SOCKG: Soil Organic Carbon KG Team

Chengkai Li, The University of Texas at Arlington; Jianzhong Su, Virginia Lin, Timothy Propst

Use case description and societal challenge being addressed

In this project, UTA researchers collaborate with USDA to develop a soil organic carbon knowledge graph (SOCKG), that supports robust soil carbon modeling for voluntary carbon markets. Carbon market provides a mechanism to incentivize climate-friendly practices that help reduce carbon in the air by keeping more carbon in soil. A market needs a currency. The currency for carbon markets is carbon credit. One carbon credit corresponds to the reduction of emission of a metric ton of carbon dioxide or carbon dioxide equivalent. Industry sectors such as energy and transportation purchase carbon credits to offset their unavoidable carbon emissions. Farmers and land managers receive rewards for adopting agricultural practices that increase soil organic carbon. For this currency to be robust and for the carbon market to be effective, it is crucial to have soil carbon modeling technologies to accurately measure soil organic carbon (SOC) content and predict SOC content change and attribute that change to agricultural practices. The key to the soil carbon modeling technology is high quality data. That's why we are embarking on the soil organic carbon knowledge graph. By bringing together existing data that are in silos to form a knowledge graph, connecting the KG with data in broader contexts, and setting up infrastructure for sustaining the KG, we are creating an important public open data resource that will play a pivotal role in enabling the nascent and fast-growing carbon market.

Knowledge graph source datasets

1) GRACEnet (Greenhouse gas Reduction through Agricultural Carbon Enhancement network). Excel data files. Each Excel file contains the data of one experiment site, including experiment description, characterization of the weather, management of the crop fields, and measurement of physical, biological, and chemical soil properties. 2) Other USDA field study data, such as those from the Nutrient Uptake and Outcome Network (NUOnet) and LTAR (Long-Term Agricultural Research). 3) National Agricultural Library Thesaurus (NALT) and Glossary. 4) Other soil carbon datasets such as the International Soil Carbon Network (ISCN) Database and the Natural Resources Conservation Service (NRCS) Soil Survey Geographic (SSURGO) database. 5) Other datasets such as KnowWhereGraph and Wikidata.

User queries / competency queries for the use case

1) What is the delta SOC change over time for each plot/replicate. 2) Which management treatment results in the greatest amount of SOC storage? (This can be answered by averaging plots/replicates and comparing management treatments.) 3) What management combinations result in the greatest amount of SOC storage? 4) Is there a relationship between crop productivity (yield) and SOC storage? Does that relationship change with crop type, management, or geographic location? 5) Which soil responses are most sensitive to management treatments? Questions 1-4 could also be applied to any soil measurement (pH/EC, bulk density, N content, GHG emissions, etc).