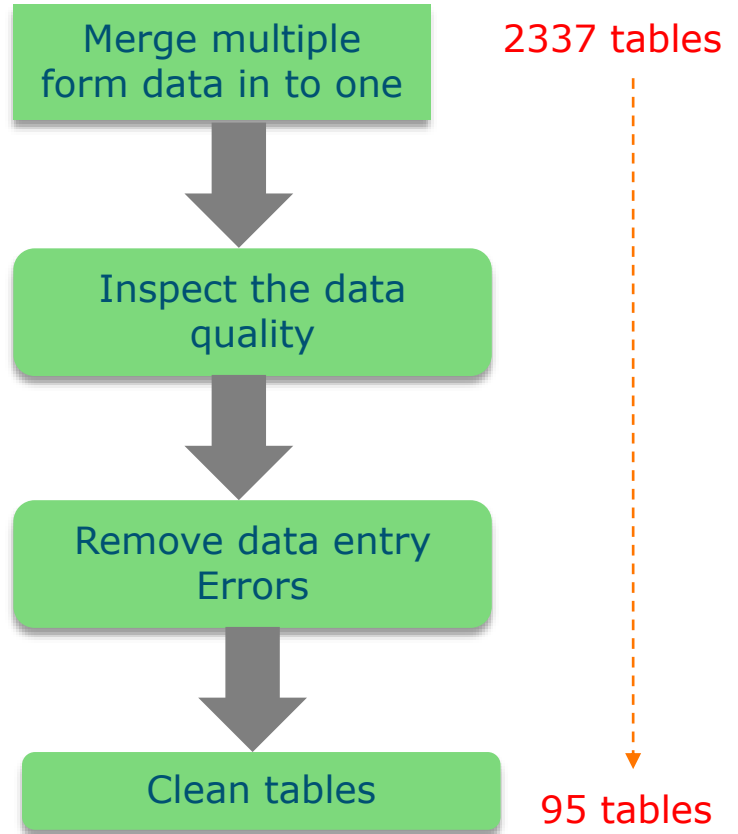


1 forms results 40 tables



# ODK data harmonization

ODK Forms
Advanced_evaluation_form_CT
CocoaSoils_dissemination_events_participants_list_final
CocoaSoils_dissemination_events_registration_final
CocoaSoils_dissemination_events_registration_final-participantDetails
CocoaSoils_dissemination_events_registration_final-trainerDetails
CocoaSoils_dissemination_events_registration_final-repeat_evaluation_criteria_mr
CocoaSoils_dissemination_events_registration_final-trainerDetails
CocoaSoils_EA_BASELINE_TOOL_FINAL
CocoaSoils_EA_BASELINE_TOOL_FINAL_French_Version
CocoaSoils_EA_LEARNING_TOOL
CocoaSoils_EA_LEARNING_TOOL_French_Version
CocoaSoils_Events_Evaluation_Tool
CocoaSoils_Events_Evaluation_Tool_French_Version
early_evaluation_form_CT
Fertilizer_Form_ST_V4
Field maize v19_05
Field preparation v19_05
Field soil samples v19_01
Initial Site Characterization_ST_v2
Management Form_ST_V2
Measurement Form_ST_v2
Nursery v19_05
PICTURE_EVIDENCE_OF_ST_PLOT_VALIDATION
Plantain evaluation v20_05
Planting 19_07
plot_check_c0gh001_V1_3
plot_check_c0ni001_V1_3
plot_check_c0ni002_V1_3
Plot_delineation_ST_v3
Plot_delineation_ST_v4
Plot coordinates v18_01
Plot maize v19_05
Seedling sample v19_05
Subplot maize v19_05
Tree_replacement_ST
VIRTUAL RAPID CHARACTERIZATION FORM

