# Introduction

Large-scale agriculture faces an intensifying set of challenges that must be overcome to maintain food security in an environmentally sustainable manner. While global food demand is projected to increase by 30-62% by 2050 with output largely keeping pace (van Dijk et al., 2021), the required level of agricultural production intensification demands increasing amounts of land (Alexander et al., 2015; Bahar et al., 2020) and inputs (Cordell et al., 2009; Lu & Tian, 2017) and is projected to cause increasingly severe ecological, environmental, and climate effects (Springmann et al., 2018). Producers must rapidly transition to sustainable production which will require ample farmer-led on-farm experimentation (Lacoste et al., 2022). Farmers often rely on limited data (e.g., average yield) to interpret experimentation results and make important management decisions and additional data may help to reduce misinterpretation of results and thus accelerate change (Cook et al., 2018). An important aspect of on-farm data interpretation is contextualization, without which a piece of data is only anecdotal. Adding layers of biophysical and environmental data can provide objective observations that can help with better understanding the conditions in which farmers are experimenting if they are meaningfully combined and analyzed. Digital agronomy can facilitate this by leveraging advancements in data science, remote sensing, engineering, and agronomy in support of farmers' self-designed experimentation efforts when determining how to best proceed. This requires large volumes of agronomic, biophysical, and environmental data to contextualize the farmers' observations and insights during their daily operations through the conclusion and interpretation of their experiments. This paper details the development of a software system designed to facilitate the retrieval, management, and use of these data in support of a novel observational research methodology applied to endogenous on-farm experiments.

This project developed a similar tool to facilitate the contextualization of observational research of farmer-led on-farm experimentation (FLOFE) by leveraging the rich set of agronomically-relevant environmental and biophysical data products available on the Google Earth Engine (GEE) (Gorelick, et al., 2017) platform. Knowing some of the drivers of the underlying site-level processes that influence experimental outcomes can help to account for previously unexplained uncertainty and spatial heterogeneity present in results. These additional layers of data can provide valuable insights into trends and patterns that may otherwise go undetected when examining FLOFE results, which can be further explored via more traditional agronomic research techniques. This two-phase research methodology has the potential to accelerate the identification of novel solutions to overcome complex challenges faced in modern agronomic systems.

Leveraging the additional layers of contextualization data in on-farm research networks can allow the comparison of previously disparate experiments, transforming them into a comprehensive body of work, increasing the scale of knowledge from farm to regional and super-regional levels. These robust datasets can then be used by various data-driven analysis methods to further elucidate drivers of agronomic processes, model outcomes given certain sets of parameters, and make trial A conducted at one location comparable to trial B conducted at another location.

This system automates the process of retrieving layers of data necessary for FLOFE contextualization. A user-friendly web interface allows users having little to no programming expertise to leverage several agronomically-relevant data products in their research efforts. Automating data retrieval greatly reduces the time and labor costs associated with retrieving large amounts spatiotemporal contextualization data. Faster and easier access means that these additional contextualization data layers can be used in more studies.

# Data and Methods

## Data

The system facilitates access to a variety of agronomically relevant data products via a web portal graphical user interface (GUI). This submits requests to a Python API backend system that retrieves the requested data according to the user-specified parameters. Table 1 lists currently available data products, their spatial and temporal resolutions, and data types.

*Table 1. Current listing of data products available within the system.*

| | **Resolution** | | |
|---|---|---|---|
| **Product** | **Spatial** | **Temporal** | **Data Type** |
| Gridded Surface Meteorological Dataset (GRIDMET) | 1/24$^{th}$ °; ~0. 4 km | daily | meteorological |
| POLARIS Probabilstic Soil Properties | 30 m | varies | soil |
| Real-Time Mesoscale Analysis (RTMA) | 2.5 km | hourly | meteorological |
| US Drought Monitor | 250 m | weekly | index |
| Sentinel 2 Surface Reflectance (L2A), Harmonized | 10 m, 20 m, 60 m | ~ 5 day | multispectral |
| USGS 3DEP Elevation 1 m | 1 m | varies | soil |
| USGS 3DEP Elevation 10 m | 10 m | varies | soil |

Product-specific and user-defined metadata are stored in a PostGRE/PostGIS relational database to facilitate data management and using data to contextualize on-farm experimentation. Backend metadata tables were structured to facilitate adding additional metadata categories as necessary. Table 2 lists metadata currently collected for each data download.

