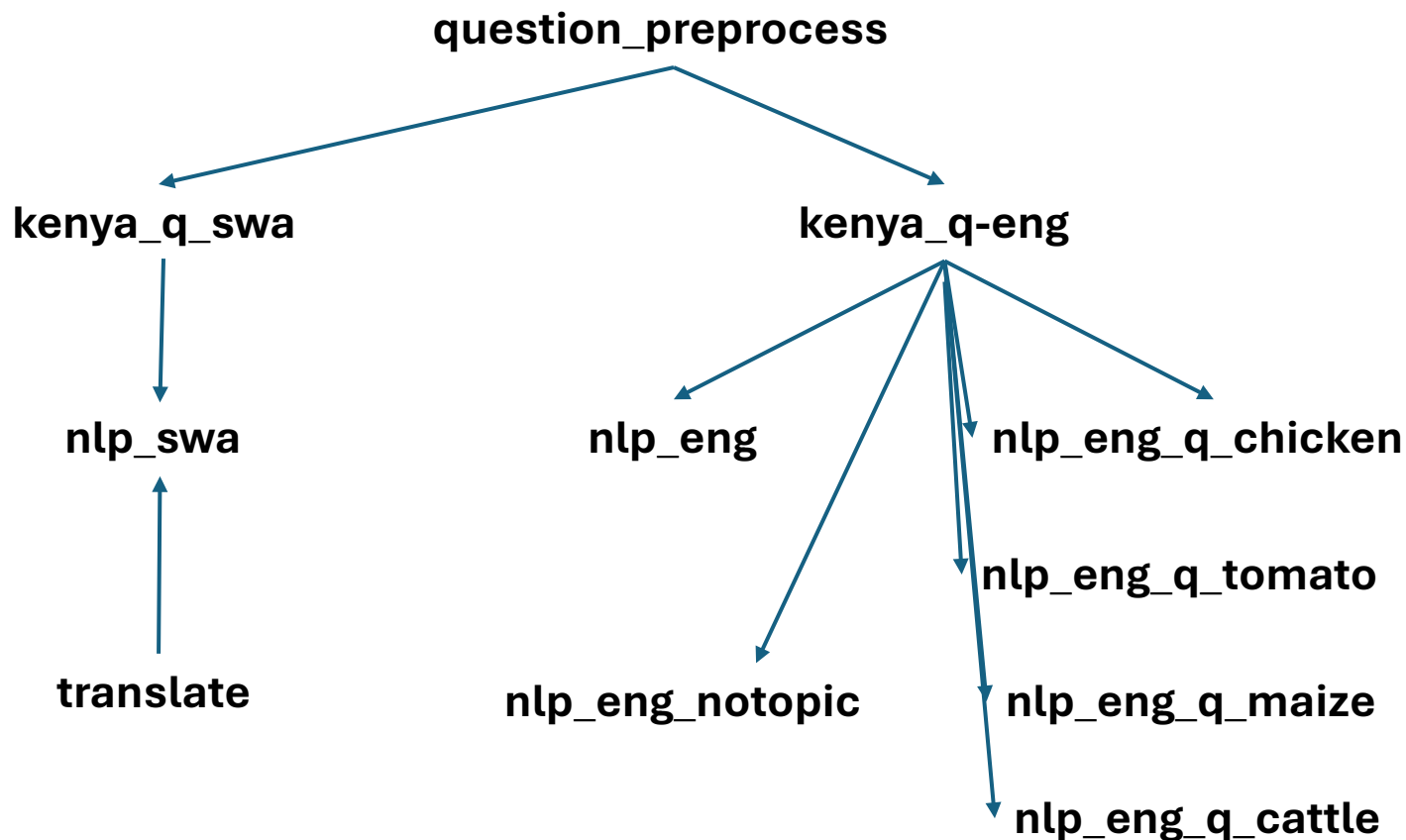


Challenge 4: Crop-Specific and Crop-Independent Questions

Identifying and Visualizing Tri-Grams of English Questions by Kenyan Farmers* in the WeFarm SMS Platform

*'Live' and 'Zombie' users

Jupyter Notebooks Run Order (in Github):