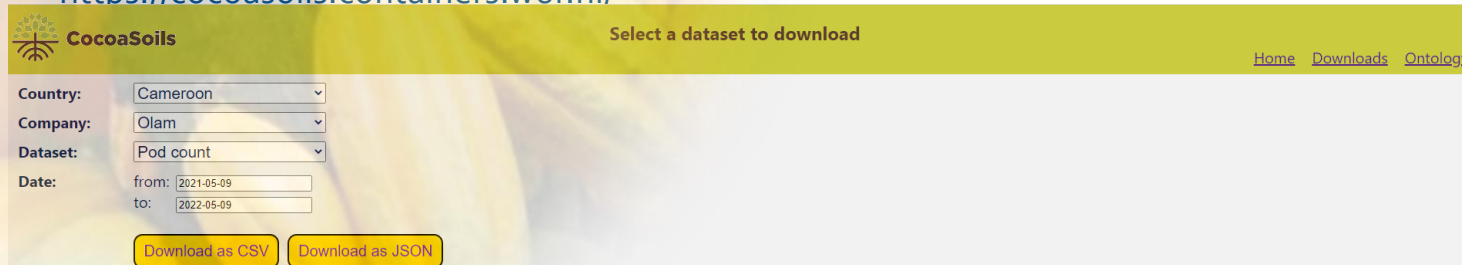


# ODK harmonization



# Data sharing

- Scientific analysis facility: set of automated analytical procedures, API data import, Quality check, export, analysis
- Public and private data filtering facilities
- Data visualization platform - a facility to visualize in an attractive way the data for experts working in cocoa research and industry
- <https://cocoasoils.containers.wur.nl/>



The screenshot shows the CocoaSoils web interface. At the top, there is a green header with the CocoaSoils logo on the left, the text "Select a dataset to download" in the center, and navigation links "Home", "Downloads", and "Ontology" on the right. Below the header, there is a form with the following fields:

- Country:** Cameroon (dropdown menu)
- Company:** Olam (dropdown menu)
- Dataset:** Pod count (dropdown menu)
- Date:** from: 2021-05-09 (text input), to: 2022-05-09 (text input)

At the bottom of the form, there are two yellow buttons: "Download as CSV" and "Download as JSON".

# Data sharing



## Open data & graph

[Home](#) [Open data & graph](#) [Consortium data](#) [Ontology](#)

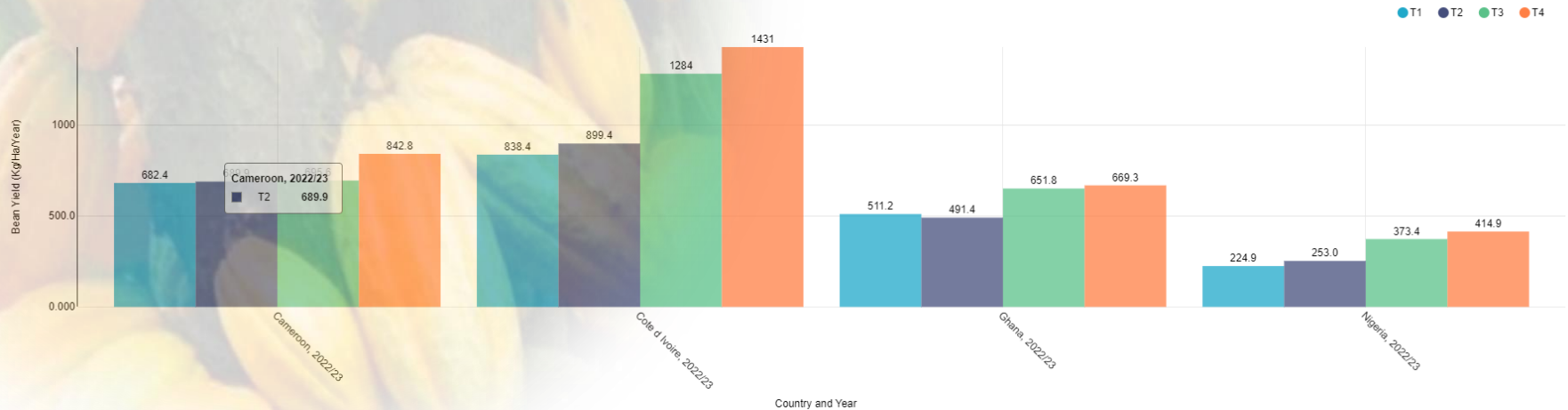
Select season:  2021/22  2022/23

T1 = current farmer practice

T2 = weeding + pruning + insecticide application + fungicide application (no fertiliser application)