*Table 2. Metadata currently collected by the system per download request*

| **Metadata** | **Description** |
|---|---|
| Data Product Type Id | The id of the data product type of a given piece of data |
| DownloadBatchId | GUID corresponding to the data request action initiated from the web portal |
| Farm Name | The name of a farm to be associated with some data |
| Field Name | The name of a field to be associated with some data |
| Project Data End Year | The year for which data for a given project were collected to, inclusive |
| Project Data Start Year | The year for which data began being collected for a given project |
| Project Name | The name of a project for which data are stored in the system. |
| Spatial Data Type | The type of spatial data, ex Raster or Vector |
| Vector Data Format | The format of vector data, ex: shp, gpkg, geojson |
| VectorDataDescribes | What a vector file describes, ex: field boundary, transect line |

** denotes optional metadata*

# System

## Overview

The system has three main components (i) a web-based user interface and (ii) an application programming interface (API) backend. A relational PostGRE/GIS database in conjunction with a server file system were used as the backend data store for all system and user data, data products, and associated metadata. A unified authentication and authorization module was developed to secure all system components. A high-level system overview is show in figure 1.

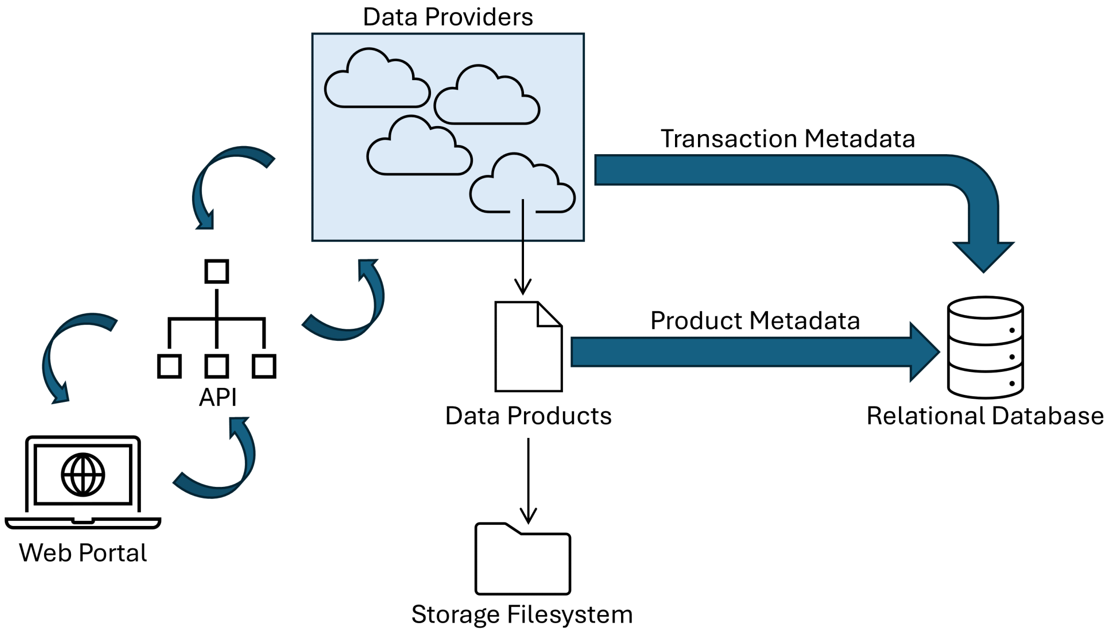


*Figure 1. System Overview*

## Authentication and Authorization

System data security and privacy were achieved through the user of a user and role-based authentication and authorization framework. Each user must provide a Google Cloud Service Account authentication token before they can initiate data requests. This ensures that it is the user's responsibility to adhere to any GEE (or other future data provider) data licensing requirements.