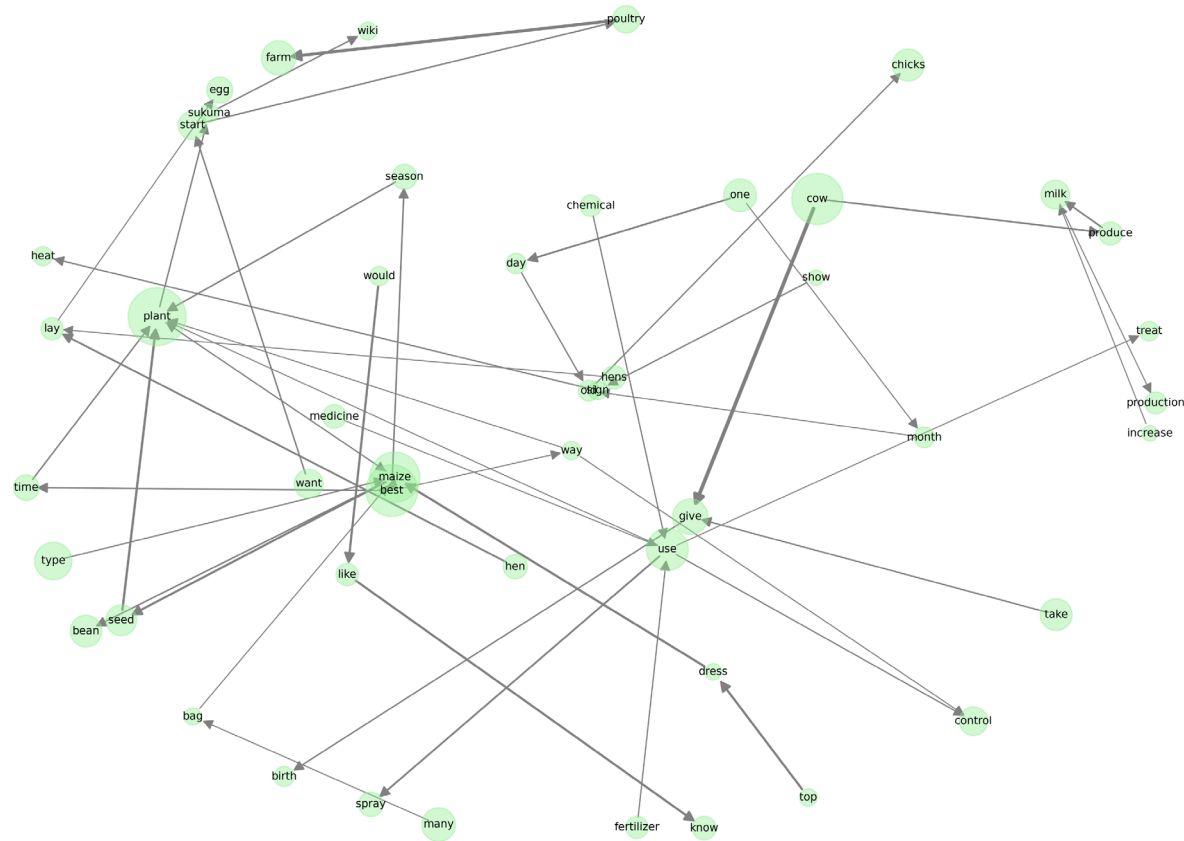
More detail about the steps, inputs and outputs are provided in these notebooks

Swahili trigrams translated in English – all Topics – more informative than bi- or quadgrams and are Similar to their English Counterparts

- cow_give_birth
- start_poultry_farm
- maize_seed_plant
- top_dress_maize (fertilizer)
- type_maize_seed
- best_maize_seed
- chemical_use_spray
- best_way_control
- hen_lay_egg
- cow_produce_milk
- best_season_plant
- plant_sukuma_wiki (collard greens)

Tri-grams of English Questions Without a Topic, Still Focused on Best Practices on Birthing Cattle, Planting Maize and Starting a Poultry Farm