T3 = weeding + pruning + insecticide application + fungicide application + current national fertiliser recommendation

T4 = weeding + pruning + insecticide application + fungicide application + offtake model recommendation

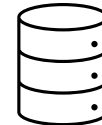


[Download Data as CSV](#)

[Download Data as JSON](#)

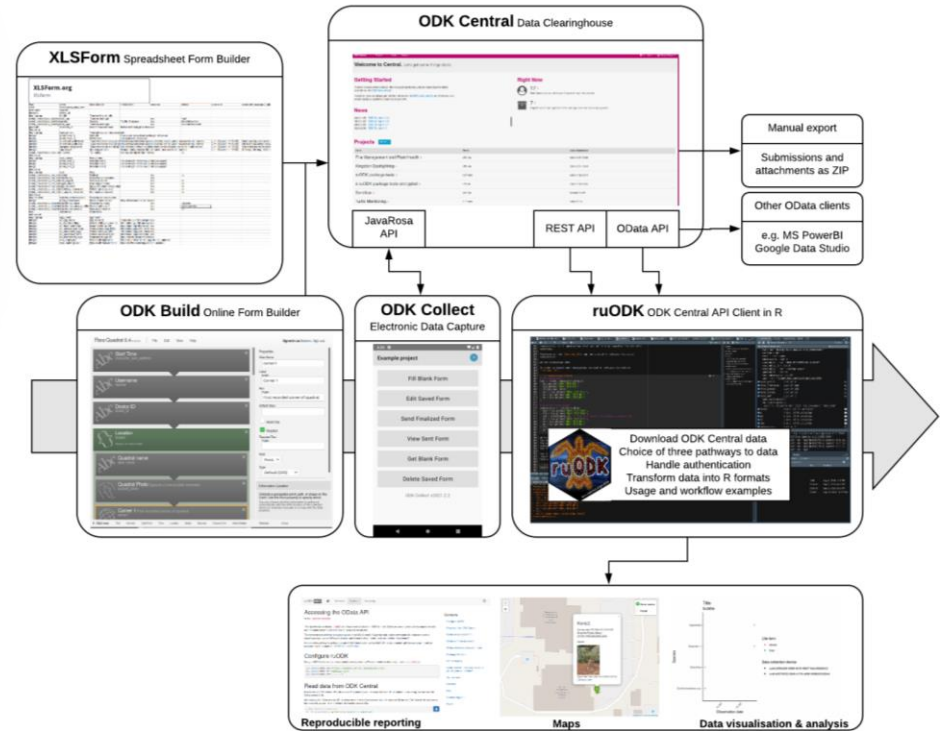
# Achievement over 5 years

- Data is essential for the sustainable cocoa production
- We have developed and operationalize digital data collection workflow: from quality data collection to semantic interoperability as cocoa ontologies
- More than 10,000 data points with images
- Developed methods & technologies are open source and can be applied to other crops
- We must enhance our efforts to ensure the continuity of our data collection system for long-term impact



# What next?

- Ensure continuity in data collection and quality assurance processes
- ODK Aggregate is no longer supported
- Urgent need to migrate to ODK Central
- Provide training to the data collectors