## Web Portal

The web portal component of the system provides an intuitive, responsive user interface (UI) to initiate data transfers. It was developed using the .NET Core Blazer Server framework in C#.  Users enter required information, such as their email address and Google service account information via a profile module. Users can search for supported data products via the data product search form after entering all required profile information. Figure 2 shows the data search form for a request example. Available search criteria are specific to each data product

## API

The API component of the system performs all search, geospatial, storage, and retrieval operations. It was developed using Python 3.11. An authentication and authorization module is present to restrict API access to requests originating from web portal-authenticated users having sufficient access privileges. Upon passing authentication and authorization checks, the user's download request is processed. Users are emailed a link to download requested data upon completion of a request. Downloaded data are indexed in the PostGRE/GIS database, storing descriptive and user-specified metadata.

## PostGRE/GIS Database

A PostGRE/GIS relational database serves as the backend data store for system configuration, user profiles, and downloaded product metadata. The GIS component of this database allows for a variety of spatial data manipulation, such as on-the-fly reprojection and allows the system to retrieve data via spatial searches. Future expansion is also facilitated by this database supporting the streaming of cloud-stored raster data to a variety of providers, which will likely be necessary as the system grows.

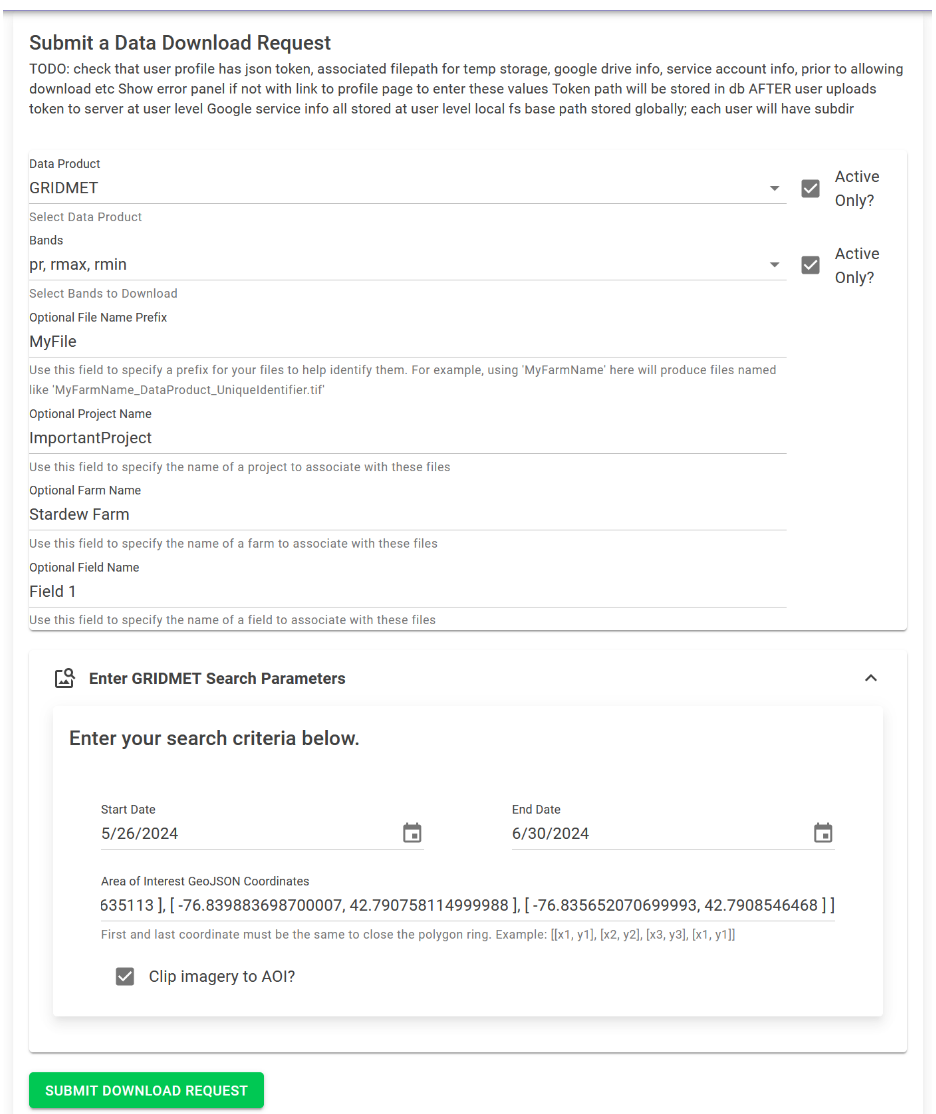


*Figure 2. Web portal data product download form.*

# Discussion

The system facilitates data retrieval in support of on-farm experimentation. The system was used to retrieve contextualization data for a 22 site-year study and greatly simplified retrieving large amounts of biophysical and environmental data that were used to provide context to the farmer-led experiments comprising this study. Future work will include UI enhancements, such as the ability to upload or draw area of interest polygons via an interactive map and an analytics module to automate many of the common analysis techniques used in evaluating on-farm data. The system is currently in late alpha stage with plans for a limited beta release later this summer.