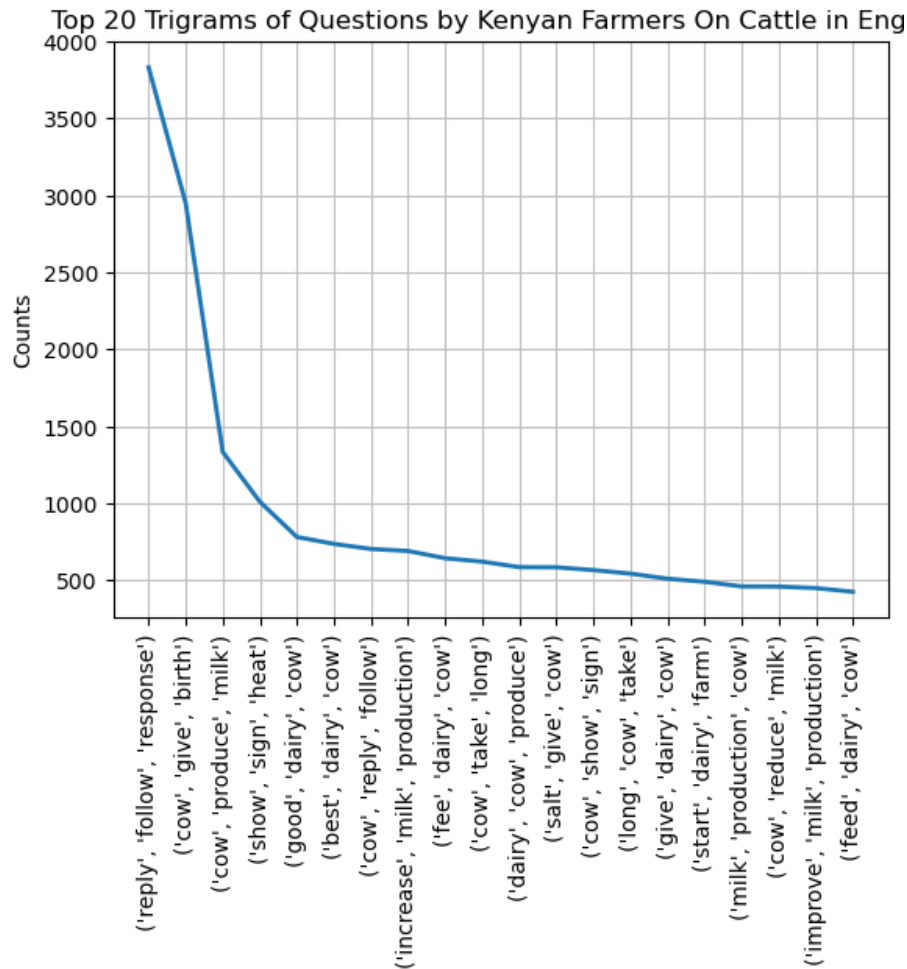
Top 30 Trigrams of Kenyan Farmer English Questions Without Topic (WeFarm, 2022)
(circle size represents word frequency, arrow width represents trigram frequency)