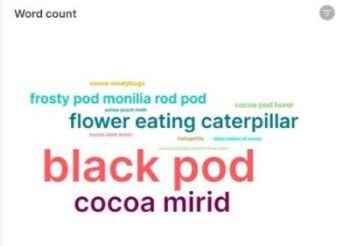
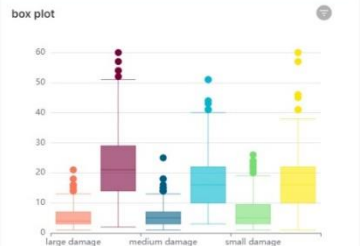
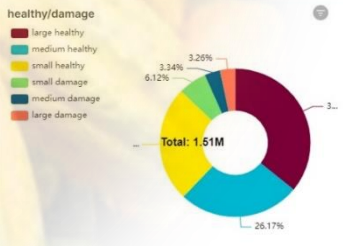


Filters

+ ADD/EDIT FILTERS

Country

4 options



Table

country_id	company	metric	AVG(value)					
			large damage	large healthy	medium damage	medium healthy	small damage	small healthy
CI	Cargill		2.71	5.34	2.61	2.71	4.84	3.17
	Mondelez		3.4	5.69	2.64	2.99	4.01	3.61
	Olam		3.05	6.33	4.12	5.57	5.82	6.41
CM	Olam		2.72	5.49	3.13	5.01	3.57	3.22
	Kuapa Kokoo		2.16	4.74	2.64	5.17	2.93	5.52
GH	Monsieur		3.52	3.63	3.45	5.14	7	4.05

All data

Show All entries

creation_date	country_id	company	site_id	plot_id	tree_id	category	value	type_of_damage
2020-09-11	GH	Rockwinds Transroyal	SQGH096	SQGH096001	SQGH09600101	large healthy	2	N/A
2020-09-11	GH	Rockwinds Transroyal	SQGH096	SQGH096001	SQGH09600101	medium healthy	5	N/A
2020-09-11	GH	Rockwinds Transroyal	SQGH096	SQGH096001	SQGH09600101	small healthy	9	N/A
2020-09-11	GH	Rockwinds Transroyal	SQGH096	SQGH096001	SQGH09600102	large healthy	5	N/A



## Filters



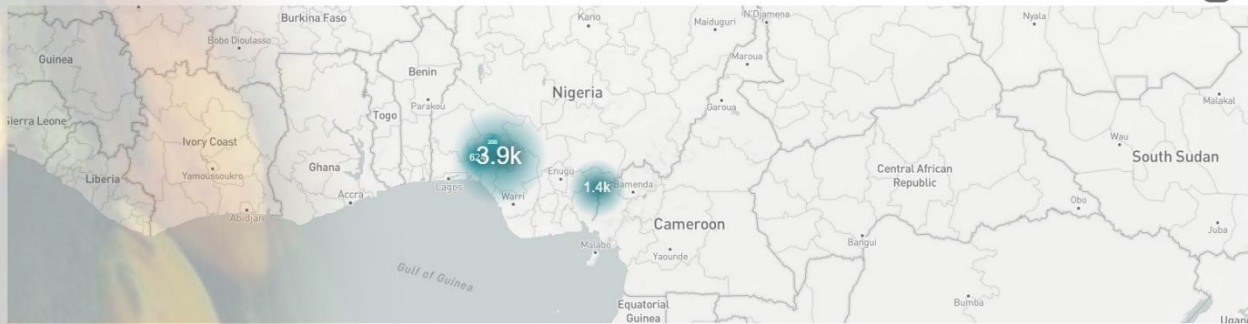
+ ADD/EDIT FILTERS

### Country

NG x



## Trees



## trend



Jan 2022

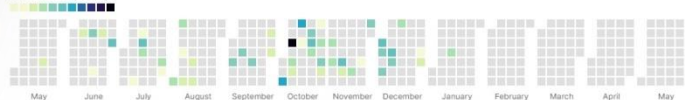
2.18



## Matrix



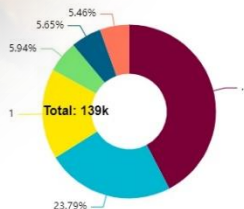
Metric: AVG(value)



