Visualizing data interoperability for food systems sustainability research—from spider webs to neural networks

Emily Jennings-Dobbs, Shavawn Forester, Adam Drewnowski

Key Insights

> **Interoperability is the key...**

Research on sustainable food systems is transdisciplinary, relying on the interconnected domains of health, nutrition, economics, society, and environment. However, crosswalks among largely siloed data are not fully developed. Our visual demonstration serves as proof of concept, identifying databases in need of expansion, integration, and harmonization for use by researchers, policymakers, and the private sector.

Defining Interoperability

"Studies on sustainable, healthy diets would benefit from greater interoperability, defined here as expanding, integrating, and harmonizing the existing data into a common modeling framework."

The current state of interoperability between data in the 4 sustainability domains, was assessed by the existence of:
- **Crosswalks** - the connection of two data sources through a common set of identifiers or indices, which can be used to link, merge, or join the data
- **Ontologies** - common terminologies within a domain, providing an organizational structure that can be used across data sources

Challenges Limiting Interoperability

"Linking data sets from multiple domains can be challenging, and there is a scarcity of tools to integrate disparate data sets into a common analytical framework"
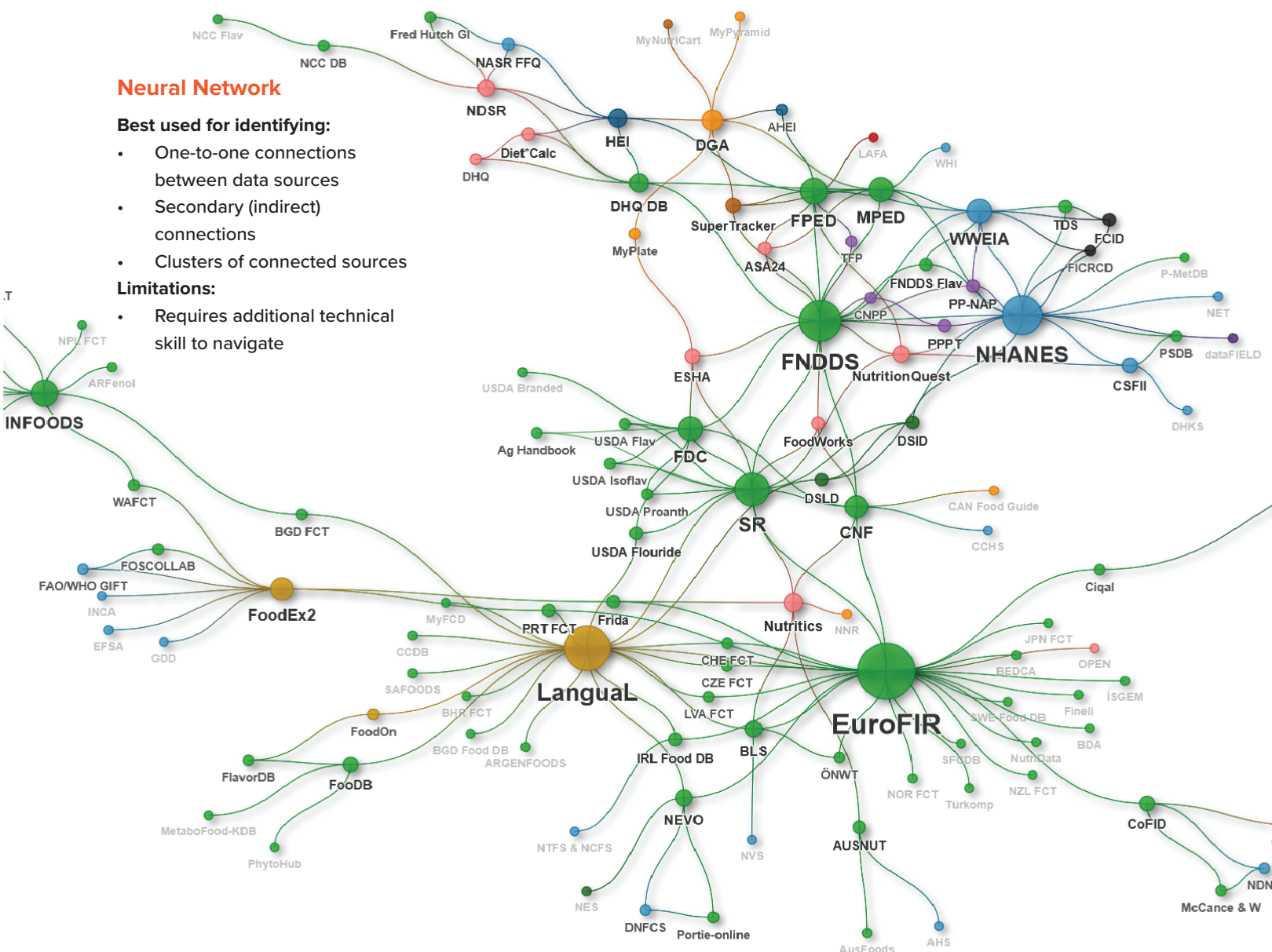
Linking agricultural, economic, and environmental data with nutrition and health databases is a major challenge because the data are:
- collected by **different agencies**
- provided to potential users in **different formats**
- stored at **different locations**

There are hundreds of different food systems data sources - each with it's own unique connections. To identify gaps in interoperability and determine the most effective path forward in interdisciplinary research, we must turn to **visualization**.

Visualizing Food Sustainability Data

"Research in the nutritional sciences may not utilize the full extent of data visualization, which is a valuable tool to transform raw data into something usable and understandable that can provide insight beyond the scope of traditional descriptive statistics."

Neural Network

Best used for identifying:
- One-to-one connections between data sources
- Secondary (indirect) connections
- Clusters of connected sources

Limitations:
- Requires additional technical skill to navigate

The neural network visualization has a dual capability of highlighting disconnectedness and enabling users to navigate and harmonize complex food systems data. This offers a transformative approach in advancing research in this critical domain.