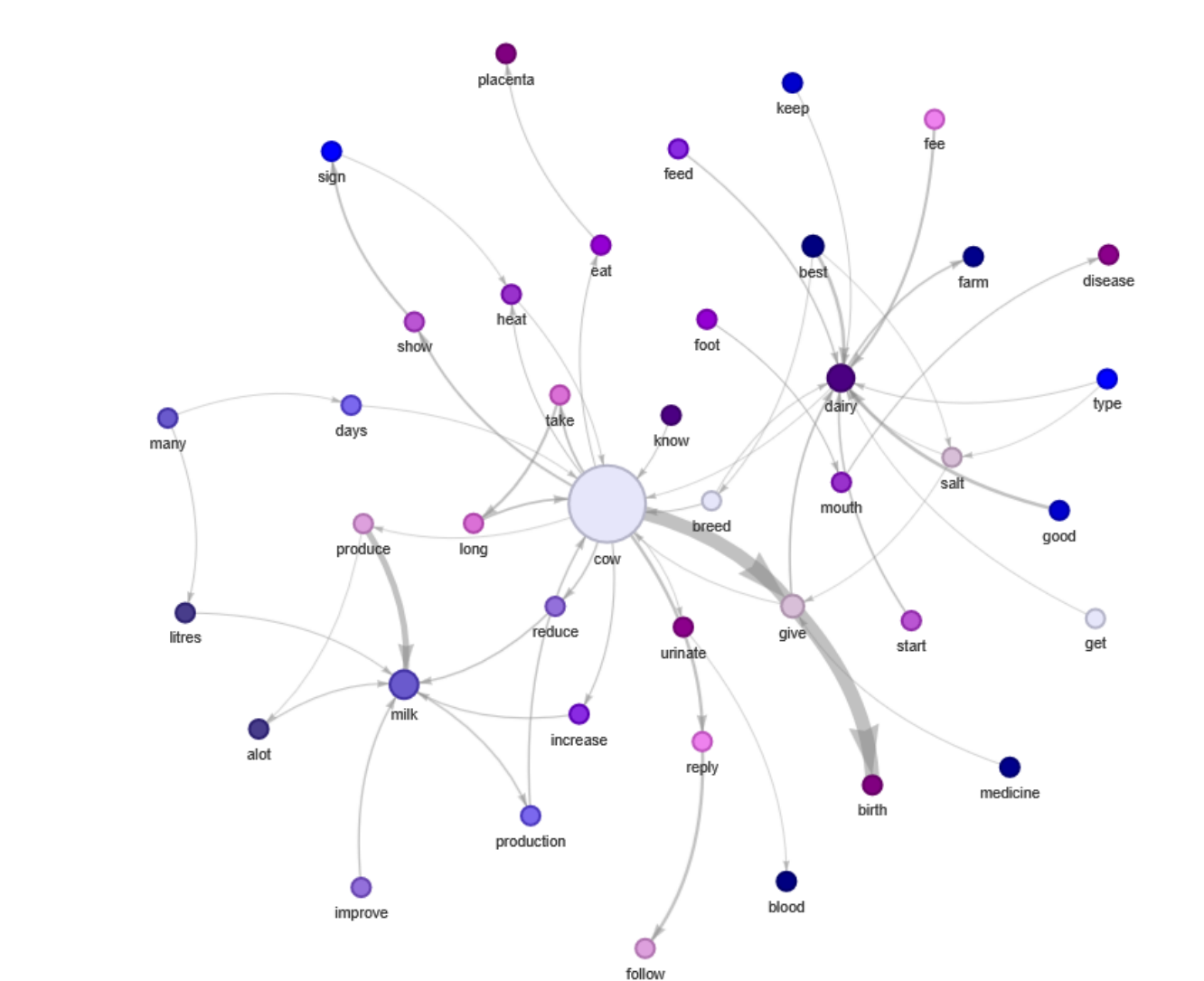


Source: WeFarm 2022 SMS Platform

DataKind Producers Direct Challenge -
December 2025, Beatrice Liu

Cattle Tri-grams focused on Giving Birth, Milk Production and Health