# References

Alexander, P., Rounsevell, M. D. A., Dislich, C., Dodson, J. R., Engström, K., & Moran, D. (2015). Drivers for global agricultural land use change: The nexus of diet, population, yield and bioenergy. *Global Environmental Change*, *35*, 138–147. https://doi.org/10.1016/j.gloenvcha.2015.08.011

Bahar, N., Lo, M., Sanjaya, M., Van Vianen, J., Alexander, P., Ickowitz, A., & Sunderland, T. (2020). Meeting the food security challenge for nine billion people in 2050: What impact on forests? *Global Environmental Change*, *62*, 102056. https://doi.org/10.1016/j.gloenvcha.2020.102056

Cook, S., Lacoste, M., Evans, F., Ridout, M., Gibberd, M., & Oberthür, T. (2018). An On-Farm experimental philosophy for farmer-centric digital innovation. *Cook, S. <https://Researchrepository.Murdoch.Edu.Au/View/Author/Cook, Simon.Html>, Lacoste, M., Evans, F. <https://Researchrepository.Murdoch.Edu.Au/View/Author/Evans, Fiona.html>ORCID: 0000-0002-7329-1289 <http://Orcid.Org/0000-0002-7329-1289>, Ridout, M., Gibberd, M. and Oberthür, T. (2018) An On-Farm Experimental Philosophy for Farmer-Centric Digital Innovation. In: 14th International Conference on Precision Agriculture, 24 - 27 June 2018, Montreal, Quebec*. 14th International Conference on Precision Agriculture, Montreal, Quebec. https://researchrepository.murdoch.edu.au/id/eprint/56711/

Cordell, D., Drangert, J.-O., & White, S. (2009). The story of phosphorus: Global food security and food for thought. *Global Environmental Change*, *19*(2), 292–305. https://doi.org/10.1016/j.gloenvcha.2008.10.009

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment*, *202*, 18–27. https://doi.org/10.1016/j.rse.2017.06.031

Lacoste, M., Cook, S., McNee, M., Gale, D., Ingram, J., Bellon-Maurel, V., MacMillan, T., Sylvester-Bradley, R., Kindred, D., Bramley, R., Tremblay, N., Longchamps, L., Thompson, L., Ruiz, J., García, F. O., Maxwell, B., Griffin, T., Oberthür, T., Huyghe, C., … Hall, A. (2022). On-Farm Experimentation to transform global agriculture. *Nature Food*, *3*(1), Article 1. https://doi.org/10.1038/s43016-021-00424-4

Lu, C., & Tian, H. (2017). Global nitrogen and phosphorus fertilizer use for agriculture production in the past half century: Shifted hot spots and nutrient imbalance. *Earth System Science Data*, *9*(1), 181–192. https://doi.org/10.5194/essd-9-181-2017

Springmann, M., Clark, M., Mason-D'Croz, D., Wiebe, K., Bodirsky, B. L., Lassaletta, L., de Vries, W., Vermeulen, S. J., Herrero, M., Carlson, K. M., Jonell, M., Troell, M., DeClerck, F., Gordon, L. J., Zurayk, R., Scarborough, P., Rayner, M., Loken, B., Fanzo, J., … Willett, W. (2018). Options for keeping the food system within environmental limits. *Nature*, *562*(7728), Article 7728. https://doi.org/10.1038/s41586-018-0594-0

van Dijk, M., Morley, T., Rau, M. L., & Saghai, Y. (2021). A meta-analysis of projected global food demand and population at risk of hunger for the period 2010–2050. *Nature Food*, *2*(7), Article 7. https://doi.org/10.1038/s43016-021-00322-9