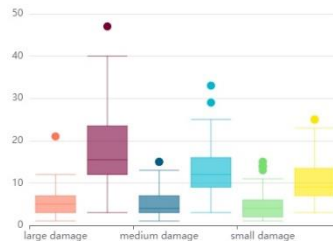
## healthy/damage



- large healthy
- medium healthy
- small healthy
- small damage
- medium damage
- large damage



## box plot



## Word count



## Table



metric	AVG(value)					
category	large	large	medium	medium	small	small

## All data



Show All entries

creation\_date country\_id company site\_id plot\_id tree\_id category value type\_of\_damage

# Linking to EU Deforestation Regulation (EUDR)

Inspiring the use of high-tech in low-tech environments

🏠 ...

## Inspiring the use of high-tech in low-tech environments

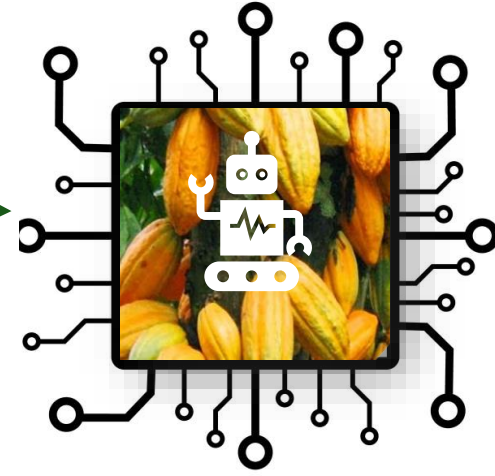
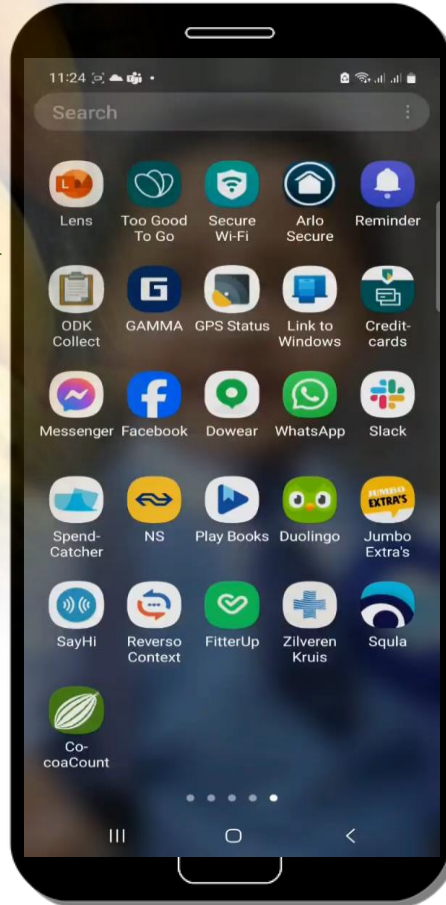
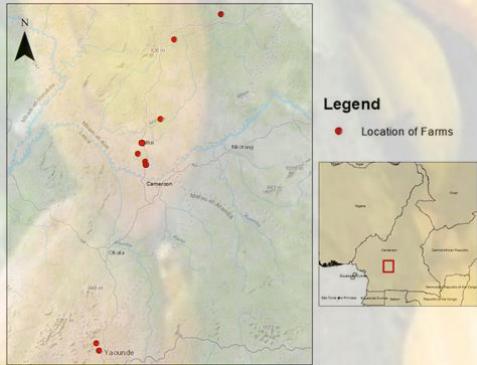
Community management of natural resources using high tech, mobile technology & connectivity solutions in small-holder farming context

Arun Pratihast & Henk Kramer

14 July 2022

<https://arcg.is/0XOjnj>

# Innovation: Farmer platforms for Cocoa pod count



Cocoa pod count



<https://hilo.wur.nl/>

Thank you for your attention !

Feedback:

[arun.pratihast@wur.nl](mailto:arun.pratihast@wur.nl)