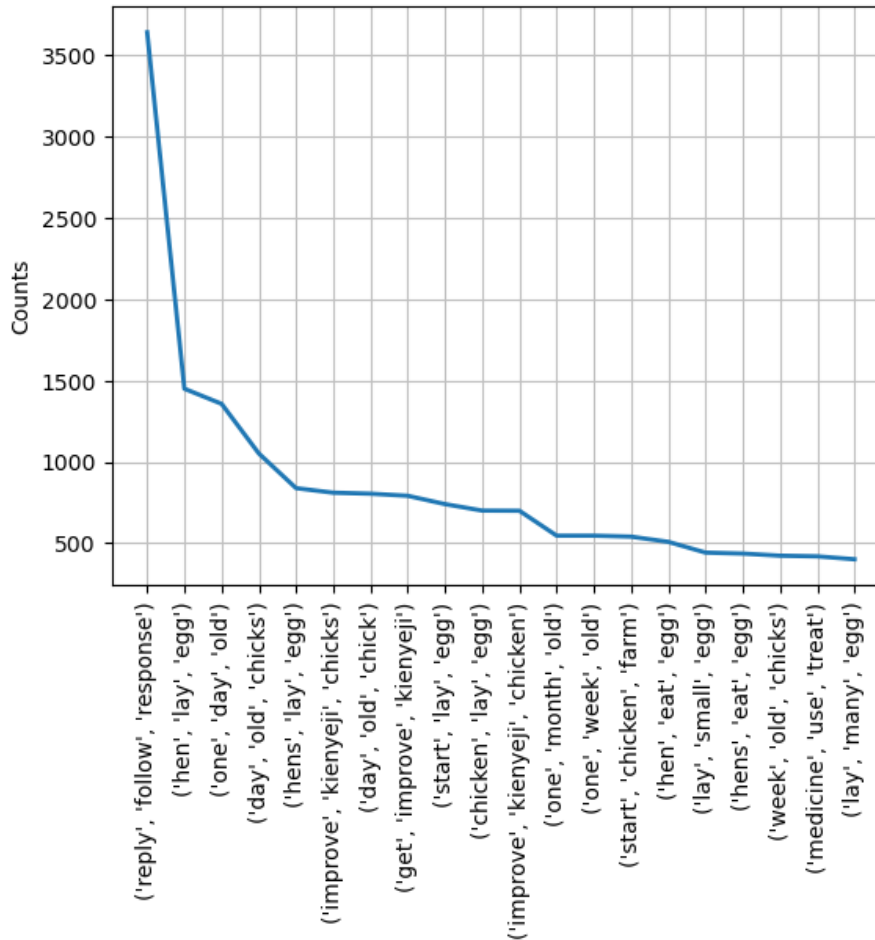




DataKind Producers Direct Challenge -
December 2025, Beatrice Liu

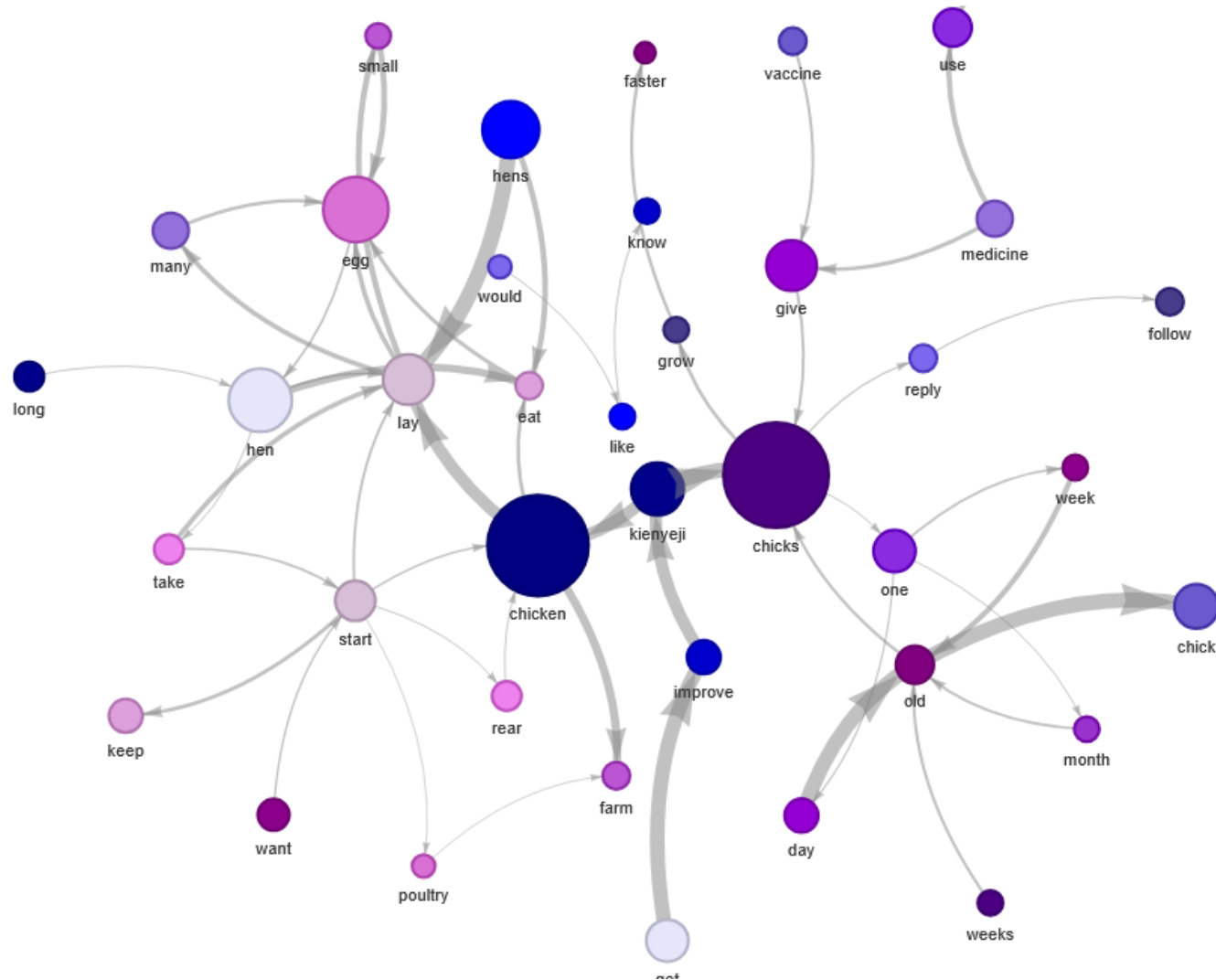
Kenyan Chicken Farms Asked About Young Chicks, Laying Eggs, and Medicines

