

AGRINER DATASET

- Existing datasets focus on a few entities
- Introducing the **AgriNER** dataset
- Sourced from Top Agricultural Journals
- 36 entities and 9 relations
- 14,307 word-tokens and 1348 entities
- Publicly available on Github
- Scores well with spaCy NER model

AgriNER: An NER Dataset of Agricultural Entities for the Semantic Web

Sayan De¹, Debarshi Kumar Sanyal², Imon Mukherjee¹

¹Indian Institute of Information Technology Kalyani, 741235, India

²Indian Association for the Cultivation of Science, Jadavpur 700032, India

CONTACT

Sayan De
sayan_jrf22@iiitkalyani.ac.in
Debarshi Kumar Sanyal
debarshisanyal@gmail.com
Imon Mukherjee
imon@iiitkalyani.ac.in

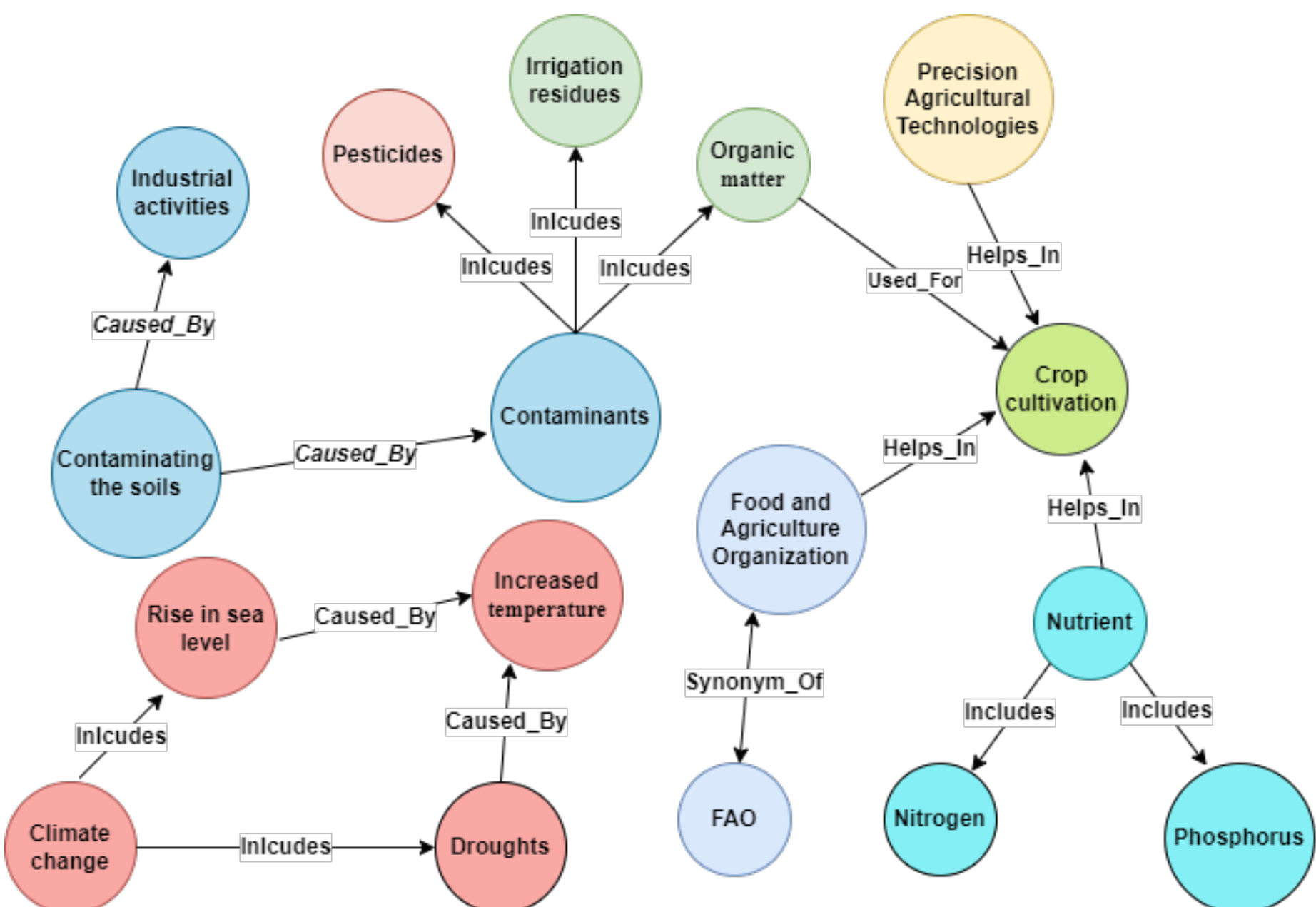
ENTITIES WITH THEIR OCCURRENCES IN AGRINER DATASET

Entity	Frequency	Entity	Frequency	Entity	Frequency
Agri_Method	4%	Other_Quantity	4%	Fruit	3%
Agri_Process	6%	Policy	<1%	Location	15%
Chemical	4%	Rainfall	1%	Money	<1%
Crop	13%	Soil	1%	Natural_Resource	2%
Disease	1%	Temp	1%	Organism	4%
Event	<1%	Vegetable	<1%	Other	<1%
Food_Item	2%	Agri_Ploution	1%	Person	3%
Humidity	<1%	Agri_Waste	2%	Quantity	2%
ML_Model	4	Citation	1%	Season	3%
Natural_Disaster	5	Date_and_Time	2%	Technology	2%
Nutrient	5%	Duration	2%	Treatment	1%
Organization	3%	Field_Area	1%	Weather	<1%

SAMPLE OF CLASSIFICATION REPORT

	Precision	Recall	F1-Score	Support
Agri_Method	1.00	0.80	0.88	5
Chemical	0.98	0.87	0.91	10
Crop	0.93	0.77	0.84	18
Duration	0.83	0.75	0.80	5
Location	0.93	0.95	0.98	19
ML_Model	0.99	0.85	0.92	7
Organism	0.90	0.97	0.94	9
Other_Quantity	0.85	0.97	0.92	6
Season	0.98	0.97	0.98	6

SOME PARTS OF THE KNOWLEDGE GRAPH USING THE DATASET



SCAN THE QR CODE