Top 20 Trigrams of Questions by Kenyan Farmers On Chicken in English



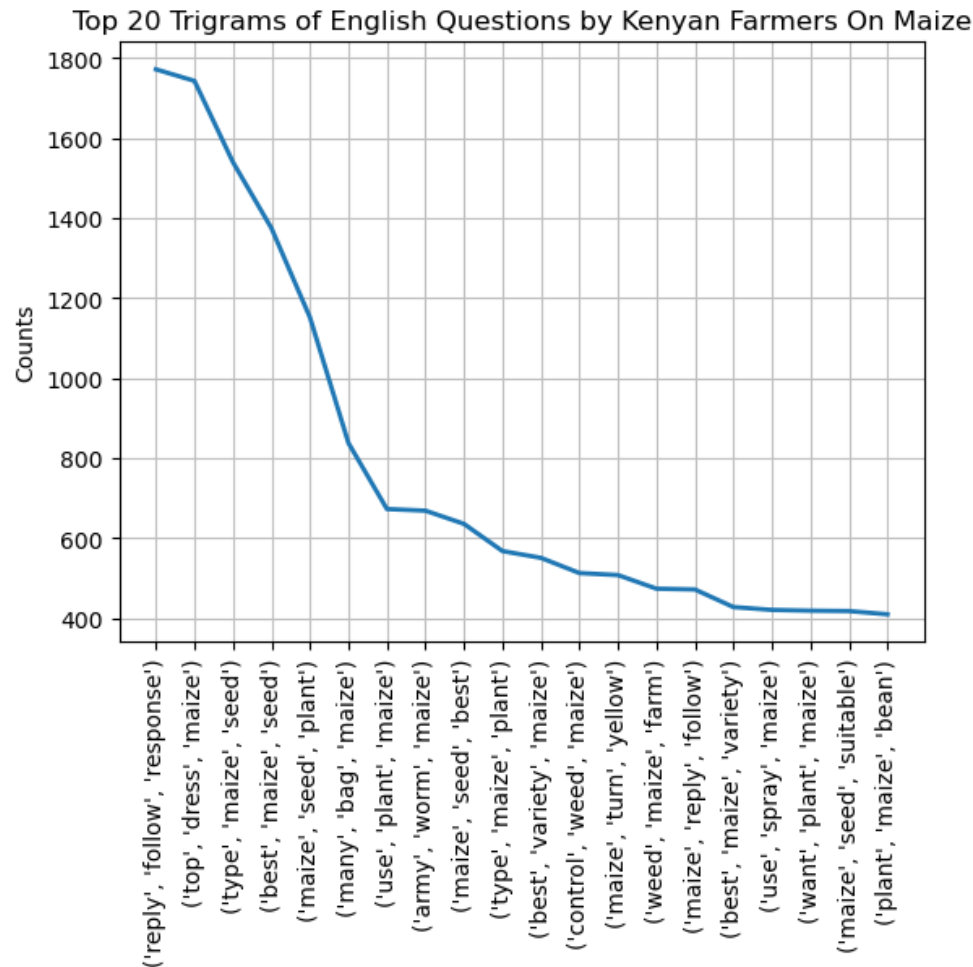
Network Graph: Top 40 Trigrams from Kenyan Farmers Questions on Chicken in English

Word circle size = word frequency, arrow width = trigram frequency, Source: WeFarm 2022 SMS Platform



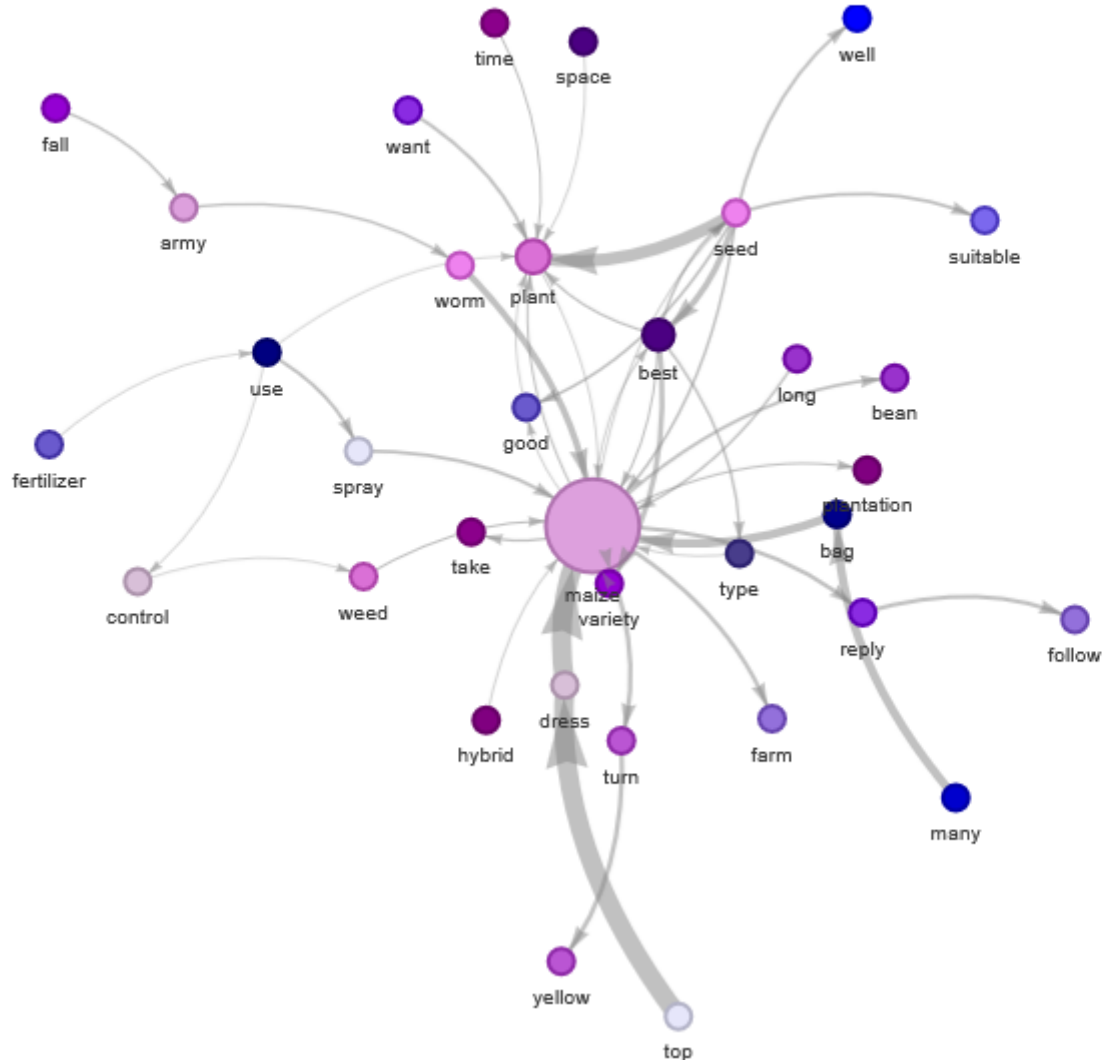
DataKind Producers Direct Challenge -
December 2025, Beatrice Liu

Maize Farmers in Kenya Asked About Best Fertilizer and Weed Control, Best Seeds and Planting



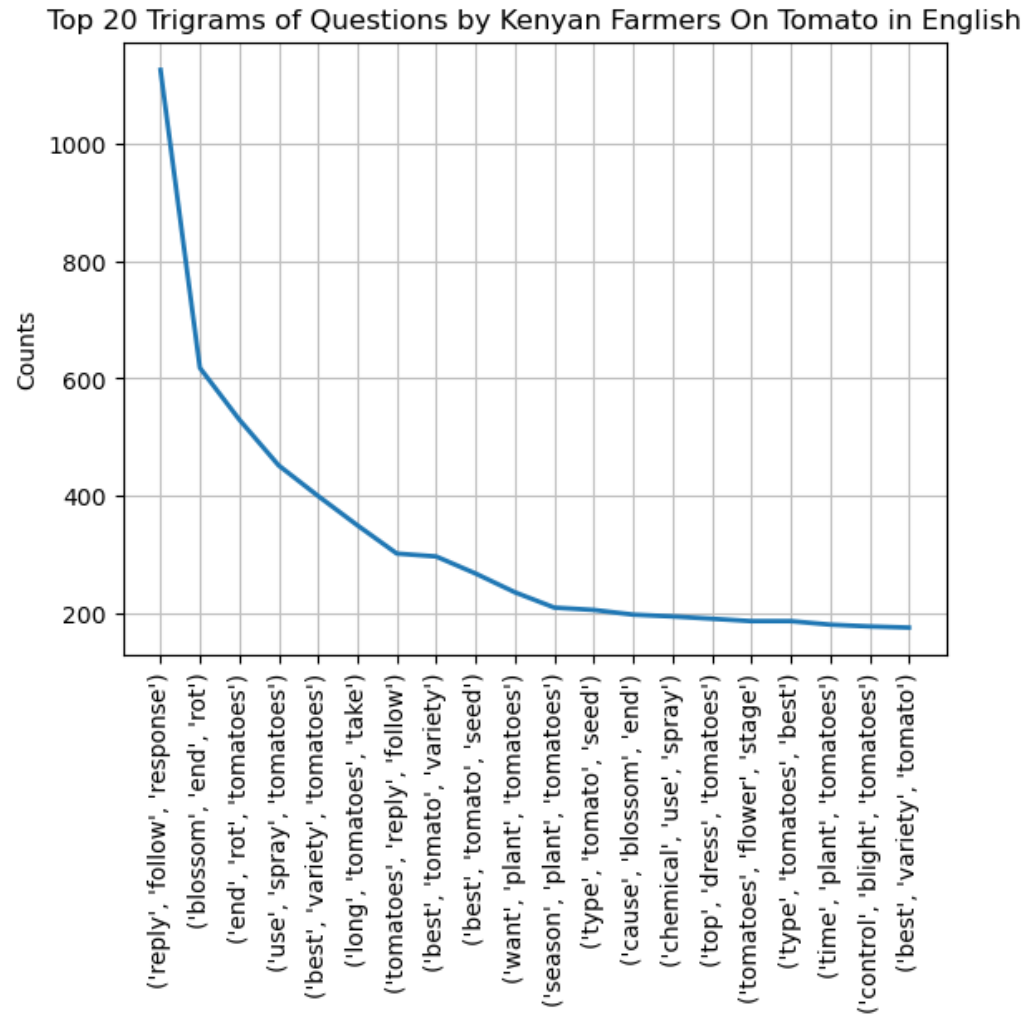
Network Graph: Top 40 Trigrams from Kenyan Farmer English Questions on Maize

Word circle size = word frequency, arrow width = trigram frequency, Source: WeFarm 2022 SMS Platform



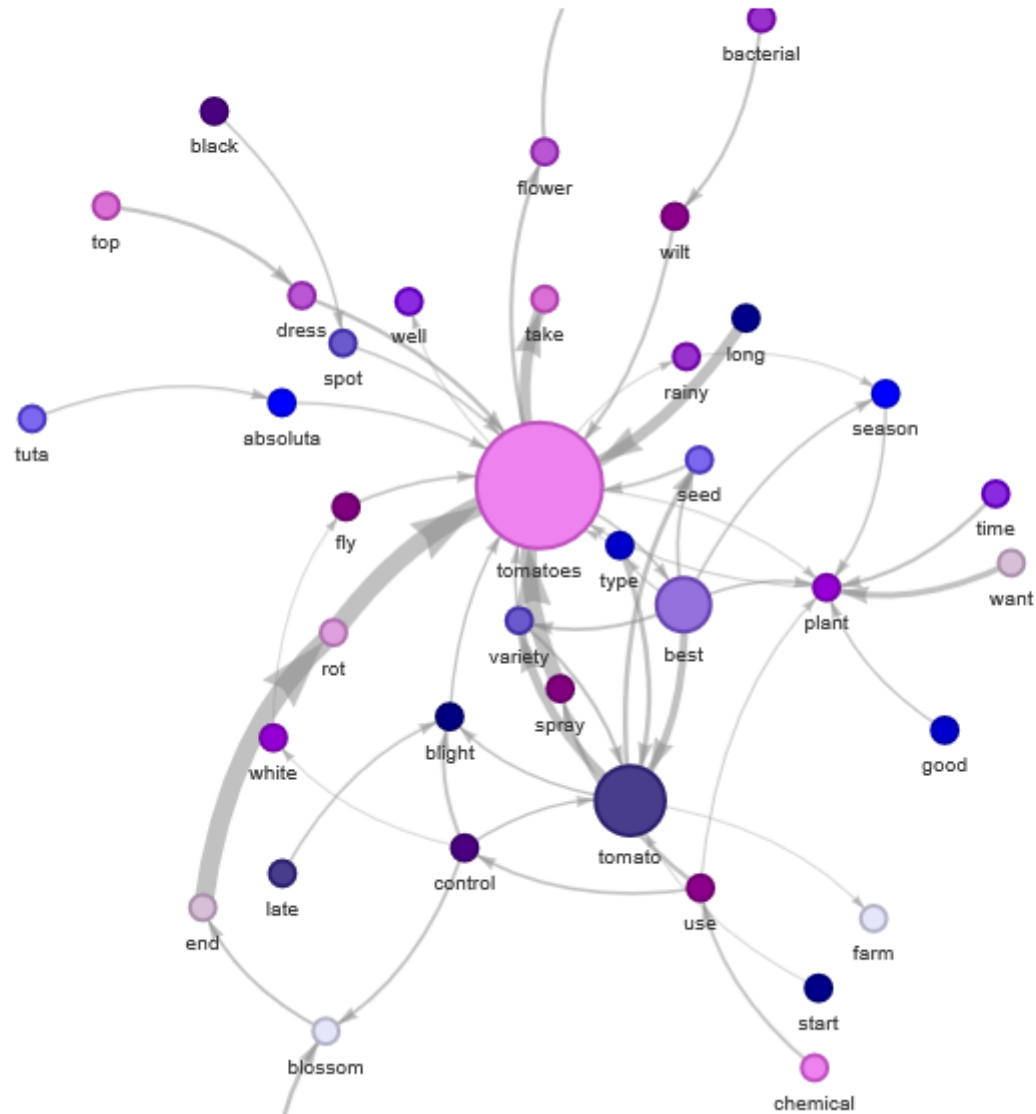
DataKind Producers Direct Challenge -
December 2025, Beatrice Liu

Tomato Farmers Asked About Planting, Best Seeds, and Keeping Them Disease-Free



Network Graph: Top 40 Trigrams from Kenyan Farmers Questions on Tomato in English

Word circle size = word frequency, arrow width = trigram frequency, Source: WeFarm 2022 SMS Platform



December 2025, Beatrice Liu

Additional Resources

- Challenge data files and visualizations in Google Drive:
https://drive.google.com/drive/folders/1tpwqTqoFfZCWvDvncJjaSbzzua0Y6Q_i?usp=sharing
- Common Swahili Stop-Words;
<https://data.mendeley.com/datasets/mmf4hnsn2n/1>
- Swahili Agriculture Corpus: KILIMO:
<https://data.mendeley.com/datasets/d4yhn5b9n6/2/files/cfd0108d-863d-460d-b52c-a51ce4101f79>
- Swahili Verb Conjugation Dataset for lemmatization:
<https://data.mendeley.com/datasets/rvt89578g5/